Science Networks
Historical Studies
45

Arie Hinkis

Proofs of the Cantor-Bernstein Theorem

A Mathematical Excursion

Birkhäuser

Science Networks. Historical Studies

Birkhäuser

Science Networks. Historical Studies
Founded by Erwin Hiebert and Hans Wußing
Volume 45

Edited by Eberhard Knobloch, Helge Kragh and Volker Remmert

Arie Hinkis

Proofs of the Cantor-Bernstein Theorem

A Mathematical Excursion

Birkhäuser

Arie Hinkis
The Cohn Institute for the History and
Philosophy of Science and Ideas
Tel Aviv University
Tel Aviv
Israel
arik@hinkis.org

ISBN 978-3-0348-0223-9 ISBN 978-3-0348-0224-6 (eBook)
DOI 10.1007/978-3-0348-0224-6
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012949083

Cover illustration: From Waller Ms de-00215, August Beer: *Über die Correction des Cosinusgesetzes bei der Anwendung des Nicol'schen Prismas in der Photometrie*, after 1850. Used with the kind permission of The Waller Manuscript Collection (part of the Uppsala University Library Collections).

Printed on acid-free paper

Springer Basel is part of Springer Science+Business Media (www.birkhauser-science.com)

To the memory of my grandparents, Haya and Haim Erlichman,
Murdered in Transnistria, with their children and grandson,
And to the memory of Adolf Lindenbaum, A delicate genius,
murdered in Ponary.
הי"ד

With transcendent love for my children and grandchildren!

"מִשְׁפָּטֶיךָ שִׁוִּיתִי" (תהלים קיט 30)

The researcher, by Tali Hinkis-Lapidus

Introduction

The chief purpose of this book is to present, in detail, a compilation of proofs of the Cantor-Bernstein Theorem (CBT) published through the years since the 1870s. Over 30 such proofs are surveyed.

CBT, an elementary but not trivial theorem of set theory, is usually stated either in its single-set formulation (if $M'' \subseteq M' \subseteq M \sim M''$ then $M \sim M'$) or in its two-set formulation (if $M \sim N' \subseteq N$ and $N \sim M' \subseteq M$ then $M \sim N$), which are equivalent.

At the turn of the twentieth century, CBT was called the Equivalence Theorem. The veterans, e.g., Bernstein, Zermelo, Hausdorff and Fraenkel, kept using this name. Bernstein (1905 p 121) says that the name was suggested by Cantor.

In Whitehead's 1902 paper, in section III written by Russell, CBT is referenced as "the Bernstein's and Schröder's theorem" (p 369, 383) after the two mathematicians who first published proofs for the theorem, both in 1898. In 1907b (p 355), Jourdain suggested that the name be 'the Schröder-Bernstein theorem'. He thought that Cantor's name should not be included in the name of the theorem because Cantor did not provide a proof of it, a mistake that has been often repeated. As it turned out, the name of Schröder should have been omitted (cf. Tait 2005) because it was found (Korselt 1911) that his proof is erroneous (or rather, as will be demonstrated, senseless). The omission already appears in Poincaré 1906a (p 27). Still, many kept using the Schröder-Bernstein theorem name for CBT. Zermelo mentioned both proofs of Schröder and Bernstein in 1932 (Cantor 1932 p 209 [6]), without any reservation, even noting that Schröder's proof dates to 1896 and that of Bernstein to 1897. Likewise, in the second edition of PM (Principia Mathematica), from 1925, the theorem is still called the Schröder-Bernstein theorem. Similar naming is used in J. Kőnig 1906, Banach 1924, Lindenbaum-Tarski 1926, Whittaker 1927, Hellmann 1961, and even Grattan-Guinness (1971a p 117 reference to #37 in the bibliography). In fact, Schröder's name remains linked to CBT even today, as a quick search of the Internet reveals.

We have not identified when the name 'the Cantor-Bernstein Theorem' was introduced. It seems to be the preferred name today. In the second edition of Fraenkel 1973, edited by A. Levy, the editor systematically changed the original

'Equivalence Theorem' of the 1958 edition, to the Cantor-Bernstein Theorem. For reasons to be detailed later it appears that the proper name for the theorem would be the Cantor-Dedekind-Bernstein Theorem.

The two-set formulation was cached already in Cantor's 1878 *Beitrag* but it was explicitly stated first in Cantor's 1887 *Mitteilungen* paper (p 413) and again in Cantor's 1895 *Beiträge* paper, §2, as Corollary B to the Comparability Theorem for cardinal numbers, presented there as Theorem A.

The single-set formulation was first stated in Cantor's letter to Dedekind of November 5, 1882.[1] In the important 1883 *Grundlagen* paper (§13),[2] which was published shortly after, an instance of the theorem was fully stated and proved. The instance was for sets of the power of (II) – the second number-class. There also the denumerable instance was mentioned. The latter, in a slightly different form, was already stated in Cantor's 1878 *Beitrag* as easy to demonstrate. It was nevertheless proved in Cantor's 1895 *Beiträge* (§6 Theorem B).[3]

Following his statement of CBT for sets of the power of (II) in *Grundlagen*, Cantor said that "this theorem has general validity". However, Cantor never published a demonstration to CBT in its general form. It was perhaps because the theorem was not stated there in full, as in the letter to Dedekind, that Bernstein (1905 p 121) references Cantor's 1887 *Mitteilungen* as the place where CBT was first proposed by Cantor. The single-set formulation was stated again in Cantor's 1895 *Beiträge* paper (§2), as Corollary C to Theorem A.[4] We will present Cantor's proof of the instance for sets of the power of (II) in Chap. 1. In Chap. 2 we will present a generalization of that proof, to sets with power in the scale of number-classes, which we believe is the proof Cantor promised in *Grundlagen*. Our generalization fills a gap in the common belief on the history of Cantor's set theory. It also led us to a new understanding of Cantor's Limitation Principle from the 1883 *Grundlagen*.

In Chap. 3 we will present evidence that Cantor knew of CBT already by the time he wrote his 1878 *Beitrag* and probably had a proof for it at least for the continuum. In Chap. 3 we will also promote the view that Cantor had, by the time of his 1878 *Beitrag*, the idea of his infinite numbers from 1883 *Grundlagen*, as well as the ideas of number-classes and his theorem about the power of the power-set. In Chap. 4 we discuss Cantor's theory of inconsistent sets and in Chap. 5 we explain how this theory provided Cantor with a proof of the Comparability Theorem for cardinal numbers. On this issue too our views are contrary to common belief on Cantor's theory. In Chap. 5 we also explain the difference between that theorem and

[1] Cantor-Dedekind 1937 p 55, Cavailles 1962 p 232, Dugac 1976 p 258, Hallett, 1984 p 59, Meschkowski-Nilson 1991 p 85f, Ferreirós 1993 p 353, 1995 p 37, Ewald 1996 vol 2 p 874.

[2] See Ewald 1996 vol 2 p 878–881 for bibliographical background on *Grundlagen*.

[3] As the proof is simple, it is reasonable to assume that Cantor possessed it already when he wrote his 1878 *Beitrag*. Cantor probably proved it in his 1895 *Beiträge* simply because he wanted at that time to present his theory in full detail.

[4] A third formulation of CBT, Corollary E, appears there as well.

Cantor's comparability of sets which leads us to discuss in Chap. 6 the scheme of complete disjunction.

Dedekind was the first to have learned of CBT and the first to provide the theorem with a general proof independent of the power concept. In Chap. 9 we bring Dedekind's proof of CBT with its historical background. However, prior to that we are led to discuss two issues regarding the relationship between Cantor and Dedekind. First (Chap. 7), with regard to the ruptures in their correspondence, we add a new explanation to those discussed in the literature. Then (Chap. 8) we discuss Cantor's criticism of Dedekind's infinite set, in view of his theory of inconsistent sets. Chapter 9 concludes the first part of the book.

In the second part of the book we bring the early published proofs of CBT. We relate to papers published in the period 1898–1901 by Schröder, Borel, Schoenflies and Zermelo. In the chapter on Schröder we also cover Korselt's criticism from 1911. In the chapter on Borel's proof we suggest a reconstruction of Bernstein's original proof upon which Borel based his own. Schoenflies' proof amalgamated the proofs of Borel and Schröder, together with a proof of Cantor communicated to him in a letter. Zermelo's proof unknowingly generalized ideas of Dedekind and stretched out the applicability of the cardinal number notion for a CBT proof. We conclude the second part with Bernstein's proof of a division theorem that we name the Bernstein Division Theorem (BDT). It says that if $k\mathfrak{m} = k\mathfrak{n}$ then $\mathfrak{m} = \mathfrak{n}$, where k is a natural number, $\mathfrak{m}$, $\mathfrak{n}$ cardinal numbers. We also touch upon an inequality version of the theorem: if $k\mathfrak{m} \geq k\mathfrak{n}$ then $\mathfrak{m} \geq \mathfrak{n}$. A research project emerging from BDT and the inequality-BDT plays a significant role in later chapters.

In part III we discuss proofs of CBT that emerged during the period 1902–1912. Most proof are related to the development of the logicist movement and the connected debate between Poincaré and Russell. First we present the CBT proof of Russell from 1902, which was circular. Then we discuss the role of CBT in the derivation of Russell's Paradox. We move to the attempts of Jourdain (1904) to extend Cantor's construction of the scale of number-classes. We point out the many defects of Jourdain's endeavor, which were, one may say, balanced by his ability to identify Cantor's theory of inconsistent sets and his proof of the Comparability Theorem. We cover Harward's criticism (1905) of Jourdain's 1904 papers and the latter's attempts to correct his 1904 exposition with new papers of 1907 and 1908. The papers of Poincaré of 1905–1906 are the centerpiece of part III. Several proofs of CBT are presented in these papers and there the criticism of an impredicative definition emerged. Poincaré's papers provoked several new proofs of CBT, mainly by Peano and Zermelo, who provided impredicative proofs. The latter gave the proof in his axiomatic system. Another proof came from J. Kőnig; it influenced his son, D. Kőnig, into a research project, on the tracks of BDT, which led to several interesting results in graph theory. We also cover Korselt's 1911 proof who applied it to solve a problem of Schröder. This problem initiated another research project through a paper from 1926 by Lindenbaum-Tarski and a paper of Sierpiński from 1947. The period surveyed in part III culminated in the Whitehead-Russell Principia Mathematica, which contains several proofs of CBT that we survey. We end part III with a discussion of Hausdorff's paradox, which we link to BDT.

Part IV of the book deals with results obtained mostly within the Polish school of logic concerning CBT, BDT and the inequality-BDT. The champions of this part are Sierpiński, Banach, Kuratowski, Lindenbaum, the Englishman Whittaker, Knaster, Sikorski, Reichbach and, foremost, Tarski.

The final part of the book includes only three chapters. One handles a proof by Hellmann from 1961. Another chapter concerns the attempts to port CBT to intuitionist context. The last chapter handles porting CBT to category theory.

My special interest in CBT arose when I first studied set theory and used Fraenkel's (1966) book, 'Abstract Set Theory', as textbook. There (p 72–79), two proofs of the theorem, of Poincarè and Whittaker, were given in detail and about a dozen more referenced. This surprised me, for I could not understand at the time, how theorems can have more than one proof and why anyone would be interested in providing a proof of an already proven theorem.

Another surprise came from the comparisons Fraenkel made between the CBT proofs that he brought. Of one of the proofs he said that it used the natural numbers, while the other was more abstract and applied the "union of a set of sets", which he linked to the proof of the Well-Ordering Theorem. Thus I learned that not only are there different proofs of a theorem, but there is an art of comparing proofs and proofs can compare even when they are not proofs of the same theorem! There and then I decided to compile one day all the proofs of CBT and study, through this compilation, how to compare proofs. This book is the outcome of that undertaking and in developing a discussion of proof comparison lies its second purpose.

Fraenkel's comparison of the CBT proofs was focused on the means whereby the proof is won, to which he tried to give a concise description. As I attempted to follow Fraenkel's example with regard to the proofs which I compiled I came to realize that what I was doing was giving metaphors to central aspects of the proof. Thus I obtained the view that proofs can be compared using their metaphors. Metaphors may involve mathematical notions (natural numbers) or template methods (complete induction, Cantor's diagonal method), as well as non-mathematical terms (pushdown an infinite stack). Thus there is room for idiosyncrasy in the generation of metaphors for proofs. The latter are especially of interest as they are rarely pointed out in mathematical texts, but are apparently used in informal discourse among peers.

Fraenkel did not compare the proofs he presented through their metaphors alone. For example, with regard to the proof of CBT of J. Kőnig Fraenkel remarked that it "is distinguished by its lucidity" and "has yielded remarkable generalizations". From the second remark I learned that the value of a proof can be appreciated only through a diachronic study, which I have contemplated in this book. In this I have followed the inspiring example of Lakatos' book, "Proofs and Refutations" (1976), where the history of proofs of Euler's polyhedron formula is presented.

From the first of the above remarks of Fraenkel, I have obtained another significant lesson. I compared J. Kőnig's proof to the first proof which Fraenkel brought. Both proofs looked equally lucid to me, so I wondered why Fraenkel deemed J. Kőnig's proof to be especially lucid. It occurred to me that

Fraenkel perhaps saw the lucidity of J. Kőnig's proof in that it shed new light on the mathematical situation of CBT. I thought that it was as if J. Kőnig identified in the situation of CBT a new gestalt. This then became my observation: proofs are characterized by gestalt, which, like metaphors, may contain non-mathematical attributes.

The gestalt explication, as a dimension in proof comparison, answered my qualm with regard to the reason why there are different proofs of a theorem and why people bother to give a new proof of a proved theorem. A proof is a description, like driving instructions. These are construed under a certain gestalt of the area in which the driving occurs. Different people may have different gestalt of the area and so they give different instructions. People are willing to offer their driving instructions because it is their own, special, dear to them, gestalt that they offer. People are even willing to debate whose instructions are better, with regard to such parameters as distance, complexity, traffic, etc. Same thing with mathematicians. They are eager to offer their gestalt. Mathematicians have an additional justification as it may turn out that one gestalt is more fruitful than another, as it happened with J. Kőnig's gestalt applied in his CBT proof.

With time I realized that I use gestalt with regard to the static elements of a proof, which are in general rather easy to identify, while metaphor I use for the dynamic elements, which, in general, change the gestalt. I will not dwell now on better explicating these notions since I believe that some experience in their use must be gained prior to such discussion.

Metaphors and gestalt ('proof descriptors') need not correspond to the entire proof; they may relate just to some of its parts. Actually, it is through descriptors that the parts of a proof are often discerned. In this capacity descriptors may thus serve as mnemonics. Descriptors are not necessarily verbal; they may include drawings or gestures. Many descriptors can be given to a proof and they need not be coherent or accurate in any objective sense, rather the opposite; informal character is essential to descriptors.

The process by which descriptors are generated from a proof, I call 'proof-processing'. I will be proof-processing the proofs presented in order to extract descriptors for their comparison.[5] With regard to the proofs of CBT I am after the differences in proof descriptors; however, I also compare proofs of different theorems that demonstrate affinity in terms of gestalt and metaphor. In such cases I believe that proof-processing of the older proof was behind the achievement of the newer proof. Mainly I will consider proofs of BDT, which received through the years several proofs related through proof descriptors to proofs of CBT.

Here Lakatos' 1976 book should be mentioned again. His book introduced me to a new subject: heuristic – patterns in the development of mathematics (Lakatos 1976 p 93, 143n4). Lakatos discerned the patterns of concept stretching, proofs and refutations, change of dominant theory, proof-analysis, hidden lemma. As a tribute

[5] However, I will not exaggerate in this and I will not undertake to extract proof descriptors from every proof presented or all proof descriptors of a proof analyzed.

to his influence on our work, for I believe that proof-processing is also a pattern for the development of mathematics,[6] I will note along our discussion, corroborating instances to elements of his theory.

The book is intended for researchers in the history and philosophy of set theory or in the methodology of the development of mathematics. Perhaps it will be of interest also to graduate students, who wish to take an excursion into the developmental area of mathematical research, which precedes textbooks.

Besides the concise historical remarks on CBT given by Fraenkel (1966), two general surveys of the history of the theorem were previously offered. First, there was a survey by Medvedev (1966, in Russian). It contains mainly anecdotal information on the early proofs of the theorem, up to 1906.[7] Second, there was a survey by Mańka-Wojciechowska (1984, in Polish). It covers not only the early proofs but also the later proofs of the 1920s that had a topological context and topological consequences. The second survey includes several remarks of a proof-processing nature. We will reference these works where appropriate.

Our main references, however, are the original papers and monographs of the mathematicians whose work we cover. In addition, we will often reference survey books, especially such that concern Cantorian set theory and its period. These include mainly Dauben 1979, Moore 1982, Hallett 1984, Meschkowski-Nilson 1991, Ferreiròs 1999, Grattan-Guinness 2000, Ebbinghaus 2007, Reference will also be made to many papers written by historians of mathematics including Grattan-Guinness, Moore, Peckhaus, Purkert, Ferreiròs, to name only a few. Occasionally we will reference some mathematics text books, such as Fraenkel 1966, Rogers 1967, Levy 1979.

The original material appeared in German, French or English. When we quote passages from texts in German or French for which available translations exist, we quote from those translations. Otherwise, the translations are ours.

Many mathematicians are mentioned in the book. Most of them are well known and information on them can easily be found on the internet. For the others we provide whatever details we happened to gather.

The book consists of 39 chapters divided into five parts, with an Introduction and a Conclusion. Most chapters are devoted to one source. The chapters are generally in chronological order, which, however, could not always be maintained because certain developments took place in parallel. Internal references are made by chapter or section numbers.

In the presentation we tried to balance between the desire to cite the original text and the desire to be concise. With regard to notation, we tried to bring the original notation but changed it when it is no longer current or if the presentation can be greatly improved by such change, which is always noted. We took special care to point out mistakes, even if only typos, in the original papers, figuring that they are now only rarely visited so that any such comment has historical value. In certain

[6] We will expand on this in the Conclusion to the book.

[7] A survey on Medvedev is given in Anellis 1994.

cases, when a lacuna in the original proof is obvious, we suggested how it could be completed. Any short-comings in these suggestions are our responsibility and reflect nothing on the original writer. The principle example here is our suggestion how Bernstein's proof of BDT for general *k* can be completed.

Unfortunately, the collection below is not complete. Omitted are papers published after 1973, which generally aim to port CBT to various mathematical structures that emerged following developments in the 1950s. They are quite numerous. Moreover, there is no warranty that all the relevant papers from before 1973 are included, even if mentioned in Fraenkel 1966.[8] Yet we believe that we have addressed the most known proofs, which were most influential.

Finally, I enjoyed writing this book; I hope someone will enjoy reading in it.

[8] Thus we have not addressed the results concerning CBT in the context of families of sets, such as Otchan 1942. and Bruns-Schmidt 1958.

Contents

List of Figures and Tables

Chapter 1
Cantor's CBT Proof for Sets of the Power of (II)

Cantor's 1883 *Grundlagen*, is Cantor's most important paper, at least with regard to his theory of infinite numbers. Though the 1895/7 *Beiträge* is more systematic and contains many more results and details, the core ideas, which Cantor never abandoned, appear in *Grundlagen*. These include the generation principles of the infinite numbers, the limitation principles, the template for the construction of the scale of number-classes and related results. It is here that we begin to unfold the story of CBT. In the following chapters we will discuss its generalization as well as its earlier beginning.

1.1 The Generation Principles

Numbers (finite and infinite) are introduced in *Grundlagen* (§11) as generated by two principles.

The first principle of generation: Given a number α, a new number is generated, called the successor or sequent of α, and denoted by $\alpha + 1$. If α is a successor its preceding number (immediate predecessor) is denoted by α-1.

Cantor described the generation process by speaking of the setting and unification of α with a unit (*Einheit*). The description is rather redundant. Implicitly, the first principle assumes a first number, say 1.

The second principle of generation: For any definite[1] succession of defined numbers in which no greatest exists, a new number is generated, which is called the limit or sequent of those numbers.

Note that a succession contains with every number all the numbers preceding it, unlike a sequence, but a sequence can be completed to a succession by adding to it

[1] This notion is not explicated in *Grundlagen*. It should most likely be understood as analog to Cantor's notion of well-defined set (see Sect. 3.1).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_1, © Springer Basel 2013

all the numbers preceding a member of the sequence. The sequent to this succession is referenced also as the sequent of the sequence.

A complete and transitive order relation between the numbers generated by the two principles is implicitly instituted by the principles with the sequent always different from the numbers generated previously, and always the smallest number greater than all these preceding numbers. Under this order the numbers form a scale. The principles generate the numbers by an inductive process, similar to the process whereby the positive integers[2] are traditionally generated. Thus proofs and definitions by induction, transfinite induction, over the scale of numbers are possible. Cantor did not use the term 'transfinite induction' but used the method for both definition and proof – see the proof of the Fundamental Theorem in the next section and the introduction of ordinal exponentiation in 1897 *Beiträge*, §18, (cf. Zermelo's remark, Cantor 1932 p 355 [24], Purkert and Ilgauds 1987 p 141). At the induction step of transfinite induction, one usually has to distinguish two cases: the move from a number to its sequent and the move from a succession to its sequent. Transfinite induction was perhaps first discerned in a publication in Harward 1905 (see Chap. 18). The name seems to have originated in Hausdorff 1914a.

The set of all positive integers (those generated by the first principle only) Cantor denoted by (I) and called it the first number-class. The sequent of (I) Cantor denoted by ω. By a dialectic process (Cantor 1932 p 148) Cantor produced, using the first principle, ω numbers following every sequent generated by the second principle and then a new number by the second principle following the numbers already generated. Thus such numbers are generated as $\omega + 5$, $7{\cdot}\omega$, $\omega{\cdot}\omega$, ω^6, ω^ω. In fact, to every arithmetic combination that one can imagine of the positive integers, ω and the operations of addition multiplication and exponentiation, the corresponding number can be generated by the two principles. Nevertheless, this does not imply that the arithmetic of the infinite numbers is thereby established. $\alpha + \beta$ or $\alpha\beta$, when α and β are arbitrary infinite numbers, are not defined. In fact, though Cantor was developing his theory of infinite numbers since the early 1870s, he managed to establish their arithmetic only in 1882 (see Sect. 2.3).

The 'power of a number α' is the power of the set of its preceding numbers that we name its 'segment-set' and denote also by α. In his 1883 *Grundlagen* Cantor began to use the notion 'power of a set' as an entity. The notion was silently introduced in 1883, no doubt as obtained from the set by abstraction. Though Cantor began to mention 'abstraction' only since his 1887 *Mitteilungen* (Cantor 1932 p 379) it appears already in his unpublished paper from 1884 (Cantor 1970 p 86). Arguably,

[2] Cantor uses the term '*positiven ganzen realen Zahlen*' which Ewald translated to 'positive integers' and we adopted when discussing Cantor's papers. Cantor excluded zero from his numbers. In the 1878 *Beitrag* paper Cantor always speaks of the positive integers as a sequence (*Reihe*) and not as a set; similarly in 1883 *Grundlagen* (Cantor 1932 p 195, Ewald 1996 vol 2 p 207 [1]). In his 1887 Mitteilungen he talks already about the "set of positive integers". Perhaps because after *Grundlagen* Cantor began to differentiate between a set and its sequencing.

Cantor used abstraction even earlier: to define the irrationals from fundamental sequences of rationals in his 1874 *Eigenschaft* paper, and in the definition of the sequent by the two generation principles described above.[3]

The notions required for the application of 'power', namely, 'equal power', 'equivalence', 'greater (smaller) power' and '1–1 mapping' were introduced in Cantor's earlier papers, the 1878 *Beitrag* and the 1874 *Eigenschaft*, and incorporated into *Grundlagen* by reference.[4]

The set of all the infinite numbers that are denumerable, namely, have equal power to (I), Cantor called the second number-class and denoted it by (II). Cantor's basic theorems on the power of (II) are presented in *Grundlagen* §12, 13. We discuss them in the following section.

1.2 Proof of CBT for Sets of the Power of (II)

Lemma 1 (the Sequent Lemma)[5] *A sequence* (α_i)[6] *of numbers from* (II) *has a sequent in* (II).

Proof: To prove Lemma 1 Cantor discerned two cases. If there is a largest member in the sequence, it is denumerable and its sequent is also denumerable (it is easy to establish an equivalence between the two sets) and is the required sequent to the sequence. Otherwise, Cantor extracted from the sequence an ascending subsequence that has for every member of the sequence a member greater than it. He then completed the subsequence to a succession. The succession is partitioned by the subsequence into ω partitions each of at most ω numbers. By the Denumerable Union Theorem and the denumerable CBT (see below), the succession is seen to be denumerable. Hence its sequent is denumerable and so in (II).

The construction of the subsequence given here is a classic. Cantor used the same method of choosing a subsequence from a sequence when he constructed the subsequence of nested intervals from a given sequence of points in his 1874 proof of the theorem that the continuum is not denumerable (Cantor 1932 p 117, Dauben 1979 p 52ff). The technique of choosing a subsequence is also employed to prove that ω^α is denumerable when α is denumerable. The origin of the method was no doubt in Weierstrass′ teachings of convergence (the Bolzano-Weierstrass Theorem).

---

[3] Weyl 1918 endnote 39 (to page 45) sees in Frege 1884 the origin of definition by abstraction.

[4] Cantor 1932 p 167, Ewald 1996 vol 2 p 883 [10]; cf. Sect. 3.1.

[5] The names and numbers of the theorems and lemmas are ours except the 'fundamental theorem' name, which was used by Cantor. The proofs that we bring are not exact citations from Cantor but rather complemented summary thereof.

[6] Cantor has this cute and efficient way of denoting a set by a pair of brackets, between which he puts a letter that serves him as a generic sign to denote the set's members.

We have here examples of proof-processing from within Cantor's own writings and his knowledge base.

It was Zermelo who pointed out in a remark, which he embedded in the text of *Grundlagen* (1932 p 198, not in Ewald), that the following theorem was used implicitly by Cantor in the proof of Lemma 1:

The Denumerable Union Theorem *The union of a denumerable set of denumerable sets is denumerable.*

The Denumerable Union Theorem was asserted by Cantor in his 1878 *Beitrag*[7] as easy to demonstrate and proved in his 1895 *Beiträge* (§6, (8)) in the form $\aleph_0 \cdot \aleph_0 = \aleph_0$[8] (or $\aleph_0{}^2 = \aleph_0$), namely, that the power of the set of all ordered-pairs[9] of positive integers is denumerable. Its proof is by the familiar enumeration of the rationals method (Fraenkel 1966 p 37, 101). The proof Cantor may have obtained as a student in a Weierstrass' seminar (Fraenkel 1930 p 199).[10]

In his 1908a paper (p 188 (5)) Zermelo noted that the proof of the Union Theorem makes use of the axiom of choice (rather, the denumerable axiom of choice), to choose for each element of the set a particular mapping from it to the positive integers (cf. Sierpiński 1948 p 220f). Strangely, the remark is not repeated in Cantor 1932.

Zermelo, however, failed to remark that for the application of the Union Theorem in the proof of Lemma 1, also the denumerable CBT is necessary, in case the partitions are finite. Cantor may have been aware of this need already in his 1878 *Beitrag* paper, for there, next to the Denumerable Union Theorem, he stated the denumerable CBT in the following form:

The Denumerable CBT *An infinite subset of a denumerable set is denumerable.*
Proof (1895 *Beiträge*, §6 Theorem B; cf. Fraenkel 1966 p 35): We can assume without loss of generality that the set is the set of positive integers and denote by A the subset. Define by complete induction a mapping from ω to A by choosing at each step the smallest number in A not yet chosen.[11] Because A is infinite it will not be exhausted in the process before ω is. On the other hand, every number m in A will be chosen at the m^{th} stage if it was not chosen earlier. Therefore, the mapping will be onto, which establishes the theorem.

From the above formulation of the denumerable CBT, it easy to prove its single-set formulation: If we have $M'' \subseteq M' \subseteq M \sim M''$, M denumerable, then M′ is

[7] Note that the term 'denumerable', in the context of 1878 *Beitrag*, is anachronistic; Cantor used it only since 1879 (Cantor 1932 p 169 footnote).

[8] The dot will be sometimes used in this book to signify multiplication. At other times we will simply use juxtaposition.

[9] It is anachronistic to use the term 'ordered-pair' in this context; not the notion itself. Cantor used 'indexed dummy' for such purpose in letters to Dedekind of 1873.

[10] We note in passing that the familiar method was probably due to an idea of Dedekind and that Cantor's original method was different (see Cantor's letter to Dedekind of December 2, 1973; Cavailles 1962 p 188f, Meschkowski-Nilson 1991 p 33f (in handwriting), Ewald 1996 vol 2, p 844f).

[11] Here Lemma 2(i) (see below) for (I) is used.

infinite, because it contains M″, and so it is denumerable. Assuming the single-set formulation of CBT, the 1878 *Beitrag* formulation can be proved using the theorem that the sequence of positive integers has the smallest power, namely, is equivalent to a subset of any infinite set. This ‘Smallest Power Theorem’ Cantor stated in his 1878 *Beitrag* and proved in his 1895 *Beiträge* (§6 Theorem A) by choosing from an infinite set a denumerable subset. The proof thus uses the denumerable axiom of choice.

The proof of the denumerable CBT assumes that the subset is not exhausted by removing from it any finite number of elements. Therefore, this is the tacit meaning of infinite set held by Cantor in 1878 *Beitrag*. The set of positive integers is such a set and therefore every set that contains an image of it, as M′ in the single-set formulation, satisfies this requirement. Note that the process of removing members from a set is a numbering process.

The method used in this proof of the denumerable CBT is important because it was proof-processed by Cantor, leveraged by transfinite induction, and applied in many instances, including in the proof of the Fundamental Theorem to be given shortly. We call it the enumeration-by method of proof (*Abzählbar durch*, Cantor 1932 p 169 footnote 1, 193, Ewald 1996 vol 2 p 885 footnote 1, 905).

The enumeration-by method is smoothly used when for both the enumerating set and the enumerated set Lemma 2 is applicable, but Cantor used it also to prove that every infinite set has a denumerable subset (1895 *Beiträge* §6 Theorem A), in which case the enumerated set does not have Lemma 2 and the axiom of choice is invoked. Cf. the proofs of the categoricity of the order-types η and θ (Cantor 1895 § 9, 11, Ferreirós 1999 p 279).

Because of Lemma 1, all the numbers defined by the two principles using ω, the positive integers and the arithmetic operations of addition, multiplication and exponentiation, are denumerable and belong to (II). Thus Cantor populated (II). This observation was used explicitly in *Grundlagen* (end of §11). In his 1882 paper (Cantor 1932 p 152, Ferreirós 1999 p 190), Cantor restated the two theorems, the denumerable union and CBT theorems, and remarked that they are the basis of every denumerability proof.

Theorem 1. *The power of* (II) *is different from the power of* (I).
Proof: Otherwise, (II) can be presented as a denumerable sequence and hence its sequent would be denumerable, by Lemma 1, and would belong to (II). However, this is a contradiction because a sequent is always different from its preceding numbers.

We call this argument, which Cantor applied here without explicitly noting it, the ‘Sequent Argument’. Note that here ‘different’ includes the possibility that the sets are not comparable. So, assuming the negation of this statement implies that the sets are equivalent and (II) is denumerable. In the later developed theory of ordinal numbers that substituted the infinite numbers, the theorem that a well-ordered set is never similar to a segment of it, replaced the Sequent Argument.

For later use let us remark here that it is obvious that (II) is not finite because it includes the numbers $\omega + n$ for every positive integer n and so it has more elements then every finite n.

Lemma 2. (i) *In every subset of* (II) *there is a smallest number*; (ii) *every descending sequence of numbers from* (II) *is finite*.

In *Grundlagen* (§2) Cantor introduced well-ordered set as a set that has the following properties: (i) it has a first member, (ii) every element except the last has a sequent, (iii) every succession (not cofinal with the set) has a sequent [in the set]. But he did not link it to the properties of Lemma 2. He did make this link in 1897 *Beiträge* §12.

Cantor only proved that Lemma 2(i) entails 2(ii) which is easy. Of 2(i) he only said that it follows from the definitions of the second number-class (II). Zermelo provided a proof of (i) in his note [4] (not in Ewald): He argued that if 1 belongs to the subset it is obviously the smallest; otherwise let A be the set of all numbers smaller than all numbers in the subset. A is clearly a succession. If A has a largest member its sequent by the first generation principle is not in A and it is thus in the set and is the smallest number of the set. Otherwise, A is a succession with no greatest number and its sequent by the second generation principle is again not in A and so is in the set and is its smallest number. Zermelo should have added that A is finite or denumerable so that its sequent is in (II). The reason is that since there is a number in the subset, which is a number in (II), A is a subset of a denumerable set and so by the denumerable CBT, A is finite or denumerable.

The proof of Lemma 2(i) changes a little when the infinite numbers (now ordinals) are no longer defined by the two generation principles but by abstraction from well-ordered sets. The set A consist of all the ordinals smaller than all the ordinals in the given set; A is well-ordered and since it is a succession (bounded by a member of the given set) it has an ordinal greater than all its members so necessarily in the given set, thus being its first element. Therewith, transfinite induction is enabled for the scale of ordinals.

The numbers in the subset are upper bounds to the numbers in A and the sequent is the least upper bound. But there is no impredicativity in its definition (see Sect. 19.3) because it is defined by an inductive argument.

Proof that (ii) implies (i) is not mentioned either by Cantor or Zermelo. It is easy assuming the axiom of dependent choices.[12] Jourdain (1904b p §3) noted that in no publication of the time was this result explicitly stated.

The Fundamental Theorem[13] *A subset of* (II) *is either finite, or has the power of* (I) *or of* (II).

Proof: To prove the Fundamental Theorem the enumeration-by method is used: Enumerate the subset by (II), namely, define by transfinite induction a 1–1 mapping from (II) to the subset by choosing at each step the smallest number in the subset not yet chosen (Lemma 2(i)). If the enumeration stops after a finite number of steps, the

[12] http://en.wikipedia.org/wiki/Well-ordered.

[13] The name was perhaps chosen because of the importance of the enumeration-by method used in it.

subset is finite. If the enumeration stops after α steps with α in (II), the subset is denumerable. Otherwise the enumeration must exhaust the subset and (II) simultaneously so that the two sets are equivalent. That a fourth case is not possible, as Cantor exclaimed triumphantly, results from the observation that any number β in the subset, being a number in (II), must be chosen at the β step of the enumeration if it was not chosen earlier. Otherwise an infinite descending sequence of numbers can be generated contrary to Lemma 2(ii).

From Theorem 1, the Fundamental Theorem and the remark on the infinity of (II), Cantor concluded that the power of (II) is the next power after that of (I), namely, it is the first non-denumerable power.

The Cantor-Bernstein Theorem for (II) *If* M *is a set of the power of* (II)*, and* $\mathrm{M}'' \subseteq \mathrm{M}' \subseteq \mathrm{M} \sim \mathrm{M}''$, *then* $\mathrm{M}' \sim \mathrm{M}$.
Proof (not detailed in *Grundlagen*): Without loss of generality we can assume that M is (II). Thus M′, as a subset of (II), is by the Fundamental Theorem either finite or denumerable or equivalent to (II). If M′ is finite then so is M″ by an easy to obtain lemma, which Cantor proved by complete induction in 1895 *Beiträge* (§5 Theorem E). Then M, or (II), would have to be finite, which is contrary to our observation above. If M′ is denumerable then M″, which as we saw cannot be finite, is denumerable by the denumerable CBT. However, this is contrary to Theorem 1. So M′ must be equivalent to (II).

1.3 The Limitation Principle

In addition to the two generation principles, Cantor introduced a third principle in *Grundlagen* (Cantor 1932 p 199, Ewald 1996 vol 2 p 911 [10]) – the Limitation Principle. It requires that a new integer generated by the two generation principles be admitted into the scale of numbers only if its power is that of "a defined number-class which was already *in existence* over its entire extent" prior to the definition of the said number. Demanding a succession to fulfill a certain test before it can generate a sequent resembles the demand that a sequence of rationals has to be a fundamental sequence before it can be taken to construe a real number (Cantor 1932 pp 184–185, Ferreirós 1999 p 125).

The positive integers are exempted silently from the Limitation Principle. ω is admitted because its power is the power of (I), which is a number-class whose full extent exists, by the admission of the positive integers. Any number obtained by the first principle of generation from an infinite number in the scale can be admitted to the scale because its power is that of its preceding number which is assumed to comply with the Limitation Principle. The sequent to the succession defined in the second case of Lemma 1 was shown to be denumerable by the Denumerable Union Theorem and the denumerable CBT, so it complies with the Limitation Principle. The sequent of A in Zermelo's proof of Lemma 2(i), is denumerable as we have argued, so it too complies with the Limitation Principle. The same applies to all the numbers that populate (II).

In Theorem 1 we assumed that there is a sequent to (II) only when we assumed that (II) is denumerable, in which case its tentative sequent complies with the Limitation Principle. However, once this assumption is rejected we cannot assume that the sequent of (II) exists. To enable the generation of the sequent of (II) we have to prove that the power of (I) + (II)[14] is equal to the power of a number-class, which necessarily means (II). This, in fact, is quite easy to obtain, for example, we can enumerate (II) by (I) + (II) as in the Fundamental Theorem. We obtain that the two sets are equivalent (we clearly have Lemma 2 for (I) + (II)). Therefore, the sequent of (I) + (II), which is also the sequent of (II), exists and in *Grundlagen* Cantor denoted it by Ω.

A mistaken view regarding the generation of Ω is Schoenflies' (1900 p 44) view that the second principle holds for sequences (Schoenflies did not mention the requirement for successions) of the type ω only (such sequences are also called fundamental sequences). He even claimed (p 48) that a new generation principle is necessary to create the numbers of the third number-class. This view, however, was not general. Thus Jourdain apprehended the issue correctly (Cantor 1915 p 59, Jourdain 1904b p 295 footnote *, 300, 301), and so did Fraenkel (Fraenkel 1930 p 243), though Fraenkel presents the principle as if it applies to any set from the beginning and not just to successions as did Cantor.

It was unfortunate that Cantor did not mention the need to prove that Ω is admissible, namely, that (I) + (II) ~ (II). Some confusion encircling the Limitation Principle would have been avoided, such as the view that the Limitation Principle is necessary for the comprehension of the number-classes. Indeed, Jourdain (Cantor 1915 p 60, 1904b p 300) and Dauben (1979 pp 98, 196) vaguely, while Hallett (1984 p 58) explicitly, seem to hold this view. However, Cantor used unrestrained comprehension, as when he speaks in G*rundlagen* (Cantor 1932 pp 193, 207 (10), Ewald 1996 pp 905 [7], 919 [10]) of various sets of functions without any tailored comprehension principles.

Having proved that Ω fulfills the Limitation Principle (in the next chapter this result will be generalized to all initial numbers) the question arises: what is the function of the Limitation Principle? Our answer is that the principle was designed to bar the monster (Lakatos 1976) of the antinomy of the number of all numbers. This entity that seems reachable by the second principle of generation, is not admissible to the scale of numbers because it cannot be proved that it is equivalent to a number-class. It seems that it was because of the Limitation Principle, that Cantor pointed out (§12) that while the procedure of defining new number-classes can go on without end, the scale of number-classes cannot be perceived as one

[14] We use here + for the union operation of two sets, an operation that we assume the reader to be familiar with, as with other similar notions. Elsewhere we may use also ∪ for the same operation. Notice that + serves us also for the sum operation between numbers which will be discussed in Sect. 2.2 in the context of the Sum Lemma.

(endnote 1). The formulation of the Limitation Principle, which distinguished between initial numbers (see the next chapter) and the number of all numbers, was a great achievement of Cantor that enabled him to master the transfinite.

We agree with Purkert-Ilgauds (1987 p 151ff; cf. Dauben 1979 p 246) that it is unlikely that Cantor was not aware of the antinomy of the number of all numbers, which appears in Leibnitz. Ferreirós' hesitation on this point (1999 pp 290–292) is not justified. Ferreirós (1999 p 272) points out correctly that the Limitation Principle links the ordinal and cardinal numbers (to use Cantor's later terminology) but he does not explain the purpose of this link.[15] Purkert (1989 p 56) thinks that because of his discovery of the antinomy of the number of all numbers, Cantor turned to philosophical arguments in *Grundlagen*. However, Cantor was not interested in having a philosophical framework for his mathematics, as if to refute Russell (in 1899, Moore 1995 p 225) for whom the antinomy of the number of all numbers "no existing metaphysics avoids". We believe, on the contrary, that to bar against the paradoxical collections Cantor employed mathematical means (the Limitation Principle or later the theory of inconsistent sets). We accept the traditional view that Cantor's philosophical discourse simply surveyed views accepting the actual infinity.

1.4 The Union Theorem

Having obtained Ω, the third number-class (III) can be defined as the set of all numbers with power equal to the power of Ω (which is the power of (II)). However, to populate it with such numbers as $45\Omega^{\Omega} + 27\Omega^{99} + 7\omega + 14$, Lemma 1 adapted to Ω is needed. Proving that these numbers have the power of Ω also settles for them compliance with the Limitation Principle. The proof of the adapted lemma is analog to the proof of Lemma 1, except that we now have to use the Union Theorem for sets of power equal to the power of Ω and CBT for such sets. The latter theorem Cantor had obtained but he did not mention the first and did not populate (III).

Regarding Cantor's awareness to the Union Theorem for non-denumerable powers we have a testimony by Bernstein who wrote (1905 p 150) that Cantor told him he had generalized his result $\aleph_0^{\,2} = \aleph_0$ (the Denumerable Union Theorem) to any aleph. The testimony surely refers to the late 1890s (at the earliest) but it is quite reasonable that Cantor had the theorem when he wrote *Grundlagen*.[16]

As things stand, however, it was Jourdain (1904b, see Sect. 17.6) who first pointed out publicly the importance of the Union Theorem to the construction of the number-classes and attempted to prove it. Jourdain's 1904a proof was not

---

[15] Ferreirós' claim (p 268) that the Limitation Principle plays a role in the proof of the Cantor-Bendixson theorem, is not clear to us.

[16] Use of the term 'aleph' ($\aleph$) in any context earlier than 1895 is anachronistic.

rigorous and was corrected by Harward in 1905, who actually gave two proofs of the theorem. Then the theorem was proved by Hessenberg (1906; cf. Fraenkel 1966 p 219). Jourdain came back to the issue in his 1908a paper (see Sect. 23.2) with a rigorous proof. These proofs, however, rely on the general CBT which was available to the mentioned mathematicians through proofs by Borel (1898, Chap. 11), Schoenflies (1900, see Chap. 12) or Zermelo (1901, see Chap. 13), but Cantor used the Union Theorem for his proof of CBT and could not rely on the latter to prove the former. Still, Cantor could have obtained the following proof, based on the Fundamental Theorem:

The Union Theorem for Ω *The union of a set of power equal to the power of* Ω*, each of its members a set of power* Ω*, is of power* Ω.
Proof: As a set representing the conditions of the theorem we take the set whose members are the columns of the set of all ordered-pairs[17] of numbers from (I) + (II). Each column is equal to the power of Ω and there are Ω such columns. The union of this set is then the matrix of all these pairs and we have to prove that its power is equal to the power of (I) + (II).

Define an order on the union set as follows: if (α,β) and (α',β') are two ordered-pairs of numbers from (I) + (II), define $(\alpha,\beta) < (\alpha',\beta')$ when $\alpha+\beta < \alpha'+\beta'$ or if $\alpha+\beta = \alpha'+\beta'$ when $\alpha < \alpha'$. Note that if $\alpha+\beta = \alpha'+\beta'$ and $\alpha = \alpha'$ then $\beta = \beta'$, for otherwise, if for example $\beta < \beta'$, $\alpha+\beta$ is in the segment-set of $\alpha+\beta'$ and thus the two must be different by the Sequent Argument.[18]

Here we assume that the sum of two finite or denumerable numbers is defined and is finite or denumerable, respectively. This is the finite or denumerable Sum Lemma. We will get back to this subject in Sect. 2.2. In *Grundlagen* Cantor took $\aleph_0 + n = \aleph_0 + \aleph_0 = \aleph_0$ as self-evident (§11). Only in 1895 *Beiträge* (§6) did he prove these equalities. Note that there, for the proof of the Denumerable Union Theorem, the closure of the finite numbers under addition is used, namely, the finite Sum Lemma.

The ordered union set has the property that every subset of it has a first number.[19] Indeed, if A is such a set, consider the set of all $\alpha + \beta$ for all (α,β) in A. This is a set of numbers from (I) + (II), because the sum of every two finite or denumerable numbers is finite or denumerable, and by Lemma 2(i) (extended to (I) + (II)) it has a smallest member, say ζ. Consider now the set of all α such that there is β so that $\alpha+\beta = \zeta$. This set is again a set of numbers of (I) + (II) and so by the extended Lemma 2(i) it has a smallest number α'. Let β' be the number such that $\alpha'+\beta' = \zeta$. Then (α',β') is the smallest number of A. (There is only one such ordered-pair (α',β') by the previous Sequent Argument.)

[17] We denote ordered-pairs by (α,β).

[18] The definition of the order (and of ρ below) is taken from Jourdain 1908a proof.

[19] Namely, the set is well-ordered but Cantor did not use this terminology with regard to Lemma 2 and so we refrain from it as well.

We can now employ the enumeration-by method to enumerate the set of ordered-pairs of numbers from (I) + (II) by (I) + (II), choosing to each number in (I) + (II) not yet assigned, the least ordered-pair not yet chosen in the definition process from the set of ordered-pairs. There are three possibilities: the process will terminate before (I) + (II) is exhausted or it will terminate before the set of ordered-pairs is exhausted or the two sets will turn out to be equivalent. We will show that the first two cases cannot occur and thus obtain the theorem.

In the first case let ζ be the first number of (I) + (II) to which no ordered-pair is assigned. Then ζ is finite or denumerable. But the set of ordered-pairs contains a subset equivalent to (I) + (II), namely, the set of ordered-pairs (α,1) with α from (I) + (II), which, by the denumerable CBT would be finite or denumerable if so is ζ, contrary to our proof in the previous section, based on Theorem 1, that Ω is of the power of (II). Thus this case cannot happen.

In the second case, let (α', β') be the first ordered-pair in the set of ordered-pairs which was not assigned to a number by the enumeration process. Let B be the set of ordered-pairs smaller than (α', β'). Every ordered-pair (α,β) in B was assigned a number during the enumeration process and so B is of the power of (I) + (II) which is not denumerable. We have $\alpha, \beta \leq \alpha + \beta \leq \alpha' + \beta' < \alpha' + \beta' + 1$. Denote $\alpha' + \beta' + 1$ by ρ. Then B is a subset of the set of all ordered-pairs of numbers smaller than ρ. However, ρ is denumerable and by the Denumerable Union Theorem so is the set of ordered-pairs of infinite numbers smaller than ρ. Then, by the denumerable CBT, B has to be denumerable, which is a contradiction because B is equivalent to (I) + (II).

1.5 The Principles of Arithmetic

We briefly digress in this section in order to point out that in fact Cantor had provided in his 1883 *Grundlagen* an axiomatic system that is equivalent to the system given by Peano 6 years later, in his famous 1889 paper.[20] Strangely, in the literature on Cantor, this observation is never made.

Cantor's first principle of generation corresponds to Peano's sequent operation and its axioms (Peano 1889 p 94). Here is how Cantor described the first principle (§11):

> The formation of the finite real integers thus rests upon the principle of adding a unity to an already formed and existing number.

For Cantor, just as for Peano, a new number is generated by the sequent operation. Obviously Cantor must assume a first number, just as Peano assumed 1 as a constant of the system. Cantor does not state the requirement that the first number is not a sequent, which in a non-formal setting can be regarded as obvious. Likewise Cantor took as self evident that a number is uniquely determined by its predecessor.

[20] Of course Cantor did not employ the logical notation introduced by Peano.

Most interesting is that Cantor had his version of an axiom for complete induction.[21] It is the above Lemma 2(i) which says that every set of numbers has a first number. It is well-known that this statement is equivalent to Peano's axiom for complete induction. We now understand why Cantor never bothered to justify Lemma 2(i): he took it as a self-evident truth, indeed, as an axiom.

The proof that Zermelo gave to Lemma 2(i) rests on the second generation principle because it needs to consider the situation when the set A of that proof has no greatest number and thus its sequent must be generated by the second principle of generation. Thus we see that when the numbers (finite and infinite) are generated by the two generation principles, no additional axiom for induction (complete or transfinite) is necessary.

[21] Cantor must have had his version of the axiom for complete induction when he read Dedekind's 1872–78 draft (Dugac 1976 p 293–309) of *Was sind und was sollen die Zahlen?* ('*Zahlen*') (Dedekind 1963, see Sect. 4.3).

Chapter 2
Generalizing Cantor's CBT Proof

Following his statement of CBT, in its single-set formulation, for sets of the power of (II), as a corollary to the Fundamental Theorem, Cantor said (Cantor 1932 p 201, Ewald 1996 vol 2 p 912 [12]):

> [T]hat this theorem has general validity, regardless of the power of the set M, seems to me highly remarkable. I shall go further into this matter in a later article and then indicate the peculiar interest which attaches to this general theorem.

However, Cantor never came back to fulfill this promise. Moreover, in his 1895 *Beiträge*, Cantor presented CBT as corollary C to the Comparability Theorem for cardinal numbers. Thus it became a generally held view in the literature on early set theory that Cantor never proved CBT directly, for sets with power other than the power of (II). For example see Zermelo in Cantor 1932 p 209 [5], Fraenkel 1966 p 77 footnote 1, Medvedev 1966 p 229f, Levy 1979 p 85, Dauben 1979 p 172, Hallett 1984 p 60 footnote 2, p 74, Ferreirós 1999 p 239, Felscher 1999; Grattan-Guinness 2000 p 94.

Cantor's statements to the contrary, made in the letter to Dedekind of November 5, 1882, and in the above cited passage, are either not mentioned (Dauben, Hallett) or brushed away as a mistake Ferreirós (1999 p 239 footnote 4). Because no generalization to Cantor's proof from *Grundlagen* (see Chap. 1) has been suggested yet,[1] nothing appears to counter this view. In this chapter we hope to correct this misconception by producing a generalization of the proof given in *Grundlagen*.[2] To this end we will show how the scale of number-classes is defined in Cantor's terms past the second number-class.

One assumption must be made though. In the passage cited, the term 'the power of the set M′ must be interpreted to mean that M can be gauged by Cantor's scale of

[1] Hallett (1984 §2.2) attempted to obtain such generalization but concluded (p 74) that it was not possible.

[2] Alone to state a view similar to ours, but with actually no details to support his thesis, is Tait (2005 p 164).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_2, © Springer Basel 2013

number-classes. It is the generalization of the assumption made for the proof given in the previous chapter that M has the power of (II). This assumption is justified by the well-ordering principle and the numbering principle which Cantor introduced in *Grundlagen* (§2). The well-ordering principle states that every set can be well-ordered and the numbering principle states (somewhat obscurely) that every well-ordered set is similar to the segment-set of one of the numbers generated by the generation principles. This number is then called the *Anzahl* of the set. Thus every set has a power in the scale of number-classes (see Chap. 4), to the definition of which we turn.

A caveat regarding the proof of the generalized Union Theorem must also be mentioned: the only evidence that Cantor had a proof of the theorem is Bernstein's testimony mentioned in Sect. 1.4. But then, Jourdain had a proof by 1908a (see Sect. 23.2), so it is not improbable to assume the same of Cantor, nearly 25 years earlier.

2.1 The Scale of Number-Classes

Our first step will be to define Cantor's scale of number-classes in its full generality as envisioned by Cantor. We begin with a definition of notation: Let us denote by (γ) the γth number-class, by ω_γ the first number in (γ), the so-called the initial number of (γ), and by U_γ the union of all (κ) for $(0<)\kappa<\gamma$. This notation is not compatible with the common notation when γ is a positive integer. Thus, after Cantor, (I), (II), (III) are used for our (1), (2), (3), and for our ω_1, ω_2, ω_3,... the common use is 1, ω, ω_1,... . We stick to the general notation when we speak in general terms and to the common notation when we speak in particular terms. A similar practice is common in the literature.[3]

ω_γ is the sequent of U_γ and thus its power is the power of U_γ and is the γth power. (γ) is the set of all numbers that have the same power as ω_γ. Cantor later (1895 *Beiträge*) denoted by $\aleph_0$ (aleph zero) the power of ω, by $\aleph_1$ the power of his ω_1 and so on. Under our generalized notation $\aleph_\gamma$ denotes the power of ω_γ but again we stick to Cantor's terminology when γ is a positive integer.[4]

Though in *Grundlagen* Cantor did not mention explicitly number-classes for limit index, his view on the matter is clear from a remark which he made in endnote 2 (Cantor 1932 p 205, Ewald 1996 vol 2 p 917):

> The number γ which gives the order of a power (in case the number γ has an immediate predecessor) stands to the numbers [*Zahlen*; Ewald has here 'number'] of the number-class that has that power in a relationship of size whose smallness mocks all description; and all the more so, the greater we take γ to be.

Since Cantor speaks here of successor γ we infer that he had in mind also powers of orders which are limit numbers. Then, if γ is a limit number and we apply Cantor's

---

[3] The convention to denote by ω_γ the initial number of the $(\gamma + 2)$th number-class (Jourdain 1904b p 295 footnote †) will not do for number-classes from the $(\omega + 1)$th on.

[4] The confused notation originated with Cantor in the following quotation.

observation to $\omega_{\gamma+1}$, we are led to conclude that ω_γ must be a singular number. Because if ω_γ is a regular number, it is equal to γ and then $\gamma + 1$ has the power of ω_γ and so it belongs to the number-class (γ) and is not mockingly smaller than the numbers in that number-class. Indeed, in the definition of ω_γ, we assumed that γ is given, namely, defined before ω_γ. Thus ω_γ cannot be equal to γ and ω_γ is a singular number.[5]

In the quoted paragraph Cantor assumes that the γth power, when γ is a successor number, has the power of the number-class (γ-1). This is the Limitation Principle applied to successor initial number, which must be proved. As a number in a number-class has the power of the initial number of its number-class, it is necessary to prove compliance with the Limitation Principle only for successor initial numbers, such as Ω. It should be stressed that this is the crux of Cantor's construction. It is in order to prove compliance with the Limitation Principle (rather the Limitation Theorem for successor initial numbers) that the bunch of theorems of §11, 12 of *Grundlagen* must be generalized.

What about the Limitation Principle for ω_γ with limit γ? Clearly, such ω_γ is not the power of a number-class. We must conclude that Cantor exempted the singular numbers from the Limitation Principle. In the same endnote from which we quoted above, Cantor said:

> every transfinite [*überendliche*][6] number, however great, of any of the higher number-classes… is followed by an aggregate of numbers and number-classes whose power is not the slightest reduced compared to the entire absolutely infinite aggregate of numbers.

Cantor further equated there this property of the absolutely infinite with the similar property of ω with regard to the positive integers. Thus the absolutely infinite was regarded by Cantor to be regular, just as ω is. Therefore, singular numbers cannot represent the absolutely infinite and no Limitation Principle is necessary for them.

Zermelo's comment (Cantor 1932 p 199, in the text, not in Ewald) that already for forming the ωth number-class Cantor's principles do not suffice, emerged probably because ω_ω is not equivalent to a number-class and not because he thought that the second generation principle is not sufficient to generate ω_ω or because of the lack of a tailored comprehension principle. For some reason Zermelo ignored the possibility that the Limitation Principle is not necessary for limit initial numbers.

---

[5] Cantor did not use the terms 'regular number' and 'singular number', introduced by Hausdorff (1914a p 130).

[6] Cantor is using here the term 'transfinite number' which gained popularity after Jourdain used "transfinite numbers" in the title of his translation (Cantor 1915) of Cantor 1895/7, instead of Cantor's *Transfiniten Mengenlehre*. But Cantor preferred in his 1897 the term 'ordinal number' and in *Grundlagen* he preferred *unendlichen realen Zahlen*, which Ewald translated into 'infinite integers' (Ewald 1996 vol 2 pp 883 [7], 908 [7]). We prefer to use 'infinite numbers' to stress that on the one hand we are attached to the *Grundlagen* presentation and on the other hand to the 'number' attribute of 1897.

Our view on the limit initial numbers is not contradicted by Cantor's proclamation in *Grundlagen* that every power is represented by a number-class (Cantor 1932 pp 167, 181, Ewald 1996 vol 2 pp 884 [14], 895 [1]), for Cantor spoke there only of successor initial numbers.

Our reconstruction of Cantor's scale of number-classes is based solely on evidence taken from *Grundlagen*. Nevertheless it is compatible with Cantor's definition of the scale in his letter to Dedekind of August 3, 1899.[7] In later chapters we will show evidence that also other results mentioned in that letter, Cantor possessed already when he wrote *Grundlagen*.

In the literature, Cantor's development of the scale of number-classes, as presented in *Grundlagen* and reconstructed here, is generally ignored and Hessenberg's presentation (1906 Chaps. 13 and 14) is adopted. In it, however, CBT is assumed (Chap. 7),[8] while in Cantor's original development this theorem was proved with the construction of the scale, to which subject we turn next.

2.2 The Induction Step

We will now generalize the theorems of Sect. 1.2 for sets with power in Cantor's scale of number-classes. The proof is by transfinite induction and the theorems must be proved in tandem because of their interdependencies. The base case given in Sect. 1.2 corresponds to $\gamma = 3$: ω_3 is Ω of *Grundlagen*; U_3 is (I) + (II), U_2 is (I)[9]; (2) is (II), (1) is (I). The induction hypothesis is that all the theorems below hold for every $\gamma', 3 \leq \gamma' < \gamma$. We have to prove that hence they hold for γ. Alone is Lemma 2 that need not be repeated for it holds for any set of numbers by the same arguments used to justify it in Sect. 1.2.

Before we begin we call the reader's attention to the difference between arguments that use Lemma 2, for example in the proof of the Fundamental Theorem below, arguments that use the Sequent Argument, for example in the proof of Theorem 1 below, and arguments that assume theorems in the pack under the induction hypothesis. Despite their differences, all these argument types use transfinite induction, but the first two use what may be called local induction, the inductive properties of the infinite numbers defined up to the induction step, and they do not invoke the induction hypothesis.[10]

---

[7] Cantor 1932 p 443, van Heijenhoort 1967 p 113, Grattan-Guinness 1974 p 128, Ewald 1996 vol 2 p 931, Meschkowski-Nilson 1991 p 407. In the first two sources the letter is dated July 28. See the Grattan-Guinness reference.

[8] Hessenberg says that his proof follows the proof of Bernstein (namely – Borel, see Sect. 11.2) with some changes. Actually his proof is similar to Peano's first (inductive) proof (see Sect. 20.1), published in March 1906. Another case of simultaneity of proofs.

[9] U_1 is the empty set. We assume by convention that when we use U_k, we have $\kappa > 1$.

[10] Compare to Zermelo's three principles of complete induction in theorems I, III, V, of his 1909 paper (p 192).

Theorem 1. *If γ is a successor number, the power of* (γ-1) *is different from the power of* U_κ *for any* $\kappa < \gamma$.

Proof: Otherwise, by Lemma 1 (see below) for γ-1, assumed by the induction hypothesis, (γ-1) would have a sequent in it, a contradiction by the Sequent Argument.

The Fundamental Theorem. *A subset of* U_γ *is either finite, or has the power of* U_κ *for some* $\kappa \leq \gamma$.

Proof: Let (α') be the subset; note that it is ordered according to the size of its members. Enumerate (α') by U_γ: let α_1 be the smallest number in (α') (Lemma 2(i)), α_2 be the next larger (Lemma 2(i)), and so on (Lemma 2(i)). Having defined all α_ζ with $\zeta < \zeta' \varepsilon U_\gamma$, and if ($\alpha'$) is not yet exhausted, $\alpha_{\zeta'}$ is defined to be the first member of (α') not yet selected (Lemma 2(i)). Now, if (α') is exhausted by this process after a finite number of steps it is finite. If it is exhausted at the ζ step and ζ belongs to (κ), $\kappa < \gamma$, then the power of (α') is the power of ζ, which is the power of ω_κ, which is the power of U_κ. That ω_κ complies by the Limitation Principle, when κ is a successor number, follows from the Limitation Theorem (see below) for κ, assumed by the induction hypothesis. If (α') is not exhausted at any ζ, then, since for every α' there is a ζ such that α' is α_ζ (if not for $\zeta < \alpha'$ then for $\zeta = \alpha'$; here Lemma 2 (ii) is used), the procedure renders the equivalence of (α') and U_γ.

The Limitation Theorem. *If γ is a successor number, the power of* (γ-1) *is equal to the power of* U_γ*, which is the power of* ω_γ*, which thus fulfills the Limitation Principle.*

Proof: If γ is a successor number then as (γ-1) $\subset U_\gamma$, and as (γ-1) is not finite, and is not, by Theorem 1, of the power of some $U_\kappa, \kappa < \gamma$, ($\gamma$-1) must be, by the Fundamental Theorem, equivalent to U_γ, so ω_γ fulfills the Limitation Principle.

The Limitation Theorem was proved for Ω in the previous chapter separately from the other theorems, because Cantor did not mention it. It is important to realize that without the Limitation Theorem, the scale of numbers, and thus the scale of number-classes, does not exist. Moreover, all the theorems in this section, including CBT, are necessary to establish the Limitation Theorem. Thus, Cantor's proof of CBT is part of his construction of the scale of numbers and number-classes and not a separate theorem that can rely on the existence of the said scale. This point is not noted in the literature on Cantor's set theory.

The Different Alephs Theorem. *The power of* U_γ *is different from that of any* $U_\kappa, \kappa < \gamma$.

Proof: If γ is a successor number this follows from Theorem 1 and the Limitation Theorem. Otherwise, we would have that ω_γ (exempt from the Limitation Principle) belongs to (κ), a contradiction by the Sequent Argument.

The name of the previous theorem was given to it because the power of the U_γ is $\aleph_\gamma$ and thus the theorem says that all the alephs are different. From the preceding theorems it follows that the power of (γ-1), for γ successor, is the next following the power of $U_{\gamma\text{-}1}$. From this results the Next-Aleph Theorem follows: $\aleph_{\gamma+1}$ is the next aleph following $\aleph_\gamma$.

The Cantor-Bernstein Theorem. *If* M *is a set of the power of* U_γ, *and* $M'' \subseteq M' \subseteq M \sim M''$, *then* $M' \sim M$.
Proof: Without loss of generality we can assume that M is U_γ and thus M' is a subset of U_γ. If M' is not equivalent to M then by the Fundamental Theorem it is either finite or of the power U_κ for some $\kappa < \gamma$. M' cannot be finite as it contains a copy of (I) contained in M'' because $M'' \sim M$. Therefore M'' can, without loss of generality, be regarded as a subset of U_κ and thus either finite (which it is not by the mentioned argument), or, by the Fundamental Theorem for κ, of the power of U_ρ for some $\rho \leq \kappa$. However, this is in contradiction to the Different Alephs Theorem.

The naive reaction of someone educated in college set theory, is that CBT for sets of infinite numbers must be trivial. Our proof shows that this is not the case in the context of early Cantorian set theory.

The enumeration-by method, and in particular the enumeration of the subset by the whole set as utilized in the proof of the Fundamental Theorem, is the metaphor of Cantor's proof of CBT. Cantor's gestalt is that every set can be enumerated. It seems that Cantor's voyage into the infinite began with the maxim "the part is smaller than or equal to the whole" replacing the antique "the part is smaller than the whole" (see Schröder 1898 p 336).[11]

In view of the Limitation Theorem, the Fundamental Theorem and CBT can be phrased, for γ successor, for subsets of (γ-1).

The Union Theorem. *The set of all ordered-pairs of numbers from* U_γ (*this set is customarily denoted by* $\omega_\gamma{}^*\omega_\gamma$) *is equivalent to* U_γ.
Proof: We define the order in the set of ordered-pairs as in the proof for (I) + (II) given in the previous chapter and likewise prove that every subset has a first member under this order. Then we enumerate the set of ordered-pairs by U_γ and move to deny that the enumeration exhausts either of the sets before the other. In the definition of the order we rely on the Sum Lemma that the sum of any two numbers from U_γ is in U_γ. We will discuss this lemma below.

First assume that the set of ordered-pairs is exhausted before U_γ. Let ζ be the first number in U_γ to which no corresponding ordered-pair was assigned. $\zeta \in (\kappa)$ for some $\kappa < \gamma$ so ζ is equivalent to U_κ. The set of all $(1, \chi), \chi < \gamma$, ordered-pairs is thus equivalent to a subset of U_κ, so that, by the Fundamental Theorem for U_κ assumed under the induction hypothesis, this set, which is equivalent to U_γ (so obviously is not finite), has the power of $U_{\kappa'}$ for some $\kappa' \leq \kappa$, so $U_\gamma \sim U_{\kappa'}$, contrary to the Different Alephs Theorem.

If, on the other hand, U_γ is exhausted before the set of ordered-pairs then let (μ', χ') be the first ordered-pair not chosen in the enumeration process and B the set of all smaller ordered-pairs which are the ordered-pairs chosen by that process. B is of the power of U_γ. We have for all (μ,χ) in B, $\mu, \chi \leq \mu + \chi < \mu' + \chi' + 1$.[12] Denote $\mu' + \chi' + 1$ by ρ (Jourdain 1908a; cf. Lindenbaum-Tarski 1926 p 308f).

[11] That the part cannot be greater than the whole is provided by CBT.

[12] The definition of the sum of numbers from U_γ will be discussed below.

Then ρ belongs to some (κ), $\kappa < \gamma$, by the Sum Lemma (see below), and B is equivalent to a subset of the set of ordered-pairs of numbers smaller than ρ and this set is, under the induction assumption, equivalent to U_κ. By the Fundamental Theorem for κ assumed under the induction hypothesis, B is equivalent to some $U_{\kappa'}$ for some $\kappa' \leq \kappa$, which means that $U_\gamma \sim U_{\kappa'}$ contrary to the Different Alephs Theorem.

Therefore, it turns out that the enumeration process puts the set of ordered-pairs in equivalence with U_γ.

The Sum Lemma used in the above proof is the following:

The Sum Lemma. *If* α, β ε U_γ *then* $\alpha + \beta$ *is also in* U_γ. *In fact, the power of the sum is equal to the power of the greater of the summands, therewith the sum fulfills the Limitation Principle.*

The Sum Lemma assumes the definition of the sum operation which we will discuss below. Its proof is easy to obtain from the Union Theorem and CBT but as we require it for the proof of the Union Theorem, a direct proof is necessary.

Proof: First we prove that the power of $\alpha + \alpha$ is equal to the power of α. We partition the segment-set of α into two sets: one set contains all the finite even numbers, all the limit numbers δ and all the numbers $\delta + 2n$[13]; the other set contains the finite odd numbers and the numbers $\delta + 2n - 1(n > 1)$. Each of these sets is equivalent to α: the first by mapping the limit numbers to themselves, finite n to $2n$, and the numbers $\delta + n$ to $\delta + 2n$; the second by mapping finite n to $2n + 1$, the limit numbers to their successors, and the numbers $\delta + n$ to $\delta + 2n + 1$. Hence it is obtained that the power of $\alpha + \alpha$ is equal to the power of α.

For $\alpha + \beta$ (or $\beta + \alpha$), where $\beta < \alpha$, a similar procedure on the segment-set of β as a segment of α will deliver the result.

Interestingly, the Sum Lemma seemingly provides an alternative proof of Theorem 1 which bypasses Lemma 1 and thus AC and the Union Theorem: If we assume that $(\gamma\text{-}1) \sim U_\kappa$ for some $\kappa < \gamma$ then by the Sum Lemma, the power of ω_γ is equal to the power of $\omega_{\gamma\text{-}1} + \omega_k$ which is the power of $\omega_{\gamma\text{-}1}$, so ω_γ would belong to $(\gamma\text{-}1)$, a contradiction by the Sequent Argument. Zermelo noticed this proof in his remark [22] on Cantor's 1897 *Beiträge* (Cantor 1932 p 355) referring to Theorem D of §16, the theorem that the power of (II) is not $\aleph_0$, which is the denumerable instance of Theorem 1 (see Sect. 1.2).[14] But this proof has a lacuna: unless k is a successor number (or ω), we do not know that ω_γ fulfills the Limitation Principle.

Lemma 1. (the Sequent Lemma). *A sequence* (α_i) *of numbers in* (γ), $i < \kappa \leq \omega_\gamma$, *has a sequent in* (γ).

Proof: If the sequence has a greatest member then its sequent is the required sequent to the sequence. Otherwise, an ascending subsequence is generated from the sequence as follows: let α'_1 be the member of the sequence with smallest index

[13] Here and in the rest of this paragraph, the + sign does not signify the sum operation but just the stage in which the number is generated by the first generation principle.

[14] Strangely, Zermelo did not refer in his comment also to the proof in *Grundlagen*.

that is greater than α_1, α'_2 be the one with the smallest index that is larger than α'_1, and so on. If the first σ members of the subsequence have been so selected and there are still members in the sequence that are greater than all the members already selected for the subsequence, define α'_σ as the member with the smallest index among them. Lemma 2(i) is used in the generation of the subsequence. It is easy to see that every member of the sequence is surpassed by a member of the subsequence (here Lemma 2(ii) is used). The subsequence will be indexed by a number not greater than κ so also not greater than ω_γ. Adding to the subsequence all the numbers that lie between any two consecutive members of it and the numbers preceding its first member, a succession is obtained. If we can demonstrate that the succession has the power of ω_γ then its sequent would fulfill the Limitation Principle and be a member of (γ). Observing that the succession is composed of up to ω_γ partitions, each with up to ω_γ numbers, we can conclude that the succession is equivalent to a subset of the set of all ordered-pairs of numbers smaller than ω_γ, the power of which is equal to the power of ω_γ by the Union Theorem for U_γ proved above. The axiom of choice is necessary here to choose a mapping from each partition to ω_γ. Noting that the succession contains at least one subset of power ω_γ (the partition defined by the first number of the subsequence), the desired result follows by CBT for ω_γ proved above.

Note that in the proof of Lemma 1 for γ, the Union Theorem for γ and CBT for γ are used as well as AC. In the proof of the Union Theorem for γ the Fundamental Theorem for $\kappa < \gamma$ (induction hypothesis) and the Different Alephs Theorem for γ are used. Also the Sum Lemma is used and it does not require the induction hypothesis directly, only that the numbers in U_γ comply with the Limitation Principle. The latter holds by the Limitation Theorem for $\kappa < \gamma$ which is the induction hypothesis. In the proof of CBT for γ the Different Alephs Theorem for γ and the Fundamental Theorem for γ are used. In the proof of the Different Alephs Theorem for γ the Limitation Theorem for γ, Theorem 1 for γ and the Sequent Argument are used. In the proof of the Limitation Theorem for γ the Fundamental Theorem for γ and Theorem 1 for γ are used. The proof of Theorem 1 for γ is established by Lemma 1 for $\gamma-1$ (induction hypothesis) and the Sequent Argument. The proof of the Fundamental Theorem for γ is established by Lemma 2 and the Limitation Theorem for all $\kappa < \gamma$ (induction hypothesis). These are the interdependencies of the theorems involved.

It is remarkable how heavy Cantor's proof of CBT is, in contrast to the meager, albeit insightful, mathematics used by most mathematicians who later provided proofs for that theorem. The reason is that Cantor's basic gestalt, by 1883, was that sets are gauged by the powers in his scale of number-classes and he guided his proof by this beacon; the later proofs typically used geometric gestalt. The metaphor of the proofs for the bunch of theorems presented above can be described as helical transfinite induction (cf. Ferreiro's 1995 p 40).

2.3 The Declaration of Infinite Numbers

In Cantor's letter to Dedekind of November 5, 1882, Cantor expressed in a dramatic tone his latest findings:

> [J]ust after our latest visit in Harzburg and Eisenach[15] God almighty saw to it that I attained the most peculiar, unexpected results in the theory of manifolds[16] and in the theory of numbers – or rather have found something which has been fermenting in me for years, and which I have long sought. – It is not a question of the general definition of a point-continuum about which we have spoken and in which I think I have made further progress, but rather about something much more general, and therefore more important.–
>
> You remember I told you in Harzburg that I could not prove the following theorem:
>
> If M′ is a part of a manifold M, M″ part of M′, and if M and M′ can be reciprocally correlated one-to-one [by a 1–1 mapping] (i.e., M and M″ have the same power) then M′ has the same power as M and M″.
>
> Now I have found the source of this theorem and can prove it rigorously and with necessary generality; and this fills a large gap in the theory of manifolds.
>
> I arrive at this result through a natural extension or continuation of the sequence of real integers, . . .

Cantor continued the letter with a summary of his theory of infinite numbers as developed in *Grundlagen*.

It was Cantor's tone in this passage that compelled me to believe that indeed he had a proof of CBT. Here is how we explicate his statements: Cantor may have completed the design of the helical proof presented in the previous section, which at once gives the construction of the scale of number-classes and CBT.[17] In addition, he may have formulated the well-ordering and numbering principles which enable CBT for any set, not just for sets of numbers. Moreover, these principles, on the one hand, provide meaning to the infinity symbols (see the next chapter) of a number-class, which turn out to represent all the possible well-orderings that a set of a certain power can obtain. On the other hand, the principles enable the definition of the arithmetic operations for the infinite numbers, thereby granting them indeed the status of numbers.[18] The background to Cantor's discovery of well-ordering was his realization that a set can obtain many orders. He mentioned this with regard to the continuum in the letter to Dedekind of September 15, 1882, (Cavailles 1962 p 230ff, Dugac 1976 p 255, Ewald 1996 vol 2 p 872). The letter was sent on the

---

[15] Cantor met Dedekind twice during September 1882; first in Harzburg, their favorite vacation retreat, and then in Eisenach, in a gathering of mathematicians.

[16] Manifold (*Mannigfaltig*) was Cantor's term for set (*Menge*) before *Grundlagen*.

[17] The crucial step was probably a scheme for the proof of the Union Theorem.

[18] It does not make much sense to assume that the entire construction of *Grundlagen* was developed in that month, as some writers have suggested, because of the wealth of the ideas and technical details in *Grundlagen*. Cf. Ferreiro's 1995 p 41, Meschkowski-Nilson 1991 p 90 (3).

day Cantor traveled to Eisenach where he met Dedekind, as he recounts in the passage cited above.

The importance of the definition of the arithmetic operations, comes out in the above letter to Dedekind. In about half-way through the letter Cantor added these words regarding the infinite numbers:

> Perhaps you are surprised by my boldness in calling the things $\omega, \omega + 1, \ldots, \alpha, \ldots$ integers, and even real integers of the second class, while I gave them the more modest title infinity symbols when I used them previously.[19]
>
> But my freedom is explained by the remark that among the conceptual things α that I call real integers of the second class there are relations that can be reduced to the basic operations.–

The point is made clearer in Cantor's 1884 letter to Kronecker (Meschkowski-Nilson 1991 p 199, Ferreirós 1995 p 38):

> [the opinion that the transfinite numbers] have to be conceived as numbers is based on the possibility of determining concretely the [arithmetical] relations among them, and on the fact that they can be conceived under a common viewpoint with the familiar finite numbers.

It is clear that Cantor was ready to call his infinity symbols numbers only after he had generalized for them the arithmetic operations. A similar approach Cantor exhibited when he introduced the irrationals by way of the fundamental sequences: he first introduced them as symbols (*Zeichen*) and called them numbers only after defining the arithmetic operations on them (Dauben 1979 p 38f). The operations between infinite numbers Cantor defined in *Grundlagen* §3, after he had introduced the well-ordering principle and the numbering principle (§2).

For the definition of $\alpha + \beta$ Cantor took two sets M, M_1 which have the *Anzahlen* α, β and generated the set M + M_1 in which the members of M are ordered before the members of M_1. Conveniently, M and M_1 can be assumed disjoint. It is easy to see that M + M_1 is well-ordered and its *Anzahl* is then taken to be the sum $\alpha + \beta$. This definition gives the basis for the Sum Lemma which in turn proves that the sum complies with the Limitation Principle if α, β do. The well-ordering principle is not necessary for the definition since the well-ordering of M + M_1 can be proved directly, but the numbering principle is necessary to obtain the *Anzahl* of this set. After the definition by abstraction of the ordinal numbers was introduced, the numbering principle was no longer required.

For the definition of multiplication Cantor proceeded in a similar way[20]:

> If one takes a succession, determined by a number β, of various sets which are similar and which are similarly ordered [and pairwise disjoint] and such that each has an *Anzahl* of its

[19] Reference is here made to Cantor's 1880 paper, 1882 paper and 1883 paper, parts 2, 3, 4, in the series *Ueber unendliche, lineare Punkmannichfaltigkeiten*, Cantor 1932 pp 145, 149, 157, respectively.

[20] In 1883 *Grundlagen* Cantor denoted the result of the following definition by $\beta\alpha$, but he reversed the notation in 1895 *Beiträge*. We use the latter convention that prevailed.

> elements equal to α, then one obtains a new well-ordered set whose corresponding *Anzahl* supplies the definition of the product $\alpha\beta$.

Cantor does not argue why the product complies with the Limitation Principle. By the Union Theorem it can be established that the product has the power of the greater of the powers of α, β. Again, it is the numbering principle that is necessary for this definition, until the context is switched to ordinals.

Chapter 3
CBT in Cantor's 1878 *Beitrag*

Cantor's first paper on general set theory was his 1878 *Beitrag*. More precisely, general set theory is dealt with in the first two pages of the paper; the rest of the paper, which consists of 13 pages (no division is explicit in the text) deals with the equivalence of the unit segment and n-dimensional space.

We have seen in Sect. 1.2 that Cantor had stated in 1878 *Beitrag* as easy to prove the denumerable CBT. Our contention in this chapter is that Cantor also had at the time, knowledge of the general CBT. To argue for our contention we will point out two notions introduced in the paper that require, for their proper definition and use, the availability of CBT. We will also see an instance of direct allusion to CBT in the text of 1878 *Beitrag*. We will further produce a proof that Cantor likely had obtained at the time, for the continuum, and discuss a drawing in the paper which gives a heuristic for a general proof.

3.1 Equivalence Classes

In the first paragraph of the paper Cantor introduced the notion of equivalence between sets[1]:

> When two well-defined sets M and N fully and unequivocally, element for element, correlate with each other (which, when it is possible to do in one way, is always possible to do also in many other ways) then the two sets can be said to have equal power or that they are equivalent.

'Well-defined set' Cantor later (1882) defined thus (Cantor 1932 p 150; Hallett 1984 p 45):

> I call a manifold (totality, a set) of elements which belong to some conceptual sphere well-defined, if on the basis of its definition and as a consequence of the logical principle of excluded third[2] it must be seen as internally determined both whether some object

[1] Our translation uses common rather than literal terms (e.g., 'set' not 'manifold').

[2] Hallett translated 'middle'. Brower later implicitly attacked this definition (see Sect. 38.1).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_3, © Springer Basel 2013

> belonging to the same conceptual sphere belongs to the imagined manifold as an object or not, as well as whether two objects belonging to the set are equal to one another or not, despite formal differences in the way they are given.

Towards the end of the 1878 *Beitrag* (p 132), Cantor introduced the continuum hypothesis thus:

> The question poses itself, into how many and which classes the linear sets are partitioned when sets of equal power are put in the same partition and those of different powers are put in different partitions.[3]

The partitions of which Cantor speaks here are equivalence classes for the relation of equivalence (~) between subsets of the continuum. Cantor did not use the term 'equivalence class'. Still, Cantor used equivalence classes, implicitly, already in his 1878 *Ausdehnung* paper, when he equated two irrationals that their defining sequences become infinitely close (p 93). The notation ~, Cantor introduced in the second part of the paper. But there Cantor used variables that range over sets, and so he used ~ between such variables and not between the sets. The use of variables to represent the content of a set, unique to 1878 *Beitrag*, is reminiscent of Frege's 'course of values' notion introduced years later. The customary use of ~ to denote the equivalence relation between sets was introduced in Cantor's 1880 paper on point-sets. Cantor did not prove that ~ is reflexive, symmetric and transitive. The "rhythm" by which it became customary to introduce equivalence relations, was developed only in the 1900s (see Sect. 21.2, Chap. 22; cf. *Zahlen* p 55, Dugac 1976 p 26).

Now, CBT is important to the construction of the equivalence classes for the relation ~, for otherwise, given $C \subset B \subset A$, nothing would exclude the possibility that A, C are in one equivalence class, but B is in another. This would entail that the part is greater than the whole.

Thus, for the statement of his continuum hypothesis Cantor needed CBT at least for denumerable sets and for the continuum. In fact, we believe that without the availability of such proofs Cantor would not have stated the continuum hypothesis.[4] We will discuss the latter proof in Sect. 3.4. Moore (1989 p 87) expressed the view that CBT was "intimately connected in [Cantor's] mind with the continuum hypothesis". However, Moore did not justify his contention and apparently he had not realized the role of CBT in other places of 1878 *Beitrag*, as explicated below.

3.2 The Order Relation Between Powers

In the first paragraph of the paper Cantor also introduced the order relation between powers, together with the comparability of sets assertion:

---

[3] Cantor is not saying here that the continuum is equivalent to the second number-class as Parpart (1976 p 51) implies.

[4] P. Tannery (1884) noted that to establish the continuum hypothesis it is required to prove that the continuum does not contain a subset of greater power. Cantor must have been aware to this point.

> If two sets M and N are not of equal power, then either M has equal power with a subset of N or N has equal power with a subset of M; in the first case we name the power of M smaller, in the second case we name it greater than the power of N.

Cantor defines the order relation to hold between non-equivalent sets, just as he defined the subset relation to hold between non-equal sets: “By a subset of a set M we understand any other set M′, the elements of which are equally elements of M.” The attribute ‘other’ implies that for Cantor, unlike Dedekind, a subset is always a proper subset, not equal to the set. Incidentally, Cantor, again unlike Dedekind (1963 p 45), never explicitly explicated the equality of sets by extensionality.

The comparability of sets assertion implies that the order relation between powers is complete (i.e., any two non-equivalent sets are comparable) unlike the relation of subsets. However, Cantor made no mention of the order relation being transitive (as is the subsets relation), though he surely realized the essentiality of this property for his theory.

How a world without transitivity would look like is well depicted in the drawing “Ascending and Descending” of Escher.[5] To ensure transitivity, the mutual exclusivity (anti-symmetry)[6] of the two cases of power inequality must be included in the comparability of sets assertion. Namely, that it is not possible for M and N to be of non-equal power, while M has equal power with a subset of N and N has equal power with a subset of M. But this assertion is equivalent to the two-set formulation of CBT. Thus CBT is entailed by the comparability of sets assertion with the mutual exclusivity clause, or it must be postulated to warrant that the order relation between powers be transitive. This is the problem situation (Lakatos 1976 pp 6, 144ff) from which CBT emerged.

3.3 A Direct Allusion

A direct allusion to CBT in its two-set formulation, is found in the second paragraph of 1878 *Beitrag*, which speaks of the characterization of infinite sets by their property of being reflexive (a set is reflexive if it is equivalent to one of its proper subset):

> By that fact alone, that an infinite set M is a subset of another set N or corresponds to such by 1–1 mapping, it can in no way be concluded that its power is smaller than that of N; this conclusion is only then justified when it is known that the power of M is not equal to that of N; just as little can the case that N is a subset of M or corresponds to such by 1–1 mapping, be sufficient to gather that the power of M is greater than the power of N.

The last sentence (starting “just as...”) is redundant because it is the same as the first sentence with the letters M and N interchanged. If Cantor did include this last

---

[5] http://en.wikipedia.org/wiki/Ascending_and_Descending

[6] Cf. Kuratowski-Mostowski 1968 p 188.

sentence he must have had a reason. The only reason we can think of is that the passage is designed to hint at CBT in its two-set formulation.[7]

3.4 CBT for the Continuum

A proof of CBT for the continuum can be established upon the following Lemma G of 1878 *Beitrag* (Cantor 1932 p 127):

> If y is a variable, which assumes all values of the interval (0...1) except 0, x a variable which assumes all values of the interval (0...1) without exception, then y ~ x.

In current notation, y ranges over the open-closed interval (0,1] and x over the closed interval [0,1] and the theorem says that (0,1] ~ [0,1].

For the proof of the lemma Cantor used a geometric drawing, which we will discuss in the next section. For our discussion here, however, we will use the analytic interpretation of the drawing (see Dauben 1979 pp 64, 323 note 46; Ferreirós 1999 p 193): Cantor divided the interval [0,1], imagine it drawn on the x-axis, into the sequence of (open) intervals (0,½), (½,¾), ..., (1-$½^n$,1-$½^{n+1}$), ..., the set {0, ½, ..., 1-$½^n$,... } and the set {1}. The interval (0,1], drawn on the y-axis, is partitioned similarly with only the point 0 omitted from the denumerable set of discrete points. Every α in the interval (0, ½) of [0,1] is correlated to ½-α of the (0,1] interval; if α belongs to (1-$½^{\text{n}}$,1-$½^{\text{n+1}}$) of [0,1] it is correlated to 1-$½^{n+1}$-(α-(1-$½^n$)) of (0,1]. The members of the discrete set of points from [0,1] are correlated each to its corresponding member in the set of discrete set of points from (0,1], the first to the first, the second to the second, etc. Thus 0 from the first set corresponds to ½ in the second set. The number 1 of [0,1] correlates to the number 1 in (0,1]. So, the intervals correlate by 1–1 mapping.

For later reference we note that the discrete set of points of [0,1] is a simple chain in Dedekind's *Zahlen* (1963) terminology, and its mapping to the discrete set of points in (0,1] can be depicted as pushing 0 down the chain. We see that Cantor employed the idea of simple chains and the pushdown metaphor years before he was aware of Dedekind's work in *Zahlen* (see Chap. 9).

Obviously the above construction can be ported to any unit segment and its open-closed counterpart in the plane with abscissa β in [0,1]. Since Cantor demonstrated in 1878 *Beitrag* that the unit square is equivalent to the unit segment, these segments, being disjoint, correspond to disjoint subsets of [0,1] under the mapping from the unit square to the unit segment. Denote the image of the open-closed part of these segments by S_β.

Let us now have the conditions of the single-set formulation of CBT with M of power of the continuum. Without loss of generality we can assume that M is [0,1].[8]

[7] It is possible that an earlier draft of 1878 *Beitrag* was more specific regarding CBT and what we see as hints, are traces of the older version.

[8] For the familiar mapping of the line onto the unit segment see Fraenkel 1966 p 49.

We can further assume without loss of generality that for every S_β with β in M-M′, S_β resides in M″. Then we can define a 1–1 mapping from M onto M′ by using Cantor's mapping of $\{\beta\} \cup S_\beta$ onto S_β for every such β and the identity on all other points. Thus we obtain a proof of CBT for the continuum. To this proof we suggest the comb gestalt, because of the way the (0,1] segments are standing on the β in [0,1], and the 'folding of the comb' metaphor for the move from the comb to the sets S_β.

In our suggested proof of CBT for the continuum, each of the members of M-M′ is pushed down its chain. A similar idea Dedekind used in his proof of CBT (from July 11, 1887, see Sect. 9.2). Another proof using simple chains was later given by J. Kőnig (in 1906, see Sect. 21.2).

We have no evidence that Cantor was aware of this proof of CBT. However, we have an indication that he did. Dedekind, in his June 22, 1877, response to Cantor's first proof of the equivalence of the unit segment and the unit square, pointed out that the mapping constructed in the proof was not on.[9] As Ferreirós (1999 p 191) remarked, CBT, in its two-set formulation, could serve to complete the proof, for the unit square surely contains a set equivalent to the unit segment. It seems that Cantor was aware of this possibility for in his response to Dedekind of a day later, June 23, 1877 (Cavailles 1962 p 204f; Ewald 1996 vol 2 p 856; Ferreirós 1999 p 191), he said:

> Alas, you are entirely correct in your objection; but happily it concerns only the proof, not its content [!; cf. Lakatos 1978a]. For I proved *somewhat more* than I had realized, in that I bring a system $x_1, x_2, \ldots, x_p$ of unrestricted real variables (that are ≥ 0 and ≤ 1) into one-to-one relationship with a variable y that does not assume all values of that interval, but rather all with the exception of certain y″. However, it assumes each of the corresponding values y′ only *once*, and that seems to me to be the essential point. For now I can bring y′ into one-to-one relation with another quantity t that assumes all values ≥ 0 and ≤ 1.

Cantor clearly expresses himself here as if he held a proof of CBT for the continuum. As we know from the 1878 *Beitrag* paper, Cantor chose not to follow the route he indicated in the letter, perhaps because he did not want to mention CBT for which he did not have at the time a general satisfactory proof.

3.5 Generalized Proofs

The proof above certainly relies on specific features of the continuum. Nevertheless, it can be generalized along the lines of the proof of Dedekind or that of J. Kőnig (see Chaps. 9 and 21). Apparently, Cantor missed this possibility even in face of the hint that Dedekind cached in his *Zahlen* (1963 #63 p 63, see Sect. 9.2). Cantor may have been too focused on his enumeration metaphor and was perhaps blind to

[9] In Dedekind's letter of July 2, 1877 (Cavailles 1962 p 215, Ewald 1996 vol 2 p 863), Dedekind notes that Cantor's mapping in the first proof is continuous!

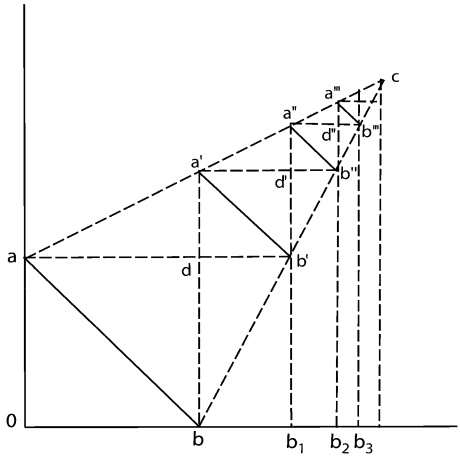


Fig. 3.1 Cantor's 1878 drawing

arguments of more geometric nature. Or maybe not? Here is the drawing that we mentioned above, which first appeared in Cantor's letter to Dedekind of June 25, 1877, (Dauben 1979 p 63; Ferreirós 1999 p 193) (Fig. 3.1).

We can obtain a proof that the two segments ac and bc are equivalent in the following way: notice that sub-segment aa′ is equivalent to sub-segment b′b″ by the lines parallel to ab′; bb′ is equivalent to a′a″ by the lines parallel to ba′. Likewise all the sub-segments of ac and bc are seen to be crisscross equivalent and therefore, because these segments are disjoint (in each sub-segment, only its right endpoint is included, except the first which includes both endpoints), their unions are also equivalent. We can also get the equivalence of aa′ with bb′ by rotating the line ab′ around d, and likewise for the other segments.

This proof captures the essence of later proofs of CBT in that it matches sub-segments, to which we will later call frames. The proof can thus be easily generalized to a proof for the two-sets CBT for any sets by taking the given mappings instead of the parallels used in the drawing-based proof. Such proof was first attempted by Schröder (1898, see Chap. 10), who, however, missed the frames gestalt, and later refined by Schoenflies (1900, see Chap. 12) after the proof given by Borel (1898, basing on Bernstein, see Chap. 11).

The drawing yields also a proof of CBT in its single-set formulation: Consider, for example, the set of all points in the figure oacbo; it is equivalent to the figure acba by mapping the triangles oabo, da′b′d, d′a″b″d′, ... onto each other (the projection from c does it) and all other points by the identity. This proof is similar to Dedekind's proof (see Sect. 9.2) and it can be generalized to any set using the given mapping instead of the projection here.

Cantor could have been aware of all these possibilities but avoided them because of their 'geometric' character. Indeed, it is said that Cantor was averse to geometric intuition (Charraud 1994, see 'intuition' in the index). Indeed, the above drawing is the only drawing in Cantor's writings we have found. Thus his definition of the real numbers (1872 *Ausdehnung*) has no geometric insinuation and he proved in 1874 *Eigenshaft* and 1878 *Beitrag*, theorems that were geometrically counter-intuitive at the time. For Cantor perhaps the theory of infinite numbers had to be developed entirely in an 'arithmetical' fashion, according to the teachings of Weierstrass. Namely, by the scale of numbers and number-classes.

Still, even though in 1878 *Beitrag* Cantor presented a shorter proof for the main result that the unit square is equivalent to the unit segment, which does not require Lemma G for which the drawing serves, Cantor decided to keep the drawing. Cantor must have felt that he needed to explain why he left the drawing after all for he said that the proof of Lemma G by way of consideration of the discontinuous curve (aba′b′a″b″....) is the simplest, to which statement Zermelo (p 127) added a question mark as if exclaiming 'really?!'. Then Cantor added (Cantor 1932 p 129 §6) that he maintained the longer proof (by way of the drawing) because the results obtained thereby "have an interest of their own". The view that the drawing was a simple heuristic to the correspondence between points of the plane and the line, cannot be maintained. The drawing is for the correspondence between [0,1] and (0,1] and not related to any points in the plane. For the proof it is connected to, it is a poor heuristic because it brings too much information. So why did Cantor keep it when he no longer needed it? We believe that the drawing conveys more than what Cantor expressly admitted and he kept it because he wanted to leave a trace of his geometric heuristic trail to CBT.

3.6 The Different Powers of 1878 *Beitrag*

On the side-line of this chapter we wish to answer the question: How many different powers had Cantor identified by the time of the 1878 *Beitrag*?

The opinion in the literature on Cantorian set theory (Dauben 1979 p 79; Moore 1989 p 82; Ferreirós 1999 pp 190, 286; Grattan-Guinness 2000 p 88f) is that when Cantor wrote his 1878 *Beitrag* he knew of only two different powers: that of the set of positive integers N and that of the continuum R, which, as he proved in his 1874 *Eigenschaft* paper, are not equivalent. Under this view it does not make sense that Cantor would introduce the comparability of sets assertion (see the quotation in 3.2), when it is trivial for the powers of N and R. Moreover, Cantor stated in 1878 *Beitrag* (p 120) the Smallest Power Theorem (see Sect. 1.2), which, again, is pointless, if all known powers are just those of R and N.

Against the common view we wish to present in this section evidence that Cantor knew at the time of his 1878 *Beitrag*, not only of the powers of R and N, but at least also of the powers of (II) and of the set of real functions, and thus of the two processes to create many more powers. Assuming that Cantor had only one of the above ideas, sets of functions or number-classes, will not provide an explanation of why Cantor introduced the comparability of sets. Either of these ideas comes within a natural scale. It is only if Cantor had both ideas, and the supposition that subsets of these sets may have still different powers, that a need to assert the comparability of sets would arise.

3.6.1 The Comparison Program

Against Cantor's later papers, the 1883 *Grundlagen* and the 1892 *Frage*, the 1878 *Beitrag* is mainly a compilation of negative results regarding the possibility of finding within Euclidean space (of any dimension) sets of powers different from the power of the continuum or of the set of positive integers. Indirectly this paper teaches us that at the time Cantor was systematically comparing natural mathematical sets for their power relations. To us it seems that 1878 *Beitrag* is a compilation of negative results because it prepared the way for those later papers where sets of new powers were dramatically introduced.

3.6.2 The Set of Real Functions

Cantor defined the exponentiation of powers only in his 1895 *Beiträge* (§4) and he did it by way of the covering set (insertion set) which is the set of all functions from a set representing the power of the exponent to a set representing the power of the base. But already in his letter to Vivanti of 1886 (Moore 1989 p 93f; Meschkowski-Nilson 1991 p 269), with regard to a paper of Paul Tannery from 1884, Cantor used the exponentiation of powers, without any definition, to state that the exponentiation of the denumerable power by itself gives the power of the continuum. Moreover, he says there that this result was known to him already 9 years earlier, namely 1877, when he was writing the 1878 *Beitrag*. A similar claim he made in a letter to Mittag-Leffler from 1884 (Meschkowski-Nilson 1991 p 205). Therefore, under the comparison program, it would have been natural for Cantor to compare the power of the set of all real functions to the power of R. Such a comparison Cantor explicitly made in endnote 10 of *Grundlagen*.[10]

Surely Cantor was aware that a similar procedure of exponentiation could be repeated with regard to the set of real functions, though he may have been skeptical of the value of such a step, for it does not give rise to functions that occur in "nature" (see Chap. 5).

3.6.3 The Set of Denumerable Numbers (II)

We next contend that Cantor knew at the time of 1878 *Beitrag*, of the denumerable infinity symbols, later the numbers that populate (II), that he had the idea to collect them into a set, namely (II), and that he could prove that (II) is not denumerable.

---

[10] A direct proof that the set of all real functions has power greater than the power of R is easy to obtain. Assume that there is a mapping from R to the set of all real functions and let f_x be the function corresponding to x. Let A be the set of numbers such that $f_x(x) = 0$. Let g be the real function that assigns to every member of A the value 1 and to every other number y the value $f_y(y) + 1$. Then there is no x such that $g = f_x$.

In 1878 *Beitrag* Cantor stated the Denumerable Union Theorem and the denumerable CBT. In the 1882 paper he declared that these two theorems, provide for any denumerability proof. Thus Cantor could prove, by the Sequent Lemma (see Sect. 1.2), the denumerability of all the infinity symbols that can be generated using finite numbers and ω, as well as the non-denumerability of (II).[11] That Cantor had the idea to generate a set of all infinite numbers of equal powers, we know from his presentation of the continuum hypothesis, in which he used the equivalence classes[12] of subsets of the continuum of equal powers.

3.6.4 Dating the Infinity Symbols

Let us now compile the evidence that Cantor had obtained by the time of 1878 *Beitrag*, the (denumerable) infinity symbols which, since the 1883 *Grundlagen*, he called the infinite numbers.

In his 1880 paper[13] Cantor introduced (p 147f) what he, since his 1882 paper,[14] called ‘infinity symbols’. There, in the 1880 paper, he stated in a footnote (p 358)[15] that he was led to the process generating these symbols 10 years earlier, namely in 1870 (Grattan-Guinness 1970 p 69; Ferreirós 1999 p 160). In his 1883 *Grundlagen* Cantor used the same process (see Sect. 1.1) whereby he defined the infinity symbols of the 1880 paper to generate what he now called the infinite real numbers (Ewald 1996 vol 2 p 883 [7]). In a footnote he explicitly associated the infinite numbers with the infinity symbols.[16]

In the mentioned footnote of the 1880 paper Cantor added that in his 1872 *Ausdehnung* paper he hinted at the discovery of the infinity symbols. With hindsight, such hints can be revealed: On p 95 Cantor said that the number concept developed there has the germ [*Keim*] for an absolutely infinite extension (cited in Dauben 1979 p 44, where it is accepted that Cantor obtained his infinity symbols by 1872). Then, on p 98 (Ferreirós 1999 p 160) he said that “here” he is interested only in point-sets whose nth derivative is finite, a remark that hints at Cantor’s later analysis of derived sets of order ∞ or higher.

---

[11] This proof leads to the proof of the Limitation Theorem for (II) which leads to the Limitation Principle and thus to regard Ω as a legitimate infinite number through which the generation of more number-classes could proceed. But Cantor may have lacked at the time a general proof of the Union Theorem.

[12] Cantor did not use the term.

[13] The second paper in the series *Ueber unendliche, lineare Punkmannichfaltigkeiten*, Cantor 1932 p 145.

[14] The third paper of the same series, Cantor 1932 p 149.

[15] Of the original publication in *Mathematische Annalen*, vol 17, not in Cantor 1932.

[16] The footnote appears apparently only in the booklet format of *Grundlagen*. It is in Ewald but neither in Cantor 1932 nor in the original *Mathematische Annalen*.

In the introduction to *Grundlagen* (the booklet; Ewald 1996 vol 2 p 881[1]), Cantor noted that the ideas that guided him in his development of his theory are already to be found in his 1874 *Eigenshaft* and 1878 *Beitrag* papers. He tells there about the length of the process and how his convictions were established gradually (preface to Grundlagen, Ewald 1996 vol 2 pp 881 [3], 890 [11]). The notions of equivalence and the relations of equal/smaller/greater power, though used in *Grundlagen*, are not defined in it, but incorporated (§1) by reference to the 1874 *Eigenshaft* and 1878 *Beitrag* papers. Thus, Cantor presented *Grundlagen* as their continuation.

In a letter to Jourdain dated August 31, 1905 (Grattan-Guinness 1971b p 127), Cantor said that he probably had conceptualized the ordinal numbers in 1877. Obviously, Cantor's use of the term "ordinal number" in relation to 1877 is anachronistic, for Cantor started to use this term only after 1883 *Grundlagen*, but it is clear from the letter that in 1877 Cantor had knowledge of the entities that later he would call ordinal numbers. The discrepancy between the 1905 testimony in the letter to Jourdain and the footnote of the 1880 paper is settled if we understand the earlier dating as relating to the discovery of the denumerable infinite numbers and the later dating as relating to the non-denumerable infinite numbers, enabled with the establishment of (II) in 1877.

With regard to Cantor's dating of his results Ferreirós said (Ferreirós 1999 p 160 footnote 2) that Cantor's word on the dating of his findings "cannot be trusted completely". As an example of wrong dating Ferreirós says that Cantor insisted on the dating of his 1883 *Grundlagen* article to October 1882, in spite of later changes. However, Ferreirós' example is not relevant to Cantor's later dating of the development of his theory. Dating in real-time of his individual papers had for Cantor the importance of establishing his priority, for example, in the early 1880s, vis-à-vis du Bois-Reymond and Harnack, as Ferreirós himself describes (1999 pp 162, 163ff; Dauben 1979 p 93). Retrospective dating was not overloaded with similar concerns. In 1905 it was merely a matter of historical interest to know if Cantor was dating the theory of infinite numbers to 1877 or to 1882.

As final piece of evidence to our contention that Cantor had infinity symbols by 1877, we bring the following remarks on finite and infinite sets that he made in the second paragraph of the paper:

> When a set is contemplated as finite, i.e., it consists of a finite number of [*Anzahl von*] elements, the concept of power[17] corresponds, as it is easy to see, to that of number [*Anzahl*] and consequently to that of the positive integer [*Zahl*]; thus two such sets have the same power if and only if the number [*Anzahl*] of their elements is the same. A subset of a finite set has always a smaller power than that of the set itself; this contention stops entirely at the infinite, that is, sets which consist of an infinite number of [*Anzahl von*] elements.

[17] Cantor slipped here into using 'power' as an entity, outside its defined context in 1878 *Beitrag* of 'equal' or 'smaller/greater' power. Cf. Fraenkel 1966 p 61.

In this passage both the finite and infinite sets are characterized by the number of (*Anzahl von*) their elements and *Anzahl* is regarded as a separate notion from that of power, even for finite sets, for which the two notions correspond to the notion of positive integer. Thus *Anzahl* here appears to refer to the ordinal aspect of *Zahl*, obtained by the numbering of a set. Indeed, compare the above passage with the following excerpts, from *Grundlagen*:

> For finite sets power coincides with the *Anzahl* of elements, because, as everybody knows, such sets have the same *Anzahl* of elements in every ordering. (§1 [11])
>
> A set consisting of infinitely many elements will in general give rise to different *Anzahlen*. (§2 [3])

The quotations from 1878 *Beitrag* and from *Grundlagen* clearly match. Thus, if *Anzahl* for infinite sets is given in *Grundlagen* by the infinite numbers (Dauben 1979 p 101ff), we may apprehend Cantor's *Anzahl* of 1878 *Beitrag* as referring to the infinity symbols.

Ferreirós' remark (1999 p 233 footnote 1, 269f; cf. 1995 §2.1) that Cantor was influenced to separate the ordinal and cardinal character of number by Dedekind in their 1882 meeting, or that Cantor started to define finite and infinite only after the publication of Dedekind's *Zahlen*, is contrary to our view and to the quoted passage (and perhaps even to his own opinion in 1999 p 189). Another objection to our view that can be raised is this: Cantor used "infinite *Anzahl*" in most of his papers from 1872 to 1882 (for example: Cantor 1932 pp 117, 153) and the special meaning of *Anzahl* is not relevant to its use in those places. Our answer is this: Infinite *Anzahl* had a mundane meaning, common among Cantor's readers, before it acquired its special meaning of 1883 *Grundlagen*. Taken literally, the meaning of "infinite *Anzahl*" is "infinite number" which literally involved a contradiction in terms for the pre 1883 *Grundlagen* era (Cantor 1932 p 206). However, from the contexts referenced above it seems that the mundane meaning was simply 'infinitely many' in the sense of 'not finitely many'. This is how Tiles (1989, p 11) has translated *Anzahl*. Also Dedekind explicated infinity by *Anzahl* (Dugac 1976 p 294, Ferreirós 1999 p 109, though he probably used it for cardinal meaning). Thus the above objection does not refute our view that in 1878 *Beitrag* Cantor used *Anzahl* as in 1883 *Grundlagen*. It serves only to point out that even after 1878 *Beitrag* Cantor used *Anzahl* also in its mundane meaning. The separation in 1878 *Beitrag* of the meaning of *Anzahl* for finite numbers from *Zahl*, is enough to testify on the work in progress regarding a change in the meaning of *Anzahl*.

3.6.5 The Berlin Circumstances

If Cantor developed his infinity symbols in the 1870s, the question naturally arises, why didn't he publish anything about them until the 1880 paper? One answer, we believe, is this: Cantor probably presented his early ideas to his teachers Weierstrass and Kronecker in December 1973, during his visit to Berlin. The negative reaction that he received, especially from Kronecker, shattered his confidence in his theory. Cantor wrote to Dedekind about the visit in his letter of December 27, 1873 (Cantor-Dedekind 1937 p 20; Cavailles 1962 p 193). There he mentioned the 'Berlin circumstances' of which he said that he would speak with Dedekind in person (Ferreirós 1999 p 183). Cantor's worry that Kronecker will refuse to print his 1878 *Beitrag* (Dauben 1979 p 69) cannot be explained by the content of that paper, which is mostly occupied with standard mathematical notions. Only if Kronecker knew more of Cantor's infinite numbers theory he could have extracted from the paper what Cantor was trying there to advance. Because of the early rejection of his ideas, Cantor preferred to bring them first under the guise of indices of concrete mathematical objects – the derived sets (cf. Dauben 1979 p 81f).[18]

[18] We note in passing our belief that Cantor had obtained both his infinity symbols and the set named after him (and thus, perhaps, the first example of non-denumerable set), by proof-processing, in 1869–1870, Dirichlet's Condition (Ferreirós 1999 p 150) that the of exception points must be nowhere dense.

Chapter 4
The Theory of Inconsistent Sets

In this and the next chapter we present the complex of ideas surrounding Cantor's theory of inconsistent sets. Our discussion serves to stress the possibility and importance of reading Cantor in his own context. For Cantor the scale of numbers and number-classes was the backbone of his set theory and the theory of inconsistent sets its necessary frame. It is unfortunately common to take towards Cantor's work the Zermelo-Dedekind approach, which by-passes these notions, up to and including the solution of the comparability of sets through the Well-Ordering Theorem. Such approach goes against the way Cantor had developed his theory, both conceptually and historically.

In his August 3, 1899, letter to Dedekind, Cantor defined an inconsistent set as "such that the assumption that all its elements 'are together'[1] leads to contradiction". Though Cantor did not say so explicitly, it appears plausible that such an assumption is made when the set is assumed to have a cardinal number, or an ordinal number (both numbers generated from the set by abstraction as explained in 1895/7 *Beiträge*), or when it is assumed to be a member of a set. Note that though inconsistent sets do not have ordinal or cardinal numbers, there is nothing to prevent us from speaking of the equivalence of such sets and therefore of the equalities or inequalities of their powers, as introduced in Cantor's 1878 *Beitrag*, as long as we do not consider these powers to be among the cardinal numbers generated by abstraction from consistent[2] sets. Thus Cantor speaks, see below, of equivalent inconsistent sets or the projection of one in another.

Cantor's definition of inconsistent sets may appear non-mathematical,[3] but what Cantor is actually saying is that when a contradiction is encountered, a tacit assumption that a certain set is consistent can be detected, and then the argument

[1] In the October 2, 1897, letter to Hilbert (Ewald 1996 vol 2 p 927f) the requirement is that the set be "finished".

[2] Cantor did not use systematically this adjective to qualify the not inconsistent sets.

[3] Hilbert (1904; van Heijenoort 1967 p 131) said: "in my opinion he [Cantor] does not provide a precise criterion for this distinction".

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_4, © Springer Basel 2013

becomes a proof that the identified set is inconsistent.[4] Naturally, Cantor worked under the general thesis that sets of power an aleph do not lead to contradiction. Cantor explicitly stated this thesis, in his letter to Dedekind of August 28, 1899,[5] to be on equal footing with the thesis on the consistency of the natural numbers. This thesis corresponds to the nowadays common thesis of the consistency of ZF.

In the August 3 letter, Cantor proved that the collection of all ordinals, which Cantor denoted there by Ω,[6] is an inconsistent set. Cantor argued (Theorem A in the letter) that if Ω, which is clearly well-ordered as Cantor noted, had an ordinal number, say δ, that number would have to belong to Ω and thus be smaller than itself, which is a contradiction. The contradiction seems to be to the theorem that a well-ordered set cannot be similar to one of its segments.

To segregate the inconsistent sets from the sets, Cantor gave three rules:

> Two equivalent collections either are both 'sets' or are both inconsistent. Every sub-collection of a set is a set.[7] Whenever we have a set of sets, the elements of these sets again form a set.

These rules correspond to the axiom of replacement, axiom of subsets and axiom of union in Zermelo-Fraenkel axiom system (see Frápolli 1990/1991 p 312, Bunn 1980 p 248). However, Cantor did not propose them as existence axioms; he brought them only in a segregation capacity. A similar suggestion was independently offered in 1905 by Harward (see Sect. 18.2). As Cantor's letter became public only in the 1930s, it surely did not influence Zermelo in his development of an axiom system for set theory (Zermelo 1908b).[8]

From Cantor's correspondence with Hilbert at the late 1890s[9] we know that he planned a third sequel to the 1895/7 *Beiträge*. From those letters it seems further that Cantor held back from submitting the third sequel for publication because he had not received Dedekind's reaction to its ideas. We thus take Cantor's letter to Dedekind of August 3, 1899, to be an exposition of the content of the third sequel. However, Dedekind refused to study Cantor's inconsistent sets idea, as is evident from Dedekind's letter to Cantor of August 29, 1899,[10] and he never commented to Cantor on it. Thus, unfortunately, the third sequel was never published. See Chap. 8 for more details on Cantor's attempts to discuss his theory of inconsistent sets with Dedekind in 1899.

---

[4] See Levy's proof (1979 p 11) that Russell's set is inconsistent.

[5] Grattan-Guinness 1974 p 129, Meschkowski-Nilson 1991 p 412, Ewald 1996 vol 2 p 936f.

[6] To be distinguished from Ω of *Grundlagen*.

[7] Hence, a set that contains an inconsistent set is inconsistent.

[8] Cantor did not have a rule that corresponds to the power-set axiom, say, that the set of functions from a set to a set is a set.

[9] Purkert-Ilgauds 1987 pp 154, Meschkowski-Nilson 1991 pp 388–466 passim, Ewald 1996 vol 2 pp 926–930.

[10] Dedekind 1930–32 vol. 3 p 448, Cantor 1932 p 449, Grattan-Guinness 1974 p 129, Dugac 1976 p 261, English translation: Ewald 1996 vol 2 p 937ff. See Sect. 7.4.

Cantor's doubts at his theory, which were perhaps not only because of the need to differentiate two types of sets but also because of the need to introduce a new postulate (see below), persisted past 1899. For while he repeated in his letter to Jourdain of November 4, 1903, his arguments from the August 3 letter to Dedekind, he refused to allow Jourdain to publish that letter (Grattan-Guinness 1971a p 117f), though he encouraged Jourdain to publish his own, similar, results (see Chap. 17).

4.1 Inconsistent Sets Contain an Image of Ω

One immediate result from the segregation rules is that the collection of all alephs is inconsistent, because it is equivalent to Ω. A more subtle result, also given in the letter, is that a set, which cannot be gauged by the scale of alephs, is inconsistent. We cite Cantor:

> If we take a definite multiplicity [*Vielheit*] V and assume that no aleph corresponds to it as its cardinal number, we conclude that V would be inconsistent.
>
> For we generally see that, on the assumption made, the whole system Ω is projectible into the multiplicity V, that is, there must exist a submultiplicity V′ of V that is equivalent to the system Ω.
>
> V′ is inconsistent because Ω is, and the same must therefore be asserted of V.
>
> Accordingly, every transfinite consistent multiplicity, that is, every transfinite set, must have a definite aleph as its cardinal number.

From the last two rows Cantor concluded (Theorem C):

> The system[11] of all alephs is nothing but the system of all transfinite cardinal numbers.[12]

The argument is clearly incomplete because its core lemma, proof that if no aleph corresponds to V (and V is not finite) then Ω is projectible into it, is missing. We call this lemma the 'Inconsistency Lemma'. Zermelo, in his lengthy remark to the theorem (in Cantor 1932), noted this point and suggested that Cantor apparently planned to define, by an inductive process, a mapping from Ω into V, a procedure which requires the axiom of choice since V is not assumed well-ordered.[13]

---

[11] Cantor uses this Dedekindian term 'system' and other terms (e.g., 'multiplicity' above) to avoid referring to inconsistent sets by the term 'set'.

[12] In the attachment for Schoenflies, to his letter of June 28, 1899, to Hilbert (Meschkowski-Nilson 1991 p 403), Cantor noted without any argument that Theorem A follows from the thesis that all cardinal numbers are alephs, which in turn follows from the fact that the collection of all alephs is inconsistent ('not finished' in the terminology of that letter). Schoenflies did nothing with this remark which is indeed quite obscure. It can be deciphered only by reference to the letter to Dedekind.

[13] This procedure was indeed used by Jourdain in his proof of a similar result (1904a p 70, cf. p 67; see Chap. 17).

Thus Zermelo reduced Cantor's discussion to the context of his own set theory whereby the comparability of sets follows from the axiom of choice. Zermelo's remark established the view that Cantor's argument was a failed attempt to prove the Well-Ordering Theorem (Purkert-Ilgauds 1987 pp 67, 139; Ferreirós 1999 p 277, 294f).

Zermelo's suggestion appears sensible, for Cantor often used a similar method (the 'enumeration-by' method) for the construction of mappings, as in the proof of the Fundamental Theorem of *Grundlagen* (§13; see Sect. 1.2). Cantor even used the method when the enumerated set is not well-ordered, as in the proof that every infinite set contains a denumerable subset (1895 *Beiträge* §6), leaning unknowingly on the axiom of choice. However, for the proof of the Inconsistency Lemma, it appears that Cantor was not following the tracks Zermelo was to lay in front of him. Rather, it seems that Cantor intended to postulate Corollary D from §2 of 1895 *Beiträge*,[14] which runs as follows:

> D. If, with two sets M and N, N is equivalent neither to M nor to a subset of M, there is a subset N_1 of N that is equivalent to M.

Here is how we obtain a proof for the Inconsistency Lemma by postulating Corollary D, or even just D′ which is D with M one of the sets U_γ (see Sect. 2.1) and no restriction on N, which can thus be inconsistent. The proof is by transfinite induction. If V is not equivalent to U_2 (I) or any of its subsets, then by Corollary D, U_2 is equivalent to a subset of V. Assume V is not equivalent to U_γ or any of its subsets, then by the induction hypothesis, for every $\kappa<\gamma, U_\kappa$ is equivalent to a subset of V and by Corollary D, U_γ is equivalent to a subset of V. By the Different Alephs Theorem (see Sect. 3.2) all the images of the U_γ are different, though not necessarily disjoint. Now let us denote by S_γ the image of U_γ, and let $S'_\gamma = S_\gamma$-$S_{\gamma-1}$, for γ a successor number. Then the power of S'_γ is equal to the power of (γ). Therewith every number-class has an image in V and all these images are disjoint. Therefore Ω has an image in V, as required. Therefore, V is inconsistent by the segregation rules.

Bernstein, in his recollections on his visit to Dedekind on Pentecost of 1897 (see Sect. 8.2), while repeating the mistaken view that Cantor attempted to prove that every set can be well-ordered, compared Cantor's proof, probably for the Inconsistency Lemma, with Zermelo's first proof (1904) of the Well-Ordering Theorem. The suggested comparison is interesting. Indeed, in both proofs, well-ordered sequences are generated within a set (similarity of metaphor); but whereas in Zermelo's proof all these sequences are compatible, because they are generated using the same choice function, in Cantor's proof they are disjoint images of the number-classes (dissimilarity of gestalt). Still in both cases the union of the strings (same metaphor) provides the required results. However, Zermelo could argue that if the union does not cover the entire set it can be extended while Cantor had no

[14] The context of Corollary D will be discussed in the next chapter. There we will explain our reasons for believing that Cantor intended to postulate Corollary D.

similar argument and so could only obtain that Ω has an image in the set. Note that the axiom of choice enters Cantor's proof too: through his use of the Union Theorem in constructing the number-classes.[15]

4.2 Views in the Literature

Bunn (1980 p 251) considered the possibility that Cantor would use an alternative postulate to Zermelo's suggestion of the axiom of choice (AC). Hallett (1984 p 171) expanded on this possibility and even named the 'projection postulate' Cantor's possible assumption for that proof. Hallett expressed in general a positive view of Cantor's inconsistent sets theory which he considers to be sound and not far from later developments in set theory. Our contribution to the Bunn-Hallett thesis is that the projection postulate is Corollary D.

Zermelo raised an additional objection against Cantor's proof of Theorem C: that the proof may involve use of "'inconsistent' multiplicities, indeed possibly contradictory notions, and is logically inadmissible already because of that". This objection was totally off mark. Zermelo ignored the change in Cantor's views from regarding an absolutely infinite collection as a notion external to mathematics,[16] to regarding it as an inconsistent set – a mathematical notion (Jané 1995 p 388f). It was not Cantor's aim "to restrict his set theory to consistent multiplicities" (Grattan-Guinness 1977 p 28)[17]; he only barred the inconsistent sets from the domain of certain operations on sets, such as the operation of abstraction which generates the ordinal and cardinal numbers.

With Zermelo, Purkert-Ilgauds (1987 p 156) regard Cantor's use of inconsistent sets as unacceptable and explain by Cantor's Platonism his calm in face of the antinomies. The correct view is that under the theory of inconsistent sets the antinomies do not arise. Purkert-Ilgauds view is shared by several writers (Ferreirós 1999 p 321 footnote 3) who look upon Cantor's theory from the vantage point of Zermelo's Well-Ordering Theorem. Dauben (1979 p 245) sees clearly that Cantor restricted the inconsistent sets from the operation of abstraction but cannot bring himself to accept them into set theory. Bunn (1980 p 252) maintains that Zermelo's criticism of "contradictory notions" would have been more devastating to Cantor's

[15] Note that in Zermelo's axiomatic set theory the axiom of choice applies to consistent sets only. In the next chapter we will show how the Comparability Theorem for cardinal numbers follows from Theorem C above. Thus Cantor obtained the Well-Ordering Theorem for consistent sets. But with Corollary D as postulate, Cantor also obtained the comparability of all sets, namely, also that of the inconsistent sets.

[16] Cantor 1932 p 205 endnote 2, Ewald 1996 vol 2 p 916 [2] 2nd paragraph. Cf. Sect. 2.1.

[17] This was Zermelo's aim, which he achieved by limiting his theory to the members of a certain model of his axioms. The members of the model are all consistent sets. The inconsistent sets are the classes in Zermelo's theory, and about them Zermelo did not prove even CBT.

argument had he used an existence axiom (such as corollary D), though nothing contradictory creeps into mathematics if inconsistent sets are allowed. Fortunately, the dissenting views prevailed and inconsistent sets were gradually admitted to set theory under the neutral, non-pejorative name 'classes'. For a rigorous such introduction and some historical background regarding the period after 1899, see Levy 1979, Chapter I §3, 4, Hallett 1984 p 166ff.

In view of this understanding we don't know for what purpose and on what grounds, Zermelo took the idea that "whereas of Cantor's attempt to show that every power is an aleph, apparently it was later recognized by the writer himself not to have been achieved" (Cantor 1932 p 451; this remark was not included in van Heijenoort or Ewald). Cantor's letter to Jourdain of November 4, 1903,[18] indicates otherwise. But then, Zermelo put words also in Poincaré mouth in his 1909 paper, so we believe that his remark may be ignored.

4.3 The Origin of the Inconsistent Sets Theory

Cantor's view that there are two types of sets, one which cannot be numbered and the other which can, appeared already in his review from 1885 of Frege 1884 (Ewald 1996 vol 2 p 924). In letters of the late 1890s and early 1900s, Cantor dated his finding of this distinction to the time of *Grundlagen*. Cf. letters to Hilbert of September 26, 1897,[19] and November 15, 1899,[20] to Jourdain of November 4, 1903, and to G. C. Young of March 9, 1907, (Ewald 1996 vol 2 p 925) where a clear reference to endnotes 1, 2 of *Grundlagen* is given. Thus, Cantor came by the idea of inconsistent sets at the same time he arrived at the idea of the alephs, dated in his 1895 *Beiträge* (§6) to 1882.

Cantor could have proof-processed his proof of the Inconsistency Lemma from Bolzano 1851 and Dedekind's 1872–78 drafts of his *Zahlen*. The drafts of *Zahlen* Cantor read in September 1882 (Ferreirós 1999 p 269). There (Dugac 1976 p 302, *Zahlen* p 105 #159) the theorem is proved that a set containing an image of every finite number contains an image of the set of all finite numbers and so is infinite.[21] Similarly, in Bolzano, infinite collections are characterized by the possibility to project into them every finite set (Bolzano 1851 p 6; Cavailles 1962 p 69; Ewald 1996 vol 1 p 254 §9). We know that Cantor was aware of Bolzano's 1851 work in 1882 from a letter to Dedekind of October 7, 1882 (Dugac 1976 p 256). Cantor said

---

[18] Grattan-Guinness 1971b p 116f, Moore 1978 p 309f, Meschkowski-Nilson 1991 p 433.

[19] Purkert-Ilgauds 1987 p 224, Meschkowski-Nilson 1991 p 388, Ewald 1996 vol 2 p 926f.

[20] Purkert-Ilgauds 1987 p 154, Meschkowski-Nilson 1991 p 414.

[21] The theorem uses infinite choices, as was already noted by Bettazzi (1896 p 512 footnote (1), cf. Moore 1982 p 30, Ferreirós 1999 p 313), who thought that the use of such an axiom is "inappropriate". Bettazzi was a member of the group of mathematicians associated with Peano. The latter was the first to spot and object the use of infinite choices (1890). Cf. Chap. 20.

in that letter that the book was noteworthy because "even though much, perhaps most of it, is erroneous, it has been inspiring for me through the paradoxes which it had excited in me". The only paradoxes that could have troubled Cantor at the time were the paradox of the number of all numbers (Dauben 1991 p 10; Tait 1998 p 11, § 3) and the realization that some sets cannot be numbered. Against the first the Limitation Principle was devised (probably already by 1877); against the second the theory of inconsistent sets was established.

Bernstein, in his referenced recollections, dates the finding of the inconsistency of Ω to shortly before his visit to Dedekind (Pentecost of 1897), when Cantor was working on a proof for the Well-Ordering Theorem, namely, the Comparability Theorem (see the next chapter). This discrepancy with Cantor's own dating can perhaps be explained by what Cantor said in his November 1903 letter to Jourdain, that he had known the theorem that there are no other cardinal numbers than the alephs first intuitively and later in detail. Thus it seems that the detailed proof was obtained when Cantor was writing his 1895/7 *Beiträge*. This evidence matches another in Cantor's letter to Dedekind of July 28, 1899,[22] where Cantor said: "for 2 years now I have been in possession of a proof" that there are no other powers than the alephs.

It is noteworthy that Cantor had no one with whom to confide his ideas on inconsistent sets before the late 1890s. His attempt to bring Dedekind to Halle in the early 1880s was not successful and their correspondence faded (see Chap. 7). Mittag-Leffler was no partner for discussions of new entities, as is obvious from his rejection of Cantor 1970 (Grattan-Guinness 1970). Surely Cantor could not tell of his views in his correspondence with clerics, which was intensive during the same period.[23] In this correspondence Cantor maintained the view that the absolutely infinite is a non-mathematical concept, a symbol of God. If Cantor had a hard time gaining support for his theory, it seems that among his opponents he feared the most (I mean literally feared) the objection of the Catholic Church (perhaps because he was of Marrano origin, cf. Aczel 2000, p 144). Therefore, he was using safe arguments while practicing heresy in the dark.

Interestingly, in 1905, a couple of years after Cantor presented to Dedekind and Hilbert his theory of inconsistent sets, which brings into mathematics the absolutely infinite previously identified by him with God, Cantor published his *Ex Oriente Lux* in which Cantor brought down to human origin God's son (see also Charraud 1994 206ff, 243ff). An almost 10 years earlier indication to this direction can be found in Cantor's letters to Pott from 1896 in which he says that he is accountable before God only.[24]

Note that it was not the move to define the infinite numbers (now ordinal numbers) by abstraction from well-ordered sets, instead of the principles of

[22] Cantor 1932 p 443, Grattan-Guinness 1974 p 127f, Meschkowski-Nilson 1991 p 412, Ewald 1996 vol 2 p 930f.

[23] Hallett 1984 p 166, Purkert-Ilgauds 1987 pp 155–6, Meschkowski-Nilson 1991 pp 252–258.

[24] Purkert-Ilgauds 1987 p 119, Meschkowski-Nilson 1991 p 444.

generation, which led to the inconsistent sets theory. The move to define the infinite numbers by abstraction from the well-ordered sets was motivated by the desire to have a uniform formulation in the general theory of order-types (cf. Cantor 1970). Inconsistent sets were admitted into set theory when Cantor realized the limitation of the Limitation Principle.

4.4 The End of the Limitation Principle

The Limitation Principle was probably conceived in the early days of Cantor's theory of infinite numbers (mid 1870s), when Cantor first looked at the set of all denumerable numbers and asked himself if the number of all these numbers is absolutely infinite. By proving that ω_3 (Ω in *Grundlagen*'s notation) is of the power of (II) Cantor found a simple criterion to bar the number of all numbers from the scale of numbers without restricting the development of the scale of infinite numbers and number-classes. However, the Limitation Principle only barred such sets as the set of all numbers from having a number, not from existence (by comprehension). Instead of banning the inconsistent sets from set theory, Cantor preferred to accept them as mathematical objects under certain restrictions. Unknowingly, Cantor chose Lakatos' content increasing pattern of monster readjustment, of turning a counterexample into an example, over the conservative retreat into a safe harbor approach.

Equipped with the inconsistent sets Cantor had no longer need for the Limitation Principle: the set of all numbers was an inconsistent set and for this reason, not because it is not equivalent to a number-class, it has no ordinal number. With the Limitation Principle Cantor also let go of the well-ordering principle and its ancillary numbering principle. The consistent sets can be well-ordered because their cardinal number is an aleph. For the inconsistent sets there is no reason to require that they be well-ordered. Instead of the dropped principles Cantor was planning to postulate Corollary D.

However, the idea of inconsistent sets seemed too revolutionary so in 1883 *Grundlagen* Cantor preferred the well-ordering principle, the numbering principle and the Limitation Principle.[25] In his following writings he dropped these principles but still avoided mentioning of the inconsistent sets, leaving, in fact, his presentation to be intrinsically incomplete. The negative result of this attitude was the scare of the antinomies, which led to various developments, mainly by Russell, that circled around Cantor's basic inconsistent sets solution. Because he held the notion of inconsistent sets Cantor was not impressed by the antinomies. When he finally did disclose his idea, in letters to Dedekind (in detail), Hilbert, Jourdain and Young, and to Bernstein in person, it made no impression on the public debate.

[25] It is possible that what Cantor found in October 1882 (see the letter to Dedekind cited in Sect. 2.3) was the idea of inconsistent sets and that the other ideas he had obtained earlier.

With the removal of the Limitation Principle and the introduction of definition by abstraction of the ordinals,[26] there was no longer need to establish the theorems of Sects. 1.2 and 2.2, as a prerequisite to the construction of the scale of number-classes. Still, those theorems are necessary to establish the properties of the scale. However, their proofs are no longer interdependent and can assume the scale of numbers and apply the properties of well-ordered sets, introduced in 1897 *Beiträge*, with transfinite induction. Thus a great methodological improvement was brought about with the new definition of the infinite numbers.

[26] Note that even after this move, the principles of generation of 1883 *Grundlagen* are necessary to generate new well-ordered sets, in the sequence of number-classes.

Chapter 5
Comparability in Cantor's Writings

There are two types of comparabilities in Cantor's writings: comparability of sets, referred to also as the comparability of powers, and comparability of cardinal numbers. The distinction between the two types of comparabilities is, however, not clear because Cantor presented the terms 'power' and 'cardinal number' as synonyms. We will outline now how these notions feature in Cantor's writings and how they relate to CBT.

The comparability of powers was asserted in Cantor's 1878 *Beitrag* (see Sect. 3.2). No justification was given to the comparability of sets assertion, as if it is self-evident (Cavailles 1962 p 100). As we explained in Chap. 3, in 1878 *Beitrag* CBT was tacitly assumed either as the mutual exclusivity clause of the comparability of sets or as a postulate, to warrant the transitivity of the order relation between powers.

In his 1883 *Grundlagen*, Cantor abandoned the comparability of sets assertion. Instead he introduced the well-ordering and numbering principles (§2, 3), which entail that every set can be gauged by the scale of number-classes. During the construction of the scale CBT was proved (see Chap. 2). Sets that cannot be gauged by the scale were considered 'absolutely infinite' and were barred outside mathematics.

In Cantor's letter to Mittag-Leffler of April 8, 1883 (Meschkowski-Nilson 1991 p 119), Cantor says:

> When I then wrote it [the comparability of sets assertion of 1878] I was convinced of its validity by an indescribable intuition, but I knew very well that it requires a proof; however, for many years I could not find that proof. Only since I have discovered and found the transfinite numbers, by which the ascending successive powers, which occur in nature, can be defined, I possess a proof.

Apparently Cantor speaks here of the move from the comparability of sets assertion of 1878 to the gauging comparability of 1883.

The well-ordering and numbering principles disappeared from Cantor's writings after *Grundlagen*. In his 1887 *Mitteilungen* (Cantor 1932 p 413), Cantor introduced the notion of cardinal number of a set as an entity obtained by abstraction from the

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_5, © Springer Basel 2013

set, and unfortunately[1] presented the notion of power as synonym to it (p 411). Cantor then (p 413) presented again the comparability of sets, or powers, linking this time explicitly mutual exclusivity and CBT:

> If it is firmly proved that two sets M and N are not equivalent then one of the following cases occurs: either N has a subset N' such that $M \sim N'$ or M has a subset M' such that $M' \sim N$. In the first case $\overline{\overline{M}}$ is called smaller than $\overline{\overline{N}}$ and in the second case $\overline{\overline{M}}$ is greater than $\overline{\overline{N}}$.[2]
>
> Here it cannot be stressed enough that the exclusivity of the two cases, which grounds the definition of greater and smaller cardinal numbers, essentially depends on the assumption that M and N are not of the same power. If the two sets are equivalent it can very well happen that subsets of them M' and N' exist such that $\overline{\overline{M}} = \overline{\overline{N'}}$ as well as $\overline{\overline{M'}} = \overline{\overline{N}}$. We have the theorem: If M, N are two such sets from which subsets M' and N' can be separated for which it can be demonstrated that $\overline{\overline{M}} = \overline{\overline{N'}}$ and $\overline{\overline{M'}} = \overline{\overline{N}}$, then M, N are equivalent.

It is not clear from this passage if Cantor regards CBT as consequence of mutual exclusivity or as a lemma for it. Under the first interpretation, preferred by Dauben (1979 p 172), mutual exclusivity is part of the comparability of sets assertion and CBT is a consequence of it. Dauben sees this interpretation to be in-line with Cantor's presentation of CBT as corollary to the Comparability Theorem for cardinal numbers in Cantor's 1895 *Beiträge* §2 (see below).

Under the second interpretation, preferred by us, the comparability of sets assertion in the 1887 *Mitteilungen* was stated without mutual exclusivity and the exclusivity of the cases was provided by CBT. The second interpretation can hold only if one agrees, with us, that between the 1878 *Beitrag* and the 1887 *Mitteilungen* Cantor indeed obtained a proof of CBT, as demonstrated in Chap. 2.

In his 1895 *Beiträge* §1, Cantor again introduced the notion of cardinal number of a set as an entity obtained by abstraction from the set, synonymous to the notion of power. However, Cantor did not use the notion of inequality introduced in 1878 for powers, with regard to the cardinal numbers. Instead he provided new definitions for these relations (see below). Then (§2), Cantor asserted the Comparability Theorem of cardinal numbers.

As will become clear below, Cantor wrote his 1895 *Beiträge* with an eye to the theory of inconsistent sets. Thus the notions of cardinal numbers, their order relations, the synonymity with the notion of power and the Comparability Theorem, hold only for consistent sets. But Cantor was not ready to give up on the general comparability of sets assertion of 1878 *Beitrag*, and he cached it in the 1895 paper, as we will demonstrate below.

[1] Declaring this synonymity was not necessary to the development of Cantor's theory.

[2] In 1887 *Mitteilungen* Cantor introduced the notation of double over-line on a letter denoting a set, to signify its cardinal number, obtained by the operation of abstraction first introduced in that paper.

5.1 The Definition of Order Between Cardinal Numbers

In 1895 *Beiträge* §1 Cantor argued from the definition of cardinal numbers that two cardinal numbers are equal iff the sets to which they belong are equivalent, as this notion was introduced in 1878 *Beitrag* (see Sect. 3.1). Thus the definition of equal cardinal number of 1895 matches the definition of equal power of 1878.

In §2, Cantor defined the order relation between cardinal numbers by the following:

> When for two sets M and N with cardinal numbers $\mathfrak{a}$ and $\mathfrak{b}$ the following two conditions are fulfilled:
> (1) There is no subset of M which is equivalent with N.
> (2) There is a subset N_1 of N which is equivalent with M.
>
> Then it is first obvious that the same holds true when the sets are replaced by equivalent sets[3] and then that these conditions entail a special relationship between the cardinals of the sets.

Cantor then added that this relation of $\mathfrak{a}$ and $\mathfrak{b}$ is expressed by saying that $\mathfrak{a}$ is smaller than $\mathfrak{b}$ or $\mathfrak{b}$ is greater than $\mathfrak{a}$, in signs:, $\mathfrak{a} < \mathfrak{b}$ or $\mathfrak{b} > \mathfrak{a}$.

This definition of the order between cardinal numbers differs from the definition that Cantor gave in 1878 *Beitrag* and 1887 Mitteilungen to the order relation between powers in that it has (1) instead of the condition that M and N are not equivalent (see Sect. 3.2). With this change, the mutual exclusivity of the cases $\mathfrak{a} < \mathfrak{b}$ and $\mathfrak{b} < \mathfrak{a}$, becomes evident and the transitivity of the order relation between cardinal numbers is easily established, as Cantor gladly announced (Cantor 1932 p 285, Cantor 1915 p 90). Under the earlier definition, the mutual exclusivity had to be assumed, or derived from CBT to which it was equivalent. Hence, in 1895 *Beiträge*, CBT lost part of its 1878 importance.[4]

On the other hand, the new order has the disadvantage that it is necessary to prove for it that either of the cases $\mathfrak{a} < \mathfrak{b}$ or $\mathfrak{a} > \mathfrak{b}$ can occur only when $\mathfrak{a} \neq \mathfrak{b}$, a result which was evident under the previous definition. This disadvantage is, however, minor, as the proof of the required result is simple, and was provided by Cantor without using CBT (Cantor 1932 p 284, Cantor 1915 p 89). Hence, the relations $\mathfrak{a} = \mathfrak{b}$, $\mathfrak{a} < \mathfrak{b}$, $\mathfrak{a} > \mathfrak{b}$, are mutually exclusive.

Explication of the important difference between the two definitions of order, is missing from the literature on set theory, e.g., Dauben 1980 p 207. For consistent sets, the equivalence of the new and old orders is easily proved. Though, for the case that if $\mathfrak{a} < \mathfrak{b}$ under the old definition, when M is infinite, then $\mathfrak{a} < \mathfrak{b}$ under the new definition, CBT is necessary (Fraenkel 1966 p 66). Cantor may have omitted reference to this equivalence to avoid CBT.

---

[3] Thus a cardinal number can be represented by any set of that cardinal number with regard to the order relation between cardinal numbers.

[4] It is still needed to warrant the smoothness of classes of the same cardinal or ordinal number even though those numbers are generated by abstraction – see Sect. 3.1.

5.2 The Comparability Theorem for Cardinal Numbers

In 1895 *Beiträge* §2 Cantor presented for the first time the Comparability Theorem for cardinal numbers (Theorem A):

> A. If $\mathfrak{a}$ and $\mathfrak{b}$ are any two cardinal numbers, then either $\mathfrak{a} = \mathfrak{b}$ or $\mathfrak{a} < \mathfrak{b}$ or $\mathfrak{a} > \mathfrak{b}$.

Because of the mutual exclusivity of the cases was already proved, the essence of the Comparability Theorem for cardinal numbers is that no fourth case exists. With regard to a proof of Theorem A Cantor said:

> Not until later, when we shall have gained a survey over the ascending sequence of the transfinite cardinal numbers and insight into their connection, will result the truth [*Wahrheit*] of the theorem.

It is reasonable to understand that the necessary insight to be gained before the proof can be presented is that of the scale of number-classes.[5] Many writers understood Cantor's reason for deferral of the proof of Theorem A as we do: Medvedev (1966 p 232), Mańka-Wojciechowska (1984 p 192), Hallett (1984 p 67), Ferreirós (1999 p 289). This understanding gains support from the context of the letter to Dedekind of August 3, 1899. There Cantor returned to the above Theorem A, following Theorem C of the letter. Cantor said:

> We further recognize out of C the truth [*Richtigkeit*] of the stated theorem[6] [Cantor references 1895 *Beiträge* and cites Theorem A]. For, as we have seen, these relations of magnitude obtain between the alephs.

Cantor echoes in this passage the peculiar language he used in 1895 *Beiträge*, namely, he speaks in both places of the emergence of the truth of the Comparability Theorem for cardinal numbers and not of its proof. From the similarity of language between the texts we learn that in 1895 Cantor already planned the development in the letter. Thus we must accept that Theorem A was construed from the beginning for consistent sets only, and that its validity was planned to be based on Theorem C of the letter and the comparability of the alephs, which stems from the comparability of the ordinals.

Though Cantor's letter to Dedekind of August 3, 1899, was published in 1932, the view that became commonplace by then, that Cantor had not kept his promise to present a proof of Theorem A because he could not prove the Well-Ordering Theorem, is still maintained. Thus, Fraenkel mentioned in 1930 (p 257) the unfulfilled promise, and he repeated the same in 1966 (p 68). This view, as we have seen, is wrong. Cantor did not promise to prove the Comparability Theorem, but intended all along to base its truth on the comparability of the alephs.

---

[5] This is explicitly stated in Cantor's letter to Mittag-Leffler quoted above.

[6] Our translation; Ewald's is "C makes it clear that I was right when I stated [1895] the theorem". The original is "*Wir erkennen ferner aus C die Richtigkeit des* [reference to 1895 *Beiträge*] *ausgesprochenen Satzes*".

One may ask: why did Cantor bring the Comparability Theorem at such an early stage of the presentation of his theory and did not postpone it to the sequel of *Beiträge*? The answer seems to lie in the heuristic value of Theorem A. It serves to show that the cardinal numbers enjoy the comparability property so fundamental for the notion of number.

5.3 The Corollaries to the Comparability Theorem

Following Theorem A (1895 *Beiträge* §2) Cantor gave four corollaries, which "can be very simply derived" from it, of which he said "we will here make no use":

> B. If two sets M and N are such that M is equivalent to a subset N_1 of N and N to a subset M_1 of M, then M and N are equivalent.
> C. If M_1 is a subset of a set M, M_2 is a subset of the set M_1, and if the sets M and M_2 are equivalent, then M_1 is equivalent to both M and M_2.
> D. If, with two sets M and N, N is equivalent neither to M nor to a subset of M, there is a subset N_1 of N that is equivalent to M.
> E. If two sets M and N are not equivalent, and there is a subset N_1 of N that is equivalent to M, then no subset of M is equivalent to N.

The first point to note with regard to the corollaries is that corollaries B, C, E are equivalent to each other; these are simply different formulations of CBT. Corollary B is the two-set formulation, C is the single-set formulation, and E is another formulation of the theorem. That E is equivalent to B and C is not noted in the literature on Cantor's set theory. Zermelo even said that E depends on A (Cantor 1932 p 351 [2]). Because of the equivalence of B, C, E we can speak in what follows of B, as a representative of CBT, and D as the corollaries of A.

The second point to note with regard to the corollaries is that they are formulated in the language of sets and mappings (equivalences), whereas Theorem A is in the language of cardinal numbers. The reason for this change of language is, we believe, that Cantor wanted to hint in this way to the comparability of sets, that is linked with the corollaries. It is for this reason that the corollaries were brought immediately after Theorem A.

5.4 The Comparability of Sets

The theorem corresponding to Theorem A, in the language of sets and mappings, is Theorem A*: If M, N, are any two sets, then either (1) the set M, N are equivalent, or (2) there is no subset of M which is equivalent with N and there is a subset of N which is equivalent with M, or (3) there is no subset of N which is equivalent with M and there is a subset of M which is equivalent with N. Theorem A* is equivalent to

Theorem A, but only for consistent sets that have a cardinal number. Theorem A* implies B, D, without invoking the notion of transfinite numbers.[7]

Theorem A* implies the comparability of sets assertion from 1878 *Beitrag*[8] and is implied by it if the mutual exclusivity of the inequalities in the comparability of sets assertion of 1878 is assumed (or CBT). Moreover, it is also true that B&D → A* and hence that B&D and A* are equivalent.[9] Thus it seems that the difference in language of the corollaries and Theorem A was intended to hint at the cached comparability of sets that follows from postulating B and D.

For the proof of B&D → A* we switch to the language of the propositional logic with the following notations:

Γ denotes the statement 'there is a subset of M equivalent to N';
Δ denotes the statement 'there is a subset of N equivalent to M';
Λ denotes the statement 'M and N are equivalent'.

Negation is denoted by ¬, conjunction by &, disjunction by |, implication by →, meta-equivalence (the observation that two propositions are equivalent) by ≡, = is used to signify that two signs denote the same entity.

We can now formulate A*, B, D as follows:

$A^* = \Lambda|\Delta\&\neg\Gamma|\neg\Delta\&\Gamma$; $B = \Gamma\&\Delta \rightarrow \Lambda$[10]; $D = \neg\Lambda\&\neg\Gamma \rightarrow \Delta$. We have
$B\&D = (\Gamma\&\Delta \rightarrow \Lambda)\&(\neg\Lambda\&\neg\Gamma \rightarrow \Delta) \equiv (\neg(\Gamma\&\Delta)|\Lambda)\&(\neg(\neg\Lambda\&\neg\Gamma)|\Delta) \equiv$
$(\neg\Gamma|\neg\Delta|\Lambda)\&(\Lambda|\Gamma|\Delta) \equiv (\Lambda|\Gamma|\Delta)\&(\Lambda|\neg\Gamma|\neg\Delta) \equiv$
$\Lambda\&\Lambda|\Lambda\&\neg\Gamma|\Lambda\&\neg\Delta|\Gamma\&\Lambda|\Gamma\&\neg\Gamma|\Gamma\&\neg\Delta|\Delta\&\Lambda|\Delta\&\neg\Gamma|\Delta\&\neg\Delta$.

After eliminating the contradictions from the last expression (the fifth and ninth disjuncts), replacing Λ&Λ by Λ and changing the order of some terms, B&D becomes $\Lambda|\Lambda\&\neg\Gamma|\Lambda\&\Gamma|\Lambda\&\neg\Delta|\Lambda\&\Delta|\Gamma\&\neg\Delta|\Delta\&\neg\Gamma$.

After replacing the second and third disjuncts, as well as the fourth and fifth, by the equivalent Λ, we get that the expression reduces to $\Lambda|\neg\Gamma\&\Delta|\neg\Delta\&\Gamma$, which is Theorem A*.

[7] Note that for finite M and any N and for finite N and any M (which is then necessarily finite), Corollary D can be proved directly using complete induction on the power of the set assumed finite. A proof of Corollary D that avoids comparability and uses instead Zorn's lemma, which has the heuristic advantage of avoiding the notions of infinite numbers or well-ordering, was given in Abian 1963.

[8] In the language of powers, (1) of A* is: M and N have equal power; (2) is: M has greater power than N; (3) is: N has greater power than M.

[9] Cantor noted this point in passing, with regard to Theorem A, in the attachment for Schoenflies to his letter to Hilbert of June 28, 1899 (Meschkowski-Nilson 1991 p 403).

[10] In this notation $E = \neg\Lambda\&\Delta \rightarrow \neg\Gamma \equiv \neg(\neg\Lambda\&\Delta)|\neg\Gamma \equiv (\Lambda|\neg\Delta)|\neg\Gamma \equiv (\neg\Delta|\neg\Gamma)|\Lambda \equiv \neg(\Delta\&\Gamma)|\Lambda \equiv \Delta\&\Gamma \rightarrow \Lambda = B$, so that indeed E and B are equivalent.

Since in the late 1890s it became clear that B, namely, CBT, could be proved without any reference to Cantor's scale of infinite numbers, it appears that by postulating D the comparability of sets assertion can be proved, in addition to the Comparability Theorem for cardinal numbers.

The role of Corollary D in the proof of A* and in the proof of the Inconsistency Lemma, explains why Cantor brought Corollary D at all, when it was not used in his 1895/7 *Beiträge*. Cantor left a hint that he had these results in 1895.

Ferreirós (1999 p 289) contends that Cantor's presentation of A-E in the 1895 *Beiträge* was as open problems. According to our reconstruction, for D alone Cantor did not have a direct proof but it does not make much sense that he presented D as an open problem among established results. And it was unlike Cantor to suggest open problems in his papers.

Chapter 6
The Scheme of Complete Disjunction

In the late 1890s a new gestalt for the presentation of the comparability of sets emerged. Following Fraenkel (1966 p 72f) we call it the scheme of complete disjunction.[1] It was published first by Borel in his 1898 book, in the appendix where he brought Bernstein's proof of CBT. It then appeared in two letters of Cantor from 1899, to Schoenflies and to Dedekind, and in Schoenflies' report of 1900. The scheme is noteworthy because it brought logical analysis, of the propositional calculus kind we used in Sect. 5.4, to what seems to be a pure set theoretic context. We will argue that it was Cantor who developed the scheme, following an analysis of Schröder (1896).

6.1 The Scheme and Schoenflies

In the attachment to Schoenflies, of June 28, 1899, (Meschkowski-Nilson 1991 p 401) Cantor presented the scheme as follows:

Let M and N be any two sets, then one of the following four cases always occurs:

I. No subset M_1 of M is ~ N, however, there is a subset N_1 of N (or even more such) such that $M \sim N_1$.

II. No subset N_1 of N is ~ M, however, there is a subset M_1 of M (or even more such) such that $M \sim N_1$.

III. There exists both a subset M_1 of M such that $N \sim M_1$, as well as a subset N_1 of N, such that $M \sim N_1$.

IV. There exists neither a subset M_1 of M, such that $N \sim M_1$, nor a subset N_1 of N such that $M \sim N_1$.

Cantor then linked the four cases of the scheme to the three cases of the Comparability Theorem for cardinal numbers: cases (I), (II) correspond to the

[1] Fraenkel may have picked the name from Schröder's (1898 p 346) "complete development" (*vollständigen "Entwickelung"*).

inequality cases, and cases (III), (IV) both to the equality case. Cantor mentioned the proofs of Schröder and Bernstein, stressing that Bernstein's proof is independent of logical calculus, and he brought Bernstein's proof (see Sect. 12.1). Following that proof Cantor added that so far a direct proof of case (IV) was not found and that he reckons that such proof does not exist. He then briefly outlined the proof of the Comparability Theorem, which leads indirectly to the conclusion that case (IV) is not a separate case.

On the importance of the scheme in 1899 we have the following testimony of Schoenflies, in his reminiscences of Cantor from 1922 (Schoenflies 1922 pp 101–102):

> I am reminded especially of a question regarding the general construction [of set theory]; the comparability of sets, i.e., the possible relationships between the powers of two sets M and N.[2] He [Cantor] based it on the four logically possible cases by which the subsets of M can stand to N and subsets of N can stand to M. He chose them so that first they exclude each other logically and second that one of them is necessarily realized. Today this method may seem to us self-evident or completely trivial; at that time it was in no way so; it was not included in the above mentioned treatise [the 1895/7 *Beiträge*]. To me it became known by a letter of Cantor[3]; it worked on us in Göttingen as a revelation and it wandered from hand to hand. When we now use analogous procedure to obtain complete conceptual division, we have to thank for this too, Cantor's example and his logically sharp sight.[4]

The logical, rather than mathematical, character of the scheme is what made it so compelling; i.e. that it is not dependent on the mathematical content of the statements involved.[5] Indeed, if we translate the scheme to the language of propositional logic, as in Sect. 5.4, the notions "set", "subset", "equivalent", disappear. The scheme is then: (I) $\Delta\&\neg\Gamma$ (II) $\neg\Delta\&\Gamma$ (III) $\Delta\&\Gamma$ (IV) $\neg\Delta\&\neg\Gamma$. It is easy to see why the cases are mutually exclusive (each two have contradicting conjuncts), and why one of the cases must always occur: the disjunction of the four cases is a tautology, namely, the four cases I–IV are all the cases by which two propositions and their negations can be combined (Table 6.1).

Incidentally, Cantor liked to make disjunctive statements. One example is found in his 1872 *Ausdehnung* paper (Cantor 1932 p 93), where for two converging sequences of rational numbers $\{a_n\}$ and $\{a'_n\}$ Cantor distinguished three mutually exclusive cases: that $a_n - a'_n$ gets infinitely small; that $a_n - a'_n$ remains larger than a

---

[2] Seemingly, Schoenflies was not aware of the intricacies regarding inconsistent sets and cardinal numbers discussed in the previous chapters, even though the subject was hinted in the letter he speaks about.

[3] Schoenflies clearly attributes the scheme to Cantor. The letter he mentions is the one of June 28, 1899. It seems that in Göttingen, by 1899, they have not yet received Borel's book of 1898.

[4] The scheme is used in Kant's division to analytic-synthetic, a priori-aposteriori. So what Shoenflies must have meant is the use of the scheme for mathematical statements. Still, we doubt it that complete disjunction was first applied for the analysis of concepts by Cantor or in Cantor's time. In 1906 Hessenberg refers to the scheme (p 495) as "the familiar logical scheme" of "fourfold disjunction", which, however, could mean the use of the scheme in philosophical discussions.

[5] In addition, we can imagine that the "scandalous" nature of incomparable sets implied by the fourth case, gave the scheme a special flavor which fitted well the air of decadence of the *fin du siècle*.

Table 6.1 Truth table for the scheme of complete disjunction

Γ	Δ	I	II	III	IV	I|II|III|IV
F	F	F	F	F	T	T
F	T	F	T	F	F	T
T	F	T	F	F	F	T
T	T	F	F	T	F	T

certain positive number ε; and that $a_n - a'_n$ remains smaller than a certain negative number $-\varepsilon$. A second example is in the proof from 1874 *Eigenshaft* that the continuum is not denumerable (Ferreirós 1999 pp 181–2). The disjunction is first between the cases that the sequence of nesting intervals is finite or not and in the second case between the case that the limits of the endpoints of the intervals are the same or are different. A third example appears in the proof of the Fundamental Theorem of *Grundlagen* (§13, see Sect. 1.2). A final example is the Comparability Theorem for cardinal numbers of the 1895 *Beiträge*. Disjunctive statements can be classified as a type of "rhythm"; other examples of rhythm in mathematics are: statements of complete induction, Weierstrassian statements of ε, δ and the introduction of equivalence relation.

6.2 The Scheme and Dedekind

A second version of the scheme appeared in Cantor's letter to Dedekind of August 30, 1899.[6] The letter was Cantor's immediate (and enthusiastic) response to Dedekind's letter of a day earlier in which Dedekind presented his proof of CBT by way of his chain theory from his *Zahlen*. Cantor thanked Dedekind for his proof and presented the scheme as follows:

> Two arbitrary sets M and N present from a purely logical standpoint, four mutually exclusive cases:
> (I) There is a subset of N which is equivalent to M, however, there exists no subset of M which is equivalent to N.
> (II) There is no subset of N which is equivalent to M, but there is a subset M_1 of M which is equivalent to N.
> (III) There is a subset N_1 of N which is equivalent to M and there is also a subset M_1 of M which is equivalent to N.[7]
> (IV) There is neither a subset of N that is equivalent to M nor a subset of M which is equivalent to N.[8]

[6] Cantor 1932 p 449, Cavailles 1962 p 247, Grattan-Guinness 1974 p 129.

[7] Here Cantor added a footnote that when he speaks of a subset he always means a proper subset because Dedekind's "subset" included the case when the subset is equal to the set (see *Zahlen* p 46; see footnote to Sect. 3.2).

[8] The scheme to Dedekind differs from the scheme to Schoenflies in that there the subsets are named in all the cases and here only in cases II, III. Also there the order of addressing M and N is interchanged, except in case II. Finally, here the clause '(or even more such)' is omitted.

Cantor pointed out that the first two cases correspond to the inequality parts of the Comparability Theorem and that in the last two cases the sets are equivalent. With the proof of this assertion for case (III), which is CBT, by Dedekind and others, only case (IV) remained unsolved. He thus urged Dedekind to try to prove that in case (IV) the sets are indeed equivalent (and thus finite), by the means (chain theory) which he employed to prove CBT. Cantor himself did not believe that a direct proof of the equivalence of the sets in case (IV) is possible, as he stated in the letter to Schoenflies, but he did not mention his view to Dedekind (see Sect. 7.4). He did say, however, that Schröder expressly said that he does not know how to prove this assertion and that he (Cantor) could demonstrate it only indirectly through Theorem A, namely by shutting out the possibility of a fourth case.

Another possibility which Cantor did not mention to Dedekind is the possibility to obtain the equivalence of the sets in case (IV) by postulating Corollary D. In fact, in our symbolism of Sect. 5.4, what is required to prove is that (IV) $\rightarrow \Lambda$, which is nothing but Corollary D, as the following chain of equivalences shows: (IV) $\rightarrow \Lambda = (\neg\Gamma\&\neg\Delta) \rightarrow \Lambda \equiv \neg(\neg\Gamma\&\neg\Delta)|\Lambda \equiv \Gamma|\Delta|\Lambda \equiv \Lambda|\Gamma|\Delta \equiv \neg(\neg\Lambda\&\neg\Gamma)|\Delta \equiv \neg\Lambda\&\neg\Gamma \rightarrow \Delta = D$. This observation is absent from the literature on Cantor's set theory. Dauben (1980 p 207; cf. 1979 p 242) thinks that Cantor was unable to prove the Comparability Theorem because he could not prove (IV) $\rightarrow \Lambda$ (D), ignoring the possibility that Cantor would postulate D, as his next door neighbor Bunn (1980 p 251) suggested.

6.3 The Scheme and Borel

If we compare two sets A and B, says Borel, "four cases are logically possible and they exclude each other":

I. There is A_1 [subset of A] having the same power as B, and there is no B_1 [subset of B] having the same power as A.
II. There exists no A_1 having the same power as B, and there exists B_1 having the same power as A.
III. There is A_1 having the same power as B, and there exists B_1 having the same power as A.
IV. There exists neither A_1 having the same power as B, nor B_1 having the same power as A.

Borel's scheme differs from Cantor's presentation in the naming of the sets and the interchange of cases (I), (II). The reason for this interchange is that Borel seems to prefer presenting first the positive case of the first set.[9] This pedantic attention, compared to Cantor's rather relaxed style of presentation of the scheme, suggests, in our view, that the scheme did not originate with Borel. Moreover, naming of the subsets in the four cases of the scheme is, in fact, redundant and the logical

[9] We find it intriguing to observe the somewhat obsessive tendency, shared probably by us all, to attempt to put in optimal arrangement a pack of discrete elements with symmetrical or antisymmetrical characteristics.

character of the scheme, which Cantor stressed, comes out without them, as can be seen when the scheme is transposed to propositional calculus terms as in the above truth table.

Following the presentation of the scheme Borel noted that in cases (I), (II) one can say that the power of one of the sets is "superior" to that of the other and that it can easily be demonstrated that this order relation is transitive. Clearly, Borel's introduction of the order relation between powers is less systematic than that of Cantor in his 1895 *Beiträge* (see Sect. 4.3) and this is another reason to the view that Borel was not the originator of the scheme.

With regard to cases (III) and (IV), Borel remarked that it should be proved that in these cases the sets are equivalent. Thus it appears that Borel did not accept Cantor's plan to derive CBT from the Comparability Theorem for cardinal numbers. With regard to case (IV) Borel further added that without its proof, the possibility arises for incomparable sets. He further noted (p 104) that to extract a precise and rigorous proof that in the fourth case the sets are equivalent, seems very difficult because its hypotheses are "purely negative". For this reasons, while Borel stated that he agreed to the use of the notions "equal power" or "greater (smaller) power", he was not ready to accept the notion "power" as an abstract entity. It appears, therefore, that Borel did not accept Cantor's Comparability Theorem as an analysis of the comparability of sets, providing a proof of CBT and Corollary D, and instead he was looking to establish these two theorems independently of the notion of power. Indeed, Borel describes his attitude as follows (1898 p 103f footnote 3):

> These difficulties, for the solution of which I vainly searched Cantor's publications, vividly preoccupied me on diverse occasions, and having had the honor to make the acquaintance of Mr. Cantor in the Zurich congress (August 1897) [the first international congress of mathematicians], I was eager to submit them to him.

It thus appears that Borel identified the comparability of sets (A* of Sect. 5.4) behind the Comparability Theorem for cardinal numbers, before he met Cantor in 1897. But there is nothing to suggest that Borel approached the problem using the logical gestalt that lies behind the scheme. Thus we believe that Borel was introduced to the scheme in his meeting with Cantor in 1897, the description of which Borel gave in the continuation of the cited paragraph:

> He was kind to notify me, for the third case, the theorem that we have just stated [CBT], the demonstration of which terminates this paragraph (see Sect. 11.1); this unpublished demonstration is due to Mr. Felix Bernstein and was given for the first time in the seminar of Mr. G. Cantor at Halle. Regarding the theorem itself, Mr. G. Cantor considers it as correct [*exact*] since a very long time;[10] he was kind to communicate to me verbally that his thought was the same regarding the fourth case [namely, that it entails the equivalence of the sets];

Borel apparently did not consider the possibility of postulating Corollary D. His approach must have been resistive to new axioms. Thus, in 1904, when Zermelo

[10] Because of this statement, Mańka-Wojciechowska (1984 p 192) understood Cantor's words to imply that he was not interested in a direct proof of CBT.

asserted the axiom of choice, Borel remained in the rear-guard of the realists who refused to accept the axiom even on tentative terms (Zermelo 1908a, in van Heijenoort 1967 p 186). Later (Borel 1919) he stretched his criticism, demanding that derived equivalences be given effectively.[11]

6.4 Schröder's Scheme

The first to give a logical formulation to Cantor's comparability of sets was Schröder. In his 1896 note and 1898 paper (p 330), he introduced similar propositions to our Δ, Γ, Λ (his notation: γ, δ, ε) and he asserted (it is the main content of 1896 and appears in 1898 on p 346 formula 70 among several other propositions so that it does not stand out clearly) that Cantor's Theorems A-E[12] can be combined to the proposition $\Lambda \leftrightarrow (\Gamma\&\Delta|\neg\Gamma\&\neg\Delta)$.[13] The combined proposition, admittedly elegant,[14] says that two sets are equivalent when, and only when, either one is equivalent to a (proper) subset of the other, or neither of them is equivalent to a (proper) subset of the other. Let us denote this proposition by P. P can be reduced to a simpler form through the following sequence of equivalences:

$P = (\Lambda \rightarrow (\Gamma\&\Delta|\neg\Gamma\&\neg\Delta))\&((\Gamma\&\Delta|\neg\Gamma\&\neg\Delta) \rightarrow \Lambda) \equiv$

$(\neg\Lambda|\Gamma\&\Delta|\neg\Gamma\&\neg\Delta)\&(\neg(\Gamma\&\Delta|\neg\Gamma\&\neg\Delta)|\Lambda) \equiv$

$(\neg\Lambda|\Gamma\&\Delta|\neg\Gamma\&\neg\Delta)\&((\neg\Gamma|\neg\Delta)\&(\Gamma|\Delta))|\Lambda) \equiv$

$(\neg\Lambda|\Gamma\&\Delta|\neg\Gamma\&\neg\Delta)\&((\neg\Gamma\&\Gamma)|(\neg\Gamma\&\Delta)|(\neg\Delta\&\Gamma)|(\neg\Delta\&\Delta)|\Lambda)$

The first and one before last disjuncts of the second conjunct are contradictions and can be eliminated. So we are left with:

$P = (\neg\Lambda|\Gamma\&\Delta|\neg\Gamma\&\neg\Delta)\&((\neg\Gamma\&\Delta)|(\neg\Delta\&\Gamma)|\Lambda)$. This expression can be further reduced as follows:

$P = (\neg\Lambda\&\neg\Gamma\&\Delta)|(\neg\Lambda\&\neg\Delta\&\Gamma)|(\neg\Lambda\&\Lambda)|(\Gamma\&\Delta\&\neg\Gamma\&\Delta)|(\Gamma\&\Delta\&\neg\Delta\&\Gamma)|$

$(\Gamma\&\Delta\&\Lambda)|(\neg\Gamma\&\neg\Delta\&\neg\Gamma\&\Delta)|(\neg\Gamma\&\neg\Delta\&\neg\Delta\&\Gamma)|(\neg\Gamma\&\neg\Delta\&\Lambda)$

After eliminating the disjuncts that are contradictions we are left with:

$P = (\neg\Lambda\&\neg\Gamma\&\Delta)|(\neg\Lambda\&\neg\Delta\&\Gamma)|(\Gamma\&\Delta\&\Lambda)|(\neg\Gamma\&\neg\Delta\&\Lambda)$.

[11] Russell, who was a realist too, had no problem in the acceptance of a tentative approach to the foundations of mathematics, and he was eventually joined by Gödel (1944 p 449).

[12] Actually it is A*-E but Schröder was not aware of the difference between the comparability of cardinal numbers and the comparability of sets.

[13] Schröder uses = for our ↔, juxtaposition for &, and + for |.

[14] Note that the first disjunct applies to infinite sets while the second to finite sets.

In the 1896 note, the exact relation between P and Theorem A is only sketched; in the 1898 paper, scattered arguments are given (mainly in pp 333, 345, 346). In summary, what comes out of the two papers is the following: There are eight combinations of the three propositions Λ, Γ, Δ and their negations: Λ&Γ&Δ, Λ&Γ&¬Δ, Λ&¬Γ&Δ, Λ&¬Γ&¬Δ, ¬Λ&Γ&Δ, ¬Λ&Γ&¬Δ, ¬Λ&¬Γ&Δ, ¬ Λ&¬Γ&¬Δ. The disjunction of these eight terms is a tautology, as it is easy to verify. Now, when Γ&¬Δ or ¬Γ&Δ,[15] Λ does not hold (see Sect. 5.1), so the two terms Λ&Γ&¬Δ, Λ&¬Γ&Δ are contradictions and can be dropped from the disjunction. By A,[16] the cases ¬Λ&Γ&Δ and ¬Λ&¬Γ&¬Δ cannot occur and so they too can be dropped from the disjunction. What is left is: (Λ&Γ&Δ)| (Λ&¬Γ&¬Δ)| (¬Λ&Γ&¬Δ)|(¬Λ&¬Γ&Δ), which is equivalent to P as was demonstrated above. So, by assuming A, P was obtained.

The left conjunct of P seemingly proves that in cases (III) and (IV) of the scheme, the sets are equivalent. But P is established by A (A*) and from A (A*) B and D follow. Thus Schröder's claim, in his 1896, that B follows from P by a purely logical proof, is pointless. Nevertheless, some writers, notably Schoenflies (1900 p 16), reference the 1896 note for Schröder's proof of CBT instead of the 1898 paper.[17]

To present more clearly the relation between Schröder's disjunction of the eight combinations which leads to P and Cantor's Theorems A-F we offer the following table in which we have extracted from each disjunct three implications. Next to each implication we write its equivalent form as a disjunction. We have indicated where an implication is one of Cantor's Theorems B-F (Table 6.2).

The table is arranged so that each diagonal going down to the right from a cell in row a (with column 1 following column 4 and line a following line c) has in all its implication cells the same disjunction. For example, the diagonal crossing the cells 3.a., 4.b., 1.c., has in all its cells the disjunction Λ|¬Γ|¬Δ.

Our observation that B and E are equivalent is demonstrated in this table by both belonging to the same diagonal. Similarly it is easy to observe that D and (IV) → Λ are on the same diagonal (from cell 2.a.) and therefore are equivalent. Note also that the diagonals show how Γ, Δ, Λ and their negation can be interchanged within the implications without affecting their validity.

This table captures the novelty and power of Schröder's analysis. To appreciate this point, notice the rather arbitrary distribution in the table of the Theorems B-F; true, Cantor did cover all four diagonals in his analysis but not all columns and he did not mention F together with his B-E corollaries as part of one conscious scheme.

[15] Let us call these assertion F> and F<, or plainly F. Thus not only A-E are necessary to establish P but F too; this is clear in the 1898 argumentation.

[16] Actually it is A* but Schröder did not differentiate the two.

[17] Schoenflies does mention the 1898 paper too as 'cf.' to the 1896 source.

Table 6.2 The implications table for complete disjunction

	1.	2.	3.	4.
	$\Lambda\&\Gamma\&\Delta$	$\Lambda\&\neg\Gamma\&\neg\Delta$	$\neg\Lambda\&\neg\Gamma\&\Delta$	$\neg\Lambda\&\Gamma\&\neg\Delta$
a.	$\Lambda\&\Delta \to \Gamma \equiv \neg\Lambda\vert\Gamma\vert\neg\Delta$	$\neg\Gamma\&\neg\Delta \to \Lambda \equiv \Lambda\vert\Gamma\vert\Delta$	$\mathbf{E} = \neg\Lambda\&\Delta \to \neg\Gamma \equiv \Lambda\vert\neg\Gamma\vert\neg\Delta$	$\mathbf{F}<= \Gamma\&\neg\Delta \to \neg\Lambda \equiv \neg\Lambda\vert\neg\Gamma\vert\Delta$
b.	$\Lambda\&\Gamma \to \Delta \equiv \neg\Lambda\vert\neg\Gamma\vert\Delta$	$\Lambda\&\neg\Gamma \to \neg\Delta \equiv \neg\Lambda\vert\Gamma\vert\neg\Delta$	$\mathbf{D} = \neg\Lambda\&\neg\Gamma \to \Delta \equiv \Lambda\vert\Gamma\vert\Delta$	$\neg\Lambda\&\Gamma \to \neg\Delta \equiv \Lambda\vert\neg\Gamma\vert\neg\Delta$
c.	$\mathbf{B} = \Gamma\&\Delta \to \Lambda \equiv \Lambda\vert\neg\Gamma\vert\neg\Delta$	$\Lambda\&\neg\Delta \to \neg\Gamma \equiv \neg\Lambda\vert\neg\Gamma\vert\Delta$	$\mathbf{F}> = \neg\Gamma\&\Delta \to \neg\Lambda \equiv \neg\Lambda\vert\Gamma\vert\neg\Delta$	$\neg\Lambda\&\neg\Delta \to \Gamma \equiv \Lambda\vert\Delta\vert\Gamma$

6.5 The Origin of the Scheme of Complete Disjunction

It is of interest to question how Cantor came by the scheme. Medvedev (1966 p 231) conjectured that Cantor did not have it when he wrote his 1895 *Beiträge*, else he would have used it there. Medvedev further noted the appearance of the scheme in Borel's 1898 book, but he made no connection with Schröder. We suggest that it was from Schröder's scheme that Cantor got the idea to present his comparability of sets in a logical gestalt. What Cantor added to Schröder's analysis was the insight that by removing Λ and $\neg\Lambda$ from the original disjunction members, there is no longer need for mathematical arguments to eliminate four of the original eight disjuncts. The mathematical aspect is reentered when each of the four disjuncts is placed as the premise of an implication for Λ or $\neg\Lambda$, namely: $\neg\Gamma\&\Delta \to \neg\Lambda$, $\Gamma\&\neg\Delta \to \neg\Lambda$, $\Gamma\&\Delta \to \Lambda$, $\neg\Gamma\&\neg\Delta \to \Lambda$.

But how did Cantor learn of Schröder's scheme? Both Cantor and Schröder attended the Zurich first international congress of mathematicians in August 1897 (Peckhaus 1990/1991b p 176) but it is unlikely that Cantor learned there of Schröder's 1896 and processed on the spot his complete disjunction. From Cantor's letter to Dedekind of August 30, 1899, we know that Schröder lectured on his proof of CBT at a conference in autumn 1896, "and 1½ year later" he presented his results in his 1898 paper. Since the 1898 paper appeared after Cantor's conversation with Borel, we must conclude that Cantor got word of Schröder's scheme independently of the 1898 paper.

The most likely assumption is that Cantor attended the 1896 conference and was present in the lecture, perhaps even discussed Schröder's results directly with him. In the letter to Dedekind of August 30, 1899, Cantor said that "Schröder explicitly clarified [*erklärt ausdrücklich*], his inability to prove this theorem [IV $\to$ Λ]", and the language "explicitly clarified" seems to me to indicate personal communication, despite Grattan-Guinness assessment (2000 p 174) that the relations between Cantor and Schröder were not good.

Otherwise it is possible that he was introduced to Schröder's result by letter or a preprint. Of a letter we have no trace. With regard to preprints, we must assume that preprints of the proceedings (JDMV)[18] of the 1896 conference were circulating because in Schoenflies' report on set theory (Schoenflies 1900 p 16 footnote 2), published in 1900 in the proceedings of the conference from 1899, Schröder 1896 is referenced in full, down to the page number. But the 1896 proceedings were published only in 1901! This information appears on the title page of the publication.[19] While the first three volumes of JDMV were published in a timely fashion, in the year following the conference, the fourth volume, for 1894–95, was published only in 1897. Starting from volume five, for 1896, the original publisher of JDMV, Georg Reimer, was replaced by Teubner. With this change of publisher a delay occurred: The JDMV for 1896 was published in 1901 and the JDMV for 1897 was published in 1899, the year the volume for 1898 was published too. So Cantor could have received Schröder 1896 as a preprint.

All the more, since the editors of the 1896 proceedings were A. Wangerin, Cantor's colleague from Halle, and A. Gutzmer, Cantor's best friend. So surely, if there was something important in the proceedings relevant to set theory they would have notified Cantor, in case he did not take part at the conference.

From Cantor's reference to the lecture as the proclamation place of Schröder's proof of CBT one expects that Schröder 1896 addressed this proof but instead it only sketches Schröder's scheme. Because the note is very short (less than half a page; the 1898 paper contains 62 pages) and rather cryptic, unless read in conjunction with the 1898 paper, it is unlikely that Cantor would have learned of Schröder's scheme from the note.

Here a question arises: If Schröder lectured on the proof of CBT in 1896, why didn't he include it in his note? There are two possible answers to this question. First, perhaps because of the problem in the publication of the JDMV of the previous year, Schröder did not want to delay the publication of his results. The second possible answer is that Schröder submitted the 1898 paper for publication in early 1896. The date of completion of the paper is noted as January 1896. Schröder's lecture was then an exposition of the paper rather than the paper being a summary of the lecture. Under both possibilities the note in JDMV appears as an act of obligation fulfillment only. The title of the 1898 paper is similar but not the same as the title of the 1896 note, which is the same as the title of Schröder's lecture.[20]

But the possibility that Cantor derived the scheme from Schröder's papers is not very likely. Schröder's 1898, though much more detailed, is not very readable. From the fact that Cantor concluded that Schröder's proof is very similar to the proof of Bernstein (in his letter to Dedekind of August 30, 1899) it appears that Cantor did not read Schröder closely. Also we know that Cantor resented

[18] Jaresbericht der Deutschen Mathematiker-Vereinigung.

[19] http://www.digizeitschriften.de/dms/toc/?PPN=PPN37721857X.

[20] Incidentally, Schröder's 1898 races with Burali-Forti's papers of the period to be the first paper on Cantor's set theory of the 1895 *Beiträge* and the first paper to apply symbolic logic to Cantorian set theory.

Schröder's sign language (Grattan-Guinness 2000 p 175). Thus our conclusion is that Cantor received Schröder's analysis in a personal communication at the 1896 DMV conference, or there about, and he contemplated his own version of the scheme in the following months.

One other remark, regarding Schröder's lecture and his 1896 and 1898 papers, seems relevant here: In the letter to Dedekind Cantor mentioned that Schröder's lecture was given at the 1896 conference of the *Natureforscherversammlung*; he surely meant the *Gesellschaft Deutscher Naturforscher und Ärzte* (GDNÄ). However, it is more likely that the lecture was given at the annual conference of the *Deutsche Mathematiker-Vereinigung* (DMV) which took place at the same time (September – hence autumn) and place (Frankfurt a. m.) as the conference of GDNÄ. The DMV split from the GDNÄ in 1891 much as a result of Cantor's efforts and he was the first president of DMV (Dauben 1979 p 160ff); so it is strange that he was still using the old name in 1899. Perhaps this slip of the tongue reveals some frustration built in Cantor towards the DMV by 1899, though we do not know of any evidence to justify this conclusion. Peckhaus 2004a (footnote 48), also gives the GDNÄ conference as the place where Schröder's lecture was given.[21] In the proceedings of the DMV conference Schröder's lecture is listed in the list of lectures given at the conference under the title "On G. Cantor's theorems" [*Über G. Cantor'sche Sätze*]. According to Peckhaus a similar listing is given in the proceedings of GDNÄ and Wangerin was also a co-editor of the proceedings of the GDNÄ conference. Perhaps there is some confusion here.

[21] Note that there is a mistake in the text of Peckhaus' paper to which the referenced footnote relates: the year of the conference appears as 1894 instead of 1896. The mistake is repeated in Ebbinghaus 2007 p 90.

Chapter 7
Ruptures in the Cantor-Dedekind Correspondence

Cantor first met Dedekind in 1872, during a summer vacation in Switzerland. Their correspondence, which lasted almost thirty years, is a major source of our knowledge on the development of Cantor's theories. However, as was noted especially by Ferreiro's (1993), the Cantor-Dedekind correspondence suffered from several ruptures, which Ferreiro's described and explained. To the current views regarding the ruptures in the Cantor-Dedekind correspondence, we add the *Aufgabe* complex: Cantor's pattern of approaching Dedekind with a seemingly open problem to which he already had an answer. We believe that Dedekind identified this pattern and resented it. We have noted the *Aufgabe* complex in the context of CBT and for this reason we link the subject to our work.

7.1 Views in the Literature

Cantor's letter to Dedekind, dated November 5, 1882, in which Cantor outlined his theory of transfinite numbers, soon published in his 1883 *Grundlagen*, remained unanswered and the correspondence between the two mathematicians was disrupted for 17 years. Several reasons are given in the literature for this rupture in correspondence; the most common (Grattan-Guinness 1971b p 354, 1974 pp 125–126, Dauben 1979 pp 2–3, Ferreirós 1993 p 356; Charraud 1994 p 160; Belna 2003 p 8) is that it was on Cantor's initiative, as he was disappointed over Dedekind's refusal to come to Halle. However, as Ferreirós (later) pointed out (1999 p 201 footnote 1), Dedekind's refusal was made in the beginning of 1882, while throughout 1882 the correspondence was regular. Moreover, it was Dedekind, not Cantor, who stopped the correspondence, by not answering Cantor's mentioned letter.

From Dauben 1979 (p 110) it may seem that Cantor's criticism of Dedekind's explanation of continuity, in his 1883 *Grundlagen* (Cantor 1932 p 194; Ewald 1996 vol 2 p 906 [8]), could have alienated the latter. However, the paper appeared months after the November 5 letter and Dedekind used to respond to letters within a couple of weeks, at most. More likely is the view that Dedekind wanted to

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_7, © Springer Basel 2013

distance his work from that of Cantor. Still, Dedekind did share with Cantor his draft of *Zahlen* in 1882 (Ferreirós 1999 p 269); so it appears, as Ferreirós pointed out (1999 p 201), that the occurrence of the rupture following the November 5 letter requires further explanation.[1]

But the rupture in correspondence of 1882 was not the first. Already in 1874 the correspondence between Cantor and Dedekind stopped for a couple of years. Then there was another rupture after 1899.

First to identify that there was a rupture in 1874 was Cantor himself who, in a letter to Hilbert of November 15, 1899, said that: "for many years he [Dedekind] has been cross with me, out of reasons unknown to me, and since about 1874 he had broken off our earlier correspondence that began in 1871". Cantor is not accurate here because he only became acquainted with Dedekind in 1872, and there were periods of flowing correspondence between him and Dedekind around 1877 and during 1881–1882; but he puts the finger on the ailing point: the ruptures had their roots in 1874.

To explain the 1874 rupture Ferreirós suggested (1993 p 349) that Dedekind's reason to stop answering Cantor's letters in 1874 was that Cantor failed to mention in his 1874 *Eigenshaft* paper, on the power of the continuum, the contribution that Dedekind made in his letters of the period to the proof that the set of algebraic numbers is denumerable. Ferreirós argument is based on the notes that Dedekind made to himself following the publication of Cantor's 1874 *Eigenshaft* paper,[2] regarding their relevant correspondence, stressing his own contribution to the published results.

7.2 The *Aufgabe* Complex

At the beginning of his letter to Dedekind of November 5, 1882, Cantor mentioned CBT (see Sect. 2.3). Then, at the end of the letter, after presenting summarily the content of the *Grundlagen*, Cantor restated the theorem, with minor changes, as an *Aufgabe*. Ewald translate *Aufgabe* as 'task'. Ferreirós interpreted *Aufgabe* as "research assignment" (Ferreirós 1993 p 360 note 15, 1999 p 239 footnote 4) and he suggested that at the beginning of the letter Cantor thought that he had a proof of CBT but that by the time he finished writing the letter he realized otherwise and therefore he presented the theorem again as an open problem.[3] To us this suggestion does not make much sense because, first, Cantor could have revised the letter, and second, Cantor still maintained in *Grundlagen* (Cantor 1932 p 201; Ewald 1996

[1] In October 1882 Dedekind's mother died (Meshkowski-Nilson 1991 p 85). As she was the center of her family, this may explain why Dedekind did not respond to Cantor immediately, but not why he did not respond at all.

[2] Cavailles 1962 pp 193–196, Dugac 1976 pp 116–118.

[3] A similar view was expressed by Medvedev (1966 p 233).

vol 2 p 912 [13]) his position that he has a proof for the theorem. Instead, we understand "*Aufgabe*" to mean 'assignment' of the kind one gets from a teacher – a homework exercise. Indeed before the conclusion of the theorem in its second presentation, Cantor added the words "to show that" instead of "then", using language which is typical in an exercise.

Now with the *Aufgabe* language Cantor positioned himself as an instructor to Dedekind. However, Dedekind was old enough to have been Cantor's teacher, and as a veteran teacher with a formal character, Dedekind may have found this language as inappropriate and insulting.

To revive the incidence and understand why Dedekind permitted himself to be insulted by Cantor, we must remember that it is only with hindsight that we see Cantor and Dedekind as two giants of similar stature. From the perspective of the 1870s, Dedekind surely regarded Cantor as an upcoming mathematician, with some interesting mainstream results, a taste for oddities and some overlap with his own fields of interest; another product of the Berlin power-house. Dedekind, on the other hand, studied under the Göttingen masters, Gauss, Dirichlet and Riemann, was the writer of important work in many major areas of mathematics and rivaled the masters of Berlin, Kummer and Kronecker; Cantor was not in that league. After the papers on point-sets (including *Grundlagen*), it may have become clear that Cantor treaded on new ground. For this reason his work was translated and republished. However, the importance of his theory was not clear and his association with philosophical and theological language, arguments and friends, may have placed him, in the eyes of his older contemporaries, on the side of the occult. Especially since Cantor hardly published on mathematics after 1883. Only in the late 1890s, after the appearance of his 1895/7 *Beiträge*, did the historical importance of Cantor, as we perceive it today, began to emerge and it became clear that his theory was formidable. Only then, as this theory began to attract the young mathematicians, Cantor gained his historical stature. At that point Dedekind was also influenced to respect Cantor's work and indeed he changed his tone towards Cantor, as his letters of July-August 1899 indicate.

However, there is more to the *Aufgabe* complex: we suggest that what perhaps added to Dedekind's irritation was that Cantor, shortly after presenting CBT as a conjecture in their September meeting, announced its proof in the November letter. To understand our point we must turn to the 1874 rupture.

7.3 The 1874 Rupture

To Ferreirós explanation we wish to add the *Aufgabe* angle. Cantor approached Dedekind with regard to the comparison of the continuum with the positive integers in his letter of November 29, 1873.[4] We contend that Cantor already had the answer

[4] Cavailles 1962 p 187f, Dauben 1979 pp 49–50, Meschkowski-Nilson 1991 p 31, Ewald 1996 vol 2 p 844. Same reference for the next letters.

when he wrote that letter to Dedekind and he was only testing the triviality of the problem; indeed he said, “I cannot find the explanation that I seek[5] perhaps it’s very easy”. Under our view Cantor’s letter of December 2, 1873, written after Dedekind responded that he does not know the answer, is read as a preparation for the presentation of the proof 5 days later. Cantor said: “I put my question to you because I had wondered about it already several years ago… I should like to add that I have never seriously occupied myself with it, because it has no special practical interest for me… But it would be good if it could be answered; e.g., if it could be answered with no, then one would have a new proof of Liouville theorem that there are transcendental numbers”.[6] Then, in the letter of December 7, 1873, no doubt in order to explain how he was able suddenly to come up with the answer, he said: “In the last days I have had the time to pursue more thoroughly the conjecture”. This sudden leisure sounds artificial and we maintain that Dedekind too heard the off-key tone in Cantor’s sequence of letters.

Furthermore, we believe that when in January 1874 (and then in May) Cantor approached Dedekind with the question whether a square can be correlated to a line, Dedekind suspected that the game was played on him again and he did not answer and thereby the 1874 rupture started.[7] Indeed, we see in Cantor’s posing of the latter question in January 1874 a sign that by that time, and probably earlier – even before he posed the question on the denumerability of R, Cantor had the proof for that theorem. It is perhaps because of its earlier nascence that the second part of Cantor’s 1878 *Beitrag* has the odd notation of variables for sets, which is (Ferreirós 1999 pp 192–3) “closer to the preferences [arithmetical presentation] of the Berlin school”. That Cantor was occupied with the question of the relation between R^n and R in 1873 is perhaps evidenced by the way he described the denumerability of the algebraic numbers: In his letters to Dedekind of November 29, 1873, and December 2, 1873, he considered indexed entities $a_{n_1 \ldots n_k}$,[8] with the indexes positive integers, rather than polynomials, as did Dedekind in his response and summary regarding the circumstances of his contributions to Cantor’s 1874 *Eigenshaft* paper (Cavailles 1962 pp 193–196, Dugac 1976 p 116). Cantor’s wider context could have been entities indexed by real numbers, as points in Cartesian space.

---

[5] Here Ewald omitted “*und um den is es mir zu thun*”, which Dauben translated “and while I attach great importance to it” and Cavailles “with it I am preoccupied”

[6] To modern readers this last remark seems trivial: if the continuum, is not denumerable and the algebraic numbers are denumerable then the continuum, is of a different power than the set of algebraic numbers and there are transcendental numbers, in fact, many more than algebraic numbers. But Cantor could not have known all this before he had a proof that the continuum is not denumerable while the algebraic numbers (not mentioned in the letter) are, and before he had time to digest the news that there are different infinities.

[7] Dedekind may have had an excuse because, at the time, he was the director of the institute where he was teaching and surely overwhelmed with non-mathematical work.

[8] Note that an indexed dummy is an alternative to using ordered-tuples.

Our view seems to be contradicted by Cantor's own words regarding the above result, written a couple of years later in his letter to Dedekind of June 20, 1877.[9] There Cantor answered the question he posed in January 1874 and he said, "for years I held the opposite opinion to be true". The years were probably the early 1870s not the years between 1874 and 1877. Because of his anxieties Cantor was sometimes devious in his letters to Dedekind. Cantor thus wrote Dedekind in October 1879 (Dugac 1976 p 125) that he made no notable progress in his research and yet he was then publishing his epoch-making series of papers on point-sets! Dedekind must have felt Cantor's attitude.

Fricke's statement (Scharlau 1981 26) that Cantor "hardly made a step in his work without informing Dedekind and requesting his judgment" has a limited scope: the non-denumerability of the continuum, the equivalence of space and line, the inconsistent sets. Only when he attained unorthodox views did Cantor turn to Dedekind for feedback, but not for guidance in his work.

Ferreirós (1993 p 351) suggested that there was also a 1878 rupture in the Cantor-Dedekind correspondence, lasting nearly 1 year, after Cantor did not mention in his 1878 *Beitrag* Dedekind's contributions made in correspondence during 1877. However, we don't know of any initiative on Cantor's side to communicate with Dedekind during this year and when in early 1879 Cantor sent Dedekind his proof on the invariance of dimension under continuous mapping, Dedekind answered promptly.

Back to the 1882 rupture we can now understand why Dedekind had withdrawn from continuing his communication with Cantor. He was introduced to CBT as a conjecture in September 1882 and probably found a proof for it right away, along the lines of the proof found in his *Nachlass*, dated to July 11, 1887, (see Chap. 9). Then in the November letter Cantor announced that he had found the proof, repeating the pattern of January 1874. It is also understandable why Dedekind, with the experience of January 1874 pointed out by Ferreirós, avoided immediately sharing his result with Cantor. Then, as Cantor proved CBT in *Grundlagen* for sets of the power of the second number-class, and announced that he had a proof of the theorem for any power, there was no point for Dedekind to publish an independent proof: Dedekind surely did not want to tail Cantor's research.

Note that if Cantor had a proof of CBT before September 1882, he had obtained a general proof of the Union Theorem prior to that date. It is for this reason that we have suggested that during October 1882 Cantor had obtained the well-ordering and numbering principles, thereby sealing in his view the arithmetic nature of his infinity symbols (see Sect. 2.3).

9 Dugac 1976 p 119, Ewald 1996 vol 2 p 854, Meschkowski-Nilson 1991 p 41.

7.4 The 1899 Rupture

Something changed in Dedekind's attitude towards Cantor in 1899. Cantor's letter of July 28, 1899, tells of a letter from Dedekind (extinct) in which he congratulated Cantor and his wife for the wedding announcement of two of their daughters. This letter gave Cantor the opportunity to try to reestablish the correspondence with Dedekind regarding his own work, in particular, the issues that he intended to discuss in the third sequel to his 1895/7 *Beiträge*. The content of that sequel he probably presented in his August 3, 1899, letter to Dedekind. From Cantor's letter of August 16, 1899, (Ewald 1996 vol 2 p 936, Grattan-Guinness 1974 p 129) it appears that Dedekind had sent in the meantime another letter (also extinct) to the Cantors in which he congratulated them on their silver anniversary and reported on a coming article of his. Therefore, Cantor could have gathered confidence in Dedekind's willingness to discuss mathematics again with him.

Incidentally, Cantor informed Dedekind of the silver anniversary in his letter of August 3. In the August 16 letter Cantor thanked Dedekind, in the name of his wife too, for his congratulations and said: "The days which we spent with you in Interlaken on our honeymoon 25 years ago remain in our memory always." We therefore doubt Ferreirós judgment that the honeymoon meeting "was probably less than cordial" (Ferreirós 1993 p 350). By the way, the honeymoon meeting took place in late August to early September, as Cantor's telegram of September 13, 1874 (Dugac 1976 p 228), reveals, and not around October as Ferreirós suggested. From the August 16, 1899, letter it appears that Cantor's wife must have dropped her early jealousy of Dedekind (Charraud 1994 p 58).

Dedekind, however, did not respond to the August 3 letter and so Cantor, determined, suggested in a letter of August 28, 1899, that he would come to visit Dedekind for a few days "to discuss these matters". Especially on the need to distinguish between consistent and inconsistent sets, emphasized in the August 16 letter.

This time Dedekind did answer, and immediately, on the very next day (August 29), and his letter, to which he attached a proof of CBT, is quite interesting and so we quote it in full[10]:

> Esteemed friend,
>
> A visit from you will always be welcome to me and my sister, but I am not at all ripe for a discussion of your communication: it would for the time being be quite fruitless! You will certainly sympathize with me if I frankly confess that, although I have read through your letter of August 3 many times, I am utterly unclear about your distinction of totalities into

[10] The letter, without the attached proof of CBT, appeared first in 1932 in Dedekind's collected works, Dedekind 1930–32 vol. 3 p 448; the proof appeared the very same year in Cantor's collected works edited by Zermelo, Cantor 1932 p 449. The proof, translated to French, appears in Cavailles 1962 p 246. The letter appears in Grattan-Guinness 1974 p 129 and in Dugac 1976 p 261. An English translation of the letter and proof, from which we quote, appears in Ewald 1996 vol 2 pp 937–8.

consistent and inconsistent; I do not know what you mean by the 'co-existence of all elements of a multiplicity [collection]', and what you mean by its opposite. I do not doubt that with a more thorough study of your letter a light will go on for me; for I have great trust in your deep and perceptive research. But until now, because of the uninterrupted flood of corrections that must be attended to, I have not had the time or the necessary mental energy to immerse myself in these things. Now [*Jetzt*][11] only revisions lie ahead, and I promised to use the greater peace for this 'immersion'.

When the young Bernstein visited me in Harzburg at Pentecost of 1897 he spoke about Theorem B on page 7 of the translation of Marrote,[12] and was somewhat startled when I expressed my conviction that this theorem was easy to prove with my methods (*Was sind und was sollen die Zahlen?*); but we did not talk further about his or my proof. After his departure I sat down and composed the enclosed proof of Theorem C, which is clearly the same as B. But whether Theorem A can also be proved with the same ease by my methods is a question I have not investigated. In general, for years I have occupied myself extremely seldom with these interesting things, and since my step-by-step mind [*Treppen-Verstand*] has always been slow, it will not now be easy for me to work my way into your research.

With the best greetings to you and your wife
Your devoted
R. Dedekind

The purpose of Dedekind's letter was clearly to dissuade Cantor from his visit plan. Thus Dedekind used a very special tone in the letter, which is quite unlike his tone in all his letters, to Cantor as well as to his other correspondents. He sounds here apologetic, asks for "sympathy" [*nachfühlen*], makes a "confession" [*gestehe*] and mentions his other obligations[13] and his slow "step-by-step" apprehension[14] (*Treppen-Verstand*; the notion appears in Dedekind's first preface to *Zahlen* as a general theory of understanding, Dugac 1976 p 80) as excuses for not dwelling more on Cantor's topics. In using the words 'immersion' [*versenkung*] and 'communication' [*Mittheilung*], Dedekind is echoing Cantor's words in his letter of August 28 where he said: "I hope that you have found the time to immerse [*versenken*] in my communications [*Mittheilungen*] on the system of all cardinal numbers or powers." Thereby, it seems, Dedekind wants to imply that he is attentive to Cantor. Dedekind was always formally cordial, but here and perhaps in the two extinct letters mentioned above, he radiates a more personal cordiality, like when he says that Cantor's visit is always welcome. Moreover, Dedekind appears to be very appreciative of Cantor when he expresses his trust in Cantor's "deep and perceptive research".

Our explanation to Dedekind's changed tone is that Dedekind had now more respect for Cantor after his decisive 1895/7 *Beiträge* which were written in a much more structured style (not unlike Dedekind's own) than Cantor's earlier papers and were rich in mathematical content. In addition, Dedekind may have heard that in the

[11] Ewald has here "But now".

[12] The French translation of Cantor's 1895–7 *Beiträge* paper.

[13] It is not clear to us if the "corrections" and "revisions" Dedekind refers to are to his students' homework or his article on cubic fields due for publication at Crelle's journal which is mentioned in Cantor's letter of August 16.

[14] Dedekind may have picked his attitude from Descartes' "discourse on the method". http://www.literature.org/authors/descartes-rene/reason-discourse/index.html.

first international congress of mathematics held in August 1897 Cantor's praise was sang (Grattan-Guinness 1971b pp 362, 365, but see also 1974 p 126). Not that Dedekind seemed to care much for Cantor's theory of transfinite numbers. As he said, in the above letter, he devoted little time to it. But it seems that he respected the outcome of Cantor's work. The early age and stature differences between him and Cantor were by-now eroded. As a special act of appeasement Dedekind attached to the letter a proof of CBT, thereby showing that he was not totally disinterested in Cantor's work and that when he had the time he was ready to "immerse" himself in it.

However, Cantor was not completely dissuaded by Dedekind's response and in a letter of August 31, 1899 (Cantor 1932 p 448; Cavailles 1962 p 244; Dugac 1976 p 261; Ewald 1996 vol 2 p 939), he scheduled his visit to September 4, for lunch only. The meeting indeed took place and it had affected Dedekind but not quite in the way Cantor anticipated: Dedekind accepted Cantor's criticism of his set of thinkable things (see Chap. 8) but did not embrace the idea of inconsistent sets. Anyway, the meeting was their last contact and the question arises what was the reason for this rupture. It is generally assumed that Cantor's deteriorating mental health prevented him from maintaining any regular correspondence during the period after 1899,[15] yet he was able to communicate with Hilbert, Young and Jourdain in those years, and, as it comes out in Grattan-Guinness' review (1971b p 368), Cantor even taught during considerable time spans until his official retirement in 1913. So our opinion is that the *Aufgabe* complex is again to blame.

The trigger may have been Cantor's letter to Dedekind of August 30, 1899, which was Cantor's response to Dedekind's proof of CBT attached to the letter of the 29th. There Cantor said: "It will be indeed very worthwhile if you would demonstrate also the main Theorem A [the Comparability Theorem for cardinal numbers] with the same auxiliary means", namely, the theory of chains from *Zahlen* used for Dedekind's proof of CBT. There were two problems with this request: First, Cantor's request was nagging because in his letter Dedekind said specifically that he did not investigate if Theorem A can be proved by the same method he used to prove Corollary B and that "it will not now be easy" for him to enter into Cantor's theory. Second, Cantor himself had proved Theorem A by another route, using inconsistent sets, in his letter to Dedekind of August 3. So here was Cantor again sending off Dedekind to missions already completed by himself. Moreover, the direction of research to which Cantor pointed Dedekind, was publicly deemed by Borel to be hopeless. On account of both points, judging from his past behavior, Dedekind could have been predicted to stop communicating with Cantor after his sequence of letters from August 1899.[16]

---

[15] http://www-groups.dcs.st-and.ac.uk/~history/Biographies/Cantor.html.

[16] Admittedly, there is room for a simpler explanation that there was nothing these two mathematicians had to correspond about. Cantor stopped his mathematical research and Dedekind, as is clear from his quoted letter, was not interested in Cantor's transfinite numbers theory, upon which he never commented. (But he also never commented on Cantor's topological papers which were closer to his area of interest.)

One final point to mention here, which may have affected Dedekind's refusal to meet with Cantor. It concerns Bernstein's visit to Dedekind on Pentecost of 1897. The purpose of the visit was to discuss the inconsistency of Dedekind's infinite set from *Zahlen* (see Sect. 8.2). It was thus natural for Dedekind to associate Cantor's request to meet with him to discuss inconsistent sets with the visit of Bernstein. But Dedekind may have had some resentment from the visit which is insinuated in his words "young Bernstein". Bernstein was 19 years old at the time of the visit while Dedekind was 66. On first reading this attribute in Dedekind's letter quoted above the impression is that of sympathy, of the kind that is being aroused in an elder professional when he meets a younger colleague. It gives the elder a certain paternalistic stance, which we all like to assume. Yet we do not think that in the quoted letter this language was only meant by Dedekind to color the anecdote. Rather we think that Dedekind was here tacitly scolding Cantor for sending him someone so young to discuss an intricate philosophical issue. For indeed what level of conversation could Dedekind and Bernstein hold?! Bernstein was then a student of the first or second year at the University of Halle while Dedekind was a teacher for generations and a monumental figure in mathematics. The gap that once alienated Cantor and Dedekind tripled! Had it been a technical question, such as a proof of CBT, surely Dedekind would have had no reservations at Bernstein's age; but the matter at hand was philosophical and required a broad view, maturity and experience, which Cantor and Dedekind had while Bernstein lacked, at least at that time.

Interestingly, Cantor echoed the attribute "the young" with respect to Bernstein, in his next day response letter of August 30, 1899, perhaps unconsciously in order to subdue Dedekind's use of the attribute to its first, sympathetic, interpretation.

7.5 By the Way

We wonder why, in the discussions on the Cantor-Dedekind relationship in the literature, non-mathematical factors are not weighted; as if mathematicians ought to behave only according to their mathematical affinities, disregarding any other human traits. Thus we wonder if the fact that Cantor got married and had children while Dedekind remained a bachelor living with his unmarried sister is not a factor to count. And what effect on their relationship had their differences in wealth (Dedekind described himself as not rich, Dugac 1976 p 127, while Cantor came from a wealthy family), ancestry (Dedekind could name his forefathers several generations back, Scharlau 1981 p 2, a feat that Cantor could not perform), religion and religiousness, political views (Scharlau 1981 p 12), extra-mathematical interests? Therefore, we do not think that a complete picture of the Cantor-Dedekind relationship was depicted yet, if ever it will be.

Chapter 8
The Inconsistency of Dedekind's Infinite Set

We review here the arguments raised by Cantor against Dedekind's infinite set and Dedekind's own doubts on this issue. This chapter still touches the main subject of this book, CBT, on two points: inconsistent sets and Bernstein's visit to Dedekind which brought about Dedekind's proof of CBT (see Chap. 4, Sect. 7.4, Chap. 9).

8.1 Dedekind's Infinite Set

Dedekind began his *Zahlen* (1963 p 44 #1) with the following definition:

> In what follows I understand by *thing* every object of our thought.

Among the things, Dedekind distinguished the sets ('systems' in his terminology) formed by things associated from a certain point of view. As an example of an infinite set, essential for the development of *Zahlen*, Dedekind gave the following (1963 p 64 #66):

> My own realm of thoughts, i.e., the totality S of all things which can be objects of my thought, is infinite.

For every member s of S Dedekind denotes by s′ the thought "s can be an object of my thought". Dedekind maintained that the mapping $\varphi : s \rightarrow s'$ is 1–1 and into, a reflection, so that S is infinite according to his definition of infinite set (a set is infinite if it is equivalent to its proper subset). S′ (the set of s′) is a proper subset of S because "my own ego" is a member of S which is not a thought of a thought and therefore is not in S′.

It appears that Dedekind proof-processed his definition of infinite set from Bolzano 1851 (Dugac 1976 pp 81, 88f and Cavailles 1962 p 125f). Dedekind received a copy of Bolzano's book as a gift from Cantor, attached to Cantor's letter of October 7, 1882. Dedekind's infinite set is missing from the 1872–78 drafts of *Zahlen* (Dugac 1976 p 293, cf. Sect. 4.3).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_8, © Springer Basel 2013

It should be noted here that because φ is a generating principle for possible thoughts its use as the reflection of S into itself seems to be impredicative. Of course, at the time *Zahlen* was written, there was no awareness of the problem of impredicative definitions. Dedekind also used impredicative definition in his definition of the chain generated by a set (1963 #44; Dugac 1976 p 95, see the next chapter). Dedekind may have instinctively felt the problem for he defined S as the "realm of his thought" that seems to pre-exist the repeated application of φ. If S was defined as the sets that contains Dedekind's thought of his own ego and closed under φ, the impredicativity would stand out more clearly. In that case, Dedekind's set corresponds to Zermelo's infinite set (see Sect. 24.2.6).

Frege (1893; Geach-Black 1980 p 129; Gillies 1982 p 52), and Hessenberg, (1906 p 662f; cf. Cavailles 1962 p 126), criticized Dedekind's above definitions for their subjective and psychological nature. However, Dedekind was not impressed by Frege's criticism and remained in his stance, which can be described as mentalist logicism (McCarty 1995 p 86f, Ferreirós 1999 p 244ff). Thus in the preface to the third edition of *Zahlen* Dedekind reconfirms his belief in the creative power of the mind to generate representatives, declared in the preface to the first edition and repeated in his unpublished paper (Sinaceur 1971 p 252) and in an excerpt found in his *Nachlass* (Ewald 1996 vol 2 p 833f). McCarty regards this faculty of the mind as "the first logical operation" in Dedekind's philosophy of mathematics.

Hessenberg also implied that the 'set of all things of thought' is subject to the paradox of the set of all things (1906 p 628), which is Cantor's paradox (see below). Dedekind did not use the set of all things and only assumed, like Zermelo in his 1908b paper on axiomatic set theory, a domain of discourse of all things. However, the difference between his domain and his infinite set is perhaps too subtle to be meaningful and most writers on the subject refer to a generic set, which can mean either of the sets. Thus Zermelo, after he argued with regard to his own domain that it is not a set, using Russell's set, noted that the same argument holds for Dedekind's 'set of everything thinkable' (1908b p 204 footnote 8). Zermelo refers to Dedekind's infinite set but uses the characterization of Dedekind's domain. Zermelo's argument applies to both sets. Likewise Frege, in the endnote to his 1903, (Purkert-Ilgauds 1987 p 149), noticed that Russell's Paradox has on Dedekind's infinite set the same devastating effect it has on his own (Frege's) theory.

Russell (1919 p 139) criticized the assumption that φ is 1–1, claiming that it is not always possible to have an idea (thought) of an object which is different from it or to have only one such idea. For example, if s is the thought of Socrates, s'''' cannot really be regarded as a distinct thought. Russell's view here seems to be realistic rather than formal. Ferreirós (1999 p 246) believes that since Dedekind speaks of possible thoughts in his definition of S, Russell's criticism can be "dispensed with". In view of Ferreirós point one wonders why Dedekind did not add the clause 'possible' to the definition of the domain; perhaps he did not want it to appear temporal.

Ferreirós (1999pp 226f, 234) also stresses the role of Dedekind's domain in Dedekind's definition of set, which goes back to Cantor's explication of his attribute 'well-defined' (*volldefiniert*, see Sect. 3.1) for sets, from his 1882 paper (Cantor 1932 p 150), which refers to a conceptual sphere.

Our purpose here, however, is not to discuss all these opinions in detail; we wish to concentrate on the criticism of Dedekind's infinite set emerging from Cantor's work.

8.2 Bernstein's Recollections

On Pentecost of 1897, it may have been only about 6 weeks after Bernstein found his CBT proof (see Chap. 11), Bernstein, then a student of Cantor, paid Dedekind a visit to convey his master's criticism of Dedekind's infinite set. We know of the visit from Bernstein's recollections of the visit, solicited by the editors of Dedekind's collected works and published there (Dedekind 1930–32 vol 3 p 449; English translation Ewald 1996 vol 2 p 836). We bring the document in full because of its anecdotal interest:

> The visit in question was proposed by Cantor. He had shortly before found the paradox of the set of all ordinal numbers while attempting to prove that every set can be well-ordered – a proof which he tried to carry out with techniques roughly similar to ones Zermelo, avoiding the inconsistent sets, later used in his first proof of well-ordering [the Well-Ordering Theorem]. Cantor was well aware that the discovered paradox could also be applied to the set of all things. Dedekind, in his paper, '*Was sind und was sollen die Zahlen*?' [1963], had used this set to prove the existence of infinite sets, and in such a way that, by the construction of his paper, the definition of the numbers depends on the non-contradictory existence of these sets. Cantor had most likely already written and asked him for his reaction; and since no reaction materialized (probably because of the serious illness of Dedekind in winter 1896–97) he now commissioned me to obtain one by word of mouth.
>
> Dedekind, however, was on that occasion unable to make a conclusive statement, and he said to me that in his ponderings he had almost begun to doubt whether human thought is fully rational.
>
> The following episode should be of special interest. Dedekind said, with respect to the concept of set, that he imagined a set as a closed sack that contains completely determinate things – but things which one does not see, and of which one knows nothing except that they exist and are determinate. Somewhat later,[1] Cantor gave his own conception of a set. He drew his colossal figure upright, made a magnificent gesture with his raised arm, and said, with an indeterminate gaze[2]: 'A set I imagine as an abyss'.

From Dedekind's letter to Cantor of August 29, 1899, we know that the visit took place in Harzburg, on Pentecost of 1897, which fell on Sunday, June 6. From Dedekind's unpublished paper (Sinaceur 1971) we learn that he met Bernstein on

[1] Contrary to this explicit timing, Mostowski, in Lakatos 1967 p 82, who quoted the last paragraph from Becker 1954 p 316, added that the said set descriptions were made in a conversation between Cantor and Dedekind. This additional information is not in Becker and seems to be a mistake.

[2] Ewald has here "staring into the indeterminate".

the 13th. So either Bernstein came to visit Dedekind again after a week or the dating to Pentecost should be taken leniently.

It is not clear from Bernstein's recollections if he went to Harzburg especially to meet Dedekind or if he was visiting there anyway and was asked by Cantor to pay Dedekind a visit. The second option is more likely otherwise some communication should have taken place and if Dedekind was reluctant to respond at the time he would have hardly participated in such. Yet even if there was no preliminary communication there must have been some communication in Harzburg for the visit to take place, unless the Dedekinds held an open house on Sundays, as may have been the custom. The Dedekind family had a house in Harzburg (Scharlau 1981 p 9) so the whereabouts of Dedekind was surely made known to Bernstein by Cantor. However, it was not certain that Dedekind would be in Harzburg: as we know from Cantor's letters to Dedekind of March 7, 1887 (Dugac 1976 p 258) and July 28, 1899, Dedekind was not in Harzburg in August 1886, or on Pentecost of 1899. So it seems that the mission was of tentative nature.

Bernstein's assumption of a letter from Cantor to Dedekind dated before June 1897 is not supported by the known letter exchange between Cantor and Dedekind. Yet it seems likely that there was an earlier attempt on Cantor's side to discuss inconsistent sets with Dedekind. This could have been the March 7, 1887, letter where Cantor mentioned "important mathematical questions that bulged since our last meeting". The last meeting mentioned there was the September 1882 meeting for all that we know. That the notion of infinite sets was discussed in that meeting is evidenced by the attachments to Cantor's letter to Dedekind of October 7, 1882: a promised fraction from Leibnitz where the infinite is discussed and a new finding of Cantor – a copy of Bolzano 1851. What made the question of the infinite to bulge was probably the development of Cantor's theory of inconsistent sets which was perhaps partly proof-processed from Bolzano 1851 (and perhaps the drafts of *Zahlen*, Ferreirós 1999 p 269, see Sect. 4.3). The suggested assumption fits Cantor's statement to Hilbert in a letter of November 15, 1899, that he was aware of the problem with Dedekind's infinite set (that it is inconsistent) immediately when *Zahlen* appeared in Christmas 1887.[3]

Bernstein's use of the term "paradox" (associated with the set of all ordinals) in relation to Cantor's 1897 set theory is anachronistic for this term was introduced into that context by Russell only in 1903 (Moore-Garciadiego 1981). From his use of the term 'inconsistent set' we are led to interpret his referenced paradox as Cantor's result that the set of all ordinals Ω is inconsistent. In his recollections Bernstein seems to hold the view, common at the time it was written as is manifested in Zermelo's remark to Cantor's letter to Dedekind of August 3, 1899, that an inconsistent set is a contradictory notion. This view was not Cantor's view (see Chap. 4). We thus see again that Cantor did not share his views even with his closest associate.

[3] The official publication date of *Zahlen* is 1888 but in Dedekind's unpublished paper (Sinaceur 1971 p 251) Dedekind remarks that *Zahlen* appeared on Christmas 1887.

According to Bernstein, Cantor had noticed that the inconsistency of Ω could be applied to the set of all things (let us denote it by U). Apparently Bernstein too refers to Dedekind's domain. Bernstein then says that this set Dedekind used to prove the existence of infinite sets. But, Dedekind did not assume that his domain is a set nor did he use it to prove that S exists. Therefore, it appears that Bernstein was not very accurate in his description.

Bernstein's meaning of "applied" seems to be that from the inconsistency of Ω, the inconsistency of U can be derived. This could be obtained if Ω is a subset of U because every ordinal is a thing; then as Ω is inconsistent, so is U (because of the segregation rules). Actually, because U contains only thinkable objects, the argument should use the collection of all thinkable ordinals instead of Ω. But most ordinals, just as most transcendental numbers, cannot be described and have no handle to make them thinkable. So most ordinals are not things and the inconsistency of U cannot be proved in this way. The same applies to Dedekind's infinite set S.

Bernstein's mission in his Pentecost of 1897 visit to Dedekind was to obtain "by word of mouth" Dedekind's reaction to the argument that his infinite set is inconsistent. With regard to this mission Bernstein reports Dedekind's doubts on the rationality of human thought. As Dedekind tried to reduce the basic mathematical notions to mental operations of representation (1963 p 33), his conclusion must have been that such a reduction is impossible with regard to the infinite set, which entails that an infinite set must be postulated and in this respect human thinking is not rational. From Dedekind's reaction we must conclude that the argument which considered all ordinals to be thinkable, if this was the argument provided by Bernstein, was not criticized by Dedekind and he accepted that his infinite set is inconsistent. So Dedekind did get Cantor's message already in 1897 (see below for the events in 1899).

On Dedekind's position on axiomatics, and that of his contemporaries, see Ferreirós 1999 pp 119–124, 233f, 246ff. Note that Ferreirós uses the term 'constructivism' for Dedekind's reduction of the numbers to logic, while McCarty (1995) uses this term with regard to intuitionism and maintains (pp 54, 64–5) that Dedekind was not constructivist. Also McCarty claims that Dedekind was revolutionary with his mentalist view (his psychologism), while Ferreirós (1999 p 121f) shows that this approach was developed genetically (evolutionally) in the nineteenth century.

Bernstein's anecdote on how Dedekind and Cantor depicted sets is lovely. We qualify Dedekind's image as algebraic and Cantor's as analytic. Dedekind expresses more control over the subject-matter and his image raises the desire to play with the content of the sack as if it consists of children's building blocks. Cantor faces an unknown that will always remain unknown after any investigation – the continuum. We feel some anxiety in standing next to Cantor facing the abyss (which we visualize in the form of his ternary set). As Dedekind said, á propos Cantor's 1–1 mapping of the unit square upon the unit segment, "you are compelled to admit a frightful, dizzying discontinuity in the correspondence, which dissolves everything to atoms" (Ewald 1996 p 863). Charraud (1994 p 229) associates the sensation produced by Cantor's imagery with the fantasy of the fragmented body. For Cantor himself the abyss seems to have been connected with the sensation that his "red flesh… has always terrible hunger, a true hole without bottom". Cantor used these words when he described to his son his desire to "learn [*apprendre*; we have not seen the German original] the entire world", which Charraud (1994 pp 210–211) describes in the context of Cantor's problems with the Other.

Interestingly, Dedekind's chain construction (see the next chapter) fits more Cantor's image of a set as an abyss than Dedekind's own image of a set as a sack. We find this example of crossed metaphors important for it teaches us that mathematicians are not too orthodox with regard to particular imagery: metaphors gladly coexist and people change and exchange their metaphors as they speak. Note further that Cantor expressed the anxiety of the abyss already in his 1883 *Grundlagen* (§11), where he said with regard to his two generation principles: "the appearance is given that by this way of forming new determinate infinite integers we would have to lose ourselves in the limitless". There he controlled his anxiety by the Limitation Principle that he soon after replaced by the doctrine that inconsistent sets are legitimate (see Sect. 4.4).

Bernstein's anecdote is important for yet another reason: His mention of Cantor's colossal figure (and then Kronecker (the ultimate Other) had a small figure (Dugac 1976 p 252)). We always imagined Cantor as a small person with pointed beard, like the late professor Fraenkel. In addition, the point where Cantor's gesture with his hands is described, is fabulous; it expresses something fundamental in the way Cantor should be perceived.

8.3 Cantor's Criticism

Dedekind's answer through Bernstein did not relate to Cantor's theory of inconsistent sets. Therefore, Cantor, after his correspondence with Dedekind was renewed in July 1899, repeated the message conveyed through Bernstein, in his letter of August 3, 1899. There, after defining what an inconsistent multiplicity is, Cantor said: "As we can readily see, the 'totality of everything thinkable', for example, is such a multiplicity". Cantor is again referencing Dedekind's domain. No argument is provided for the assertion made, so it seems that the wrong argument mentioned in the previous section is applied, namely, that this set is inconsistent because it contains Ω.

In his letter of August 31, 1899, in which Cantor scheduled the meeting with Dedekind to September 4, for lunch, Cantor produced a new example of an inconsistent set, linked to Dedekind's infinite set, and this time gave a full argument why this set is inconsistent. Cantor begins with the "System S* of all thinkable classes $\mathfrak{a}$",[4] where a is the equivalence class for the equivalence relation, namely,

[4] We slightly changed Cantor's notation in this paragraph.

the class of all the sets of power $\mathfrak{a}$.[5] By $M_\mathfrak{a}$ Cantor denotes a representative from the class $\mathfrak{a}$. He then states that S* is not a 'set', namely, that it is inconsistent. For if S* is a set then $T = \{M_\mathfrak{a}\}$ is a set and then $T' = \Sigma M_\mathfrak{a}$[6] is also a set. Here Cantor makes use of his segregation rules of replacement and union. Then, no doubt since T′ is thinkable, T′ is equivalent to one of the $M_\mathfrak{a}$, say to the class $M_{\mathfrak{a}'}$. Then there exists a cardinal number $\mathfrak{a}'' > \mathfrak{a}'$. Cantor references here his 1892 *Frage*, namely, the theorem that bears his name. But since $\mathfrak{a}''$ can be taken as $2^{\mathfrak{a}'}$, $\mathfrak{a}''$ is thinkable and so $M_{\mathfrak{a}''} \in T$. But then $M_{\mathfrak{a}'} \sim M' \subseteq M_{\mathfrak{a}''} \subseteq T'$ so that by CBT $\mathfrak{a}' = \mathfrak{a}''$, a contradiction. It is not likely that Bernstein delivered a similar argument during his visit to Dedekind of 1897, because Bernstein mentioned specifically the application of the inconsistency of Ω in the argument (see Sect. 8.2).

The argument, when applied to Russell's universal set as well, generates what is called Cantor's paradox. The nice thing about this argument is that one can argue that $T \subseteq S^*$, something we could not argue for Ω. T is thinkable because it can be named. So S is inconsistent by one of the segregation rules. In addition, this example can be stated in Dedekind's language of sets and mappings, avoiding the language of ordinals which Dedekind seems to have resented, a resentment that Cantor must have felt. Equivalence classes and their representatives are introduced in *Zahlen* #34. Cantor even used Dedekind's terms 'system' and Cantor's notion 'thinkable' [*Denkbaren*] can be accepted as the equivalent of 'that can be object of my thought'.

In the August 31 letter, Cantor plainly explained to Dedekind what he wants from him:

> if you find me zealous to persuade you of the necessity of dividing 'systems' into two sorts, I hope thereby to show my gratitude for the repeated inspiration and instruction I have received from your classic writings.

Cantor is not criticizing Dedekind's infinite set, nor does he suggest that Dedekind made a mistake in using an inconsistent set, as is the impression we get from Bernstein, which was perhaps Dedekind's impression. Rather Cantor wants Dedekind to accept his view that there are two types of sets: consistent and inconsistent. He does it by raising the doubt in Dedekind's mind that perhaps inconsistent sets exist, hence perhaps his infinite set is inconsistent too. In proclaiming his gratitude to Dedekind Cantor hints at Dedekind's stature as the reason why he seeks his approval. Perhaps he even hints that the discovery could be shared, with Dedekind's *Zahlen* being an example of use of inconsistent sets. If Cantor could convince Dedekind to accept this suggestion, Cantor would have gained incredible support for his theory, turning a counterexample into an example (Lakatos 1976 p 39).

---

[5] The ambiguity whether $\mathfrak{a}$ is a class or a power appears in the original.

[6] $\Sigma M_\mathfrak{a}$ denoting the union of all $M_\mathfrak{a}$.

It is possible that Dedekind could have maintained his theory of *Zahlen* with an infinite inconsistent set instead of an infinite set. But, there was another problem with his theory that bothered Dedekind.

8.4 Dedekind's Concerns

The meeting of September 4, 1899, between Cantor and Dedekind indeed took place.[7] We know this from Dedekind's famous letter to Teubner of 1904 (Landau 1917 p 54, Dugac 1976 p 130, Grattan-Guinness 1971b p 366), in which Dedekind denied his death on that date and told of his meeting with Cantor "who on that occasion dealt a mortal blow not to myself but to a mistake of mine". The blow may have affected Cantor himself more severely than it affected Dedekind for Cantor was suddenly the only authority around and this situation may have triggered his 1899 breakdown (Charraud 1994 p 204ff). Anyway, it appears that Dedekind accepted Cantor's criticism that there is a problem with his infinite set.

Besides the issue of inconsistent sets, another subject was discussed in the meeting. It had to do with Dedekind's avoidance in *Zahlen* from differentiating in symbols between an element and the singleton set which contains it. This subject was brought up by Dedekind who tells about it in an undated and unpublished paper (Sinaceur 1971 p 251; cf. Dugac 1976 p 83; Pla i Carrera 1993 pp 276–278). Lack of attention in applying extensionality (*Zahlen* p 45 #2) in this case, Dedekind noted there, may lead to a contradiction.

In the unpublished paper Dedekind says that he briefly discussed this issue with Bernstein too, during his visit on Pentecost of 1897. The same issue was raised also in a letter from Dedekind to Weber of January 24, 1888 (Dugac 1976 p 273), exactly 1 month after the appearance of *Zahlen* and thus much earlier than Frege's criticism on this subject (Frege 1893; Geach-Black 1980 p 129, Gillies 1982 p 54; cf. Frege 1895). In the letter to Weber Dedekind addressed Weber's question, probably in a previous letter, whether a set that contains itself is infinite. Dedekind limited his answer to the case that the set contains only itself which brought him to the problem of the notation of a singleton, and he suggested to discuss it with Weber in a meeting of which we have no further information. Obviously, Weber's question could not be reduced to the notational problem that preoccupied Dedekind and it seems to lead to the question of foundation (regularity), well before Mirimanoff (see Hallett 1984 p 185, Ferreirós 1999 p 370).

The problem of foundation is connected to Cantor's inconsistent sets theory. Ω was found to be an inconsistent set for otherwise its ordinal must belong to it, which leads to a contradiction. In Cantor's new example, T' is inconsistent for otherwise it

[7] Scharlau 1981 p 10 says that the meeting took place in Harzburg but Dedekind's letter of August 29 is from Braunschweig and Dugac 1976 p 130 cites a postcard from Cantor dated September 3 announcing his coming to Braunschweig the next day.

would be equivalent to one of its subsets which contains a larger set. With Dedekind's infinite set the situation is somewhat similar: since it is thinkable it must be its own member. Though in this case no contradiction results, against the background of the other examples this was probably enough to raise doubts in Dedekind's mind with regard to S. All the more since the situation of a set that contains itself seems to violate Cantor's requirement of members of a set, that they be distinct and definite (1895 *Beiträge* §1, Fraenkel 1966 p 9, Fraenkel et al. 1973 p 86), a view which Dedekind accepted in his sack metaphor.[8]

The way Dedekind perceived the situation is clearly described in his preface to the third edition of *Zahlen* from 1911 (Dedekind 1930–32 p 343, English translation: Ewald 1996 vol 2 p 796)[9]:

> When I was asked roughly eight years ago [1903] to replace the second edition [from 1893] of this work (which was already out of print) by a third, I had misgivings about doing so, because in the meantime doubts had arisen about the reliability of important foundations of my conception. Even today I do not underestimate the importance, and to some extent the correctness, of these doubts. But my trust in the inner harmony of our logic is not shattered; I believe that a rigorous investigation of the power of the mind to create from determinate elements a new determinate, their system, that is necessarily different from each of these elements, will certainly lead to an unobjectionable formulation of the foundation of my work. But I am prevented by other tasks from completing such an investigation; so I beg for leniency if the paper now appears for the third time without changes – which can be justified by the fact that interest in it, as the persistent inquiries about it show, has not yet disappeared.

Dedekind expresses here two concerns: that when elements are given, a set can be created that contains them, and, such which is different from each of them. Dedekind's first concern is with the household comprehension introduced in *Zahlen* (p 45 #2) with no justification and applied throughout, e.g., to establish the algebraic operations between sets, to which the theory of natural numbers is reduced, and to establish the infinite set. Dedekind attributed comprehension to the creative power of the mind, (the first preface to *Zahlen*, the unpublished paper (p 252), the above quoted third preface).

Dedekind's second concern addresses the issue of foundation (simplified, because the demand is not explicitly made that the set be different from its hereditary members). The problem with foundation crept into *Zahlen* without Dedekind's notice. Dedekind probably became aware of it only through his discussions with Bernstein and Cantor that magnified Weber's earlier remark.

Cantor had solutions to both Dedekind worries, inconsistent sets for the comprehension worry and the characterization of members of sets for the foundation worry. The fact that Cantor had solutions for these worries certainly served to strengthen Dedekind's feeling that something was wrong. But he cannot be blamed for

---

[8] Ebbinghaus (2007 p 85) calls Dedekind's infinite set 'inconsistent' not in Cantor's sense but to convey its contradictory character. In this he is not justified (see Fraenkel et al. 1973 p 86 footnote 4).

[9] http://www.ru.nl/w-en-s/gmfw/bronnen/dedekind1.html.

rejecting Cantor's remedies that were vaguely formulated. Dedekind even preferred a stance of belief, as expressed in the quoted passage, to an acceptance of Cantor's approach.

Dedekind points to 1903 as the time when his doubts were already in place. Perhaps as a result, Dugac suggested (1976 p 89) that "the echo arriving at Dedekind from Göttingen about research on the foundation of mathematics and set theory" was one of the reasons to postpone the third edition of *Zahlen*. But this view is not justified. In 1903, talk of the paradoxes was only a mumble, and even Russell (1903 p 357) still accepted Dedekind's proof of the existence of an infinite set as undeniable. What Dedekind in fact says is that in 1903 he was asked to replace the second edition and that in the meantime – between the second edition and 1903 – doubts had arisen. Likewise we disagree with Dugac that it was Zermelo's axiomatic set theory that gave Dedekind his confidence for a new edition of *Zahlen*. What happened in 1911 that renewed in Dedekind his conviction in the mentalist foundation of logic was probably the appearance in 1910 of Whitehead-Russell Principia Mathematica with its theory of types, which ensured that a set is different from its members, and addressed the issue of comprehension as well.

Chapter 9
Dedekind's Proof of CBT

In 1887, when he prepared his *Zahlen* for publication, Dedekind wrote down a proof of CBT. The proof, dated July 11, 1887, was found in Dedekind's *Nachlass* and was first published in 1932, in Dedekind's collected works (Dedekind 1930–32 p 447). The proof is for both the single-set and the two-set formulations. In his letter to Dedekind of November 5, 1882, Cantor only mentioned the theorem in its single-set formulation, from which the derivation of the two-set formulation is not natural. Thus the question arises how did Dedekind learn of the two-set formulation. It is possible that Cantor mentioned both formulations to Dedekind in their discussion of the theorem in September 1882, recollected in the above referenced letter to Dedekind. Or it is possible that Dedekind knew of the two formulations from his own experience developed when he was preparing the early drafts of *Zahlen* in the years 1872–78 (Dugac 1976 pp 79, 293–309). We prefer the second possibility because Dedekind's proof is indeed based on results that he obtained in that draft. Thus it seems very likely that Dedekind recognized immediately upon first hearing of CBT, that a proof for it can be obtained from his chain theory. This is not surprising because the conditions of the single-set CBT are those of a reflexive set, namely, the gestalt of a set and a subset is given and the metaphor that the two are equivalent. Associating the two contexts is thus natural.

Dedekind did not mention his proof to Cantor, perhaps because of their tainted relationship (Ferreirós 1993 p 349, see Chap. 7). Since Cantor announced in the November letter that he had obtained a proof of the theorem, there was nothing pressing Dedekind to present his own. In addition, at the time, Dedekind may have wanted to distance his work from that of Cantor for two good reasons: First, Dedekind's effort to produce manicured monographs was not compatible with the far from settled work of Cantor in the 1880s. Second, because Cantor mentioned in his 1878 *Beitrag* paper the defining property of infinite sets and the equivalence class of equivalent sets, both notions used in *Zahlen*. That Dedekind felt he had to defend the originality of his work against Cantor is evident from his omission of Cantor's name from his first preface to *Zahlen* (it is mentioned in the draft – see Cavailles 1962 pp 120–1, Ferreirós 1996 p 46 footnote 72, Ferreirós 1999 p 270). He also stressed in the second preface to *Zahlen*, the independence of his

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_9, © Springer Basel 2013

achievements from those of Cantor. Nevertheless, Dedekind did allow Cantor to read his first drafts of *Zahlen* (Ferreirós 1993 p 349, 1999 p 269). Surprisingly, Cantor did not recognize the relevance of Dedekind's chain theory to CBT. The reason this is surprising is that Cantor used similar ideas already in his 1878 *Beitrag* paper (Cantor 1932 p 119), perhaps even to prove CBT for the continuum (see Sect. 3.4).

From the first proof Dedekind extracted a lemma that he inserted into the final version of *Zahlen* as Theorem 63 (Dedekind 1963 p 63), leaving only a stretch of unpaved reasoning leading to CBT (see Noether's remark to the third proof in Dedekind 1930–32 p 448). To that theorem Dedekind added the following (atypical) restrictive qualification: "the proof of this theorem of which . . . we shall make no use[,] may be left to the reader". Early readers of *Zahlen*, which was widely read, must have taken Dedekind's advice literally and jumped over the theorem, what with its position at the end of a chapter giving it an air of dead-end. Even Zermelo, who in his 1908b paper on axiomatic set theory reconstructed the proof from the theory of *Zahlen*, appears to have not realized at the time that Dedekind cached the proof in Theorem 63. It seems that only with the discovery of the first proof in the *Nachlass*, was the role of Theorem 63 appreciated.

Ferreirós (1993 p 355, 1999 pp 239–240) suggested that Dedekind included Theorem 63 in *Zahlen*, while not explicitly stating its connection to CBT, because he wanted to test the alertness of Cantor, in the manner of seventeenth century mathematicians, and that apparently Cantor failed the test. Ferreirós concluded that the incident shows the lack of collaboration between the two mathematicians and that Cantor paid little attention to Dedekind's theory. We rather believe that Cantor's failure was a result of his ingrained gestalt that linked CBT to the scale of number-classes.

Following the visit from Bernstein on Pentecost of 1897 (see Sect. 8.2), during which Bernstein told Dedekind of his (very recent) proof of CBT, Dedekind rewrote a proof of the theorem (so this is the third proof). Two years later, attached to his letter dated August 29, 1899 (see Sect. 7.4), he sent it to Cantor. The letter containing the third proof was also found only in Dedekind's *Nachlass* and published in 1932.

In a remark to the letter Noether noted (Dedekind 1930–32 p 448) that: "Apparently Dedekind had forgotten that he was dealing with a reconstruction of an older proof". We do not think that this impression, repeated by Ewald and accepted by Medvedev (1966 p 234), is correct: first, because as it is described in the letter, Dedekind responded on the spot to Bernstein's mention of his proof of CBT that the theorem can be proved by the means of *Zahlen*. Second, in the first proof Dedekind made a note to himself to change the order of presentation and he implemented this remark in the third proof.

The third proof is for the single-set formulation, which Dedekind denoted by the letter C after its naming in the 1895 *Beiträge* §2. Dedekind remarked that it is the same as B of the same paper. However, there the equivalence of the two theorems is not mentioned and so we must conclude again that either Cantor mentioned the equivalence of the two formulations of CBT in 1882 or that Dedekind was aware of the complete meaning of CBT from his own study.

Dedekind did not send the third proof to Cantor right away when he wrote it down, perhaps because he did not want to embarrass Bernstein who, as Dedekind observed in the letter, seemed startled at Dedekind's declaration that the theorem can be proved by the means provided in *Zahlen*. Additionally, Dedekind perhaps assumed that as he explicitly referenced Bernstein to *Zahlen*, Cantor would be able to link Theorem 63 to a proof of CBT by himself. As no feedback arrived from Cantor, Dedekind took the opportunity to send the proof in 1899, when he wanted to show Cantor that he sometimes took interest in his theory (see Sect. 7.4).

In a comment to the third proof (Cantor 1932 p 451 first paragraph and [2]), Zermelo, described it as "classic". He noted the proof's reliance on Dedekind's "chain theory" upon which his own proof of 1908a was also based, without knowing about the proof of Dedekind. He then expressed his wonder "why neither Dedekind nor Cantor have decided at the time to publish this, after all, not unimportant proof" and added that this "is today [1932] not rightly understood". But then, in 1899 it was superfluous to publish another proof of a theorem that had two published proofs, by Schröder and Bernstein, which proofs Cantor regarded in his letter to Dedekind of August 30, 1899, as similar to the proof of Dedekind. Besides Cantor declared CBT to be proved in his *Grundlagen* and then to be a corollary from the Comparability Theorem for cardinal numbers in his 1895 *Beiträge*.

What remains a mystery is why the information on Dedekind's proof did not become public property. Apparently Cantor never showed Dedekind's letter with the proof to any of his colleagues: to Bernstein, with whom he collaborated in those years, to Schoenflies, on account of his 1900 exposition of set theory, or to Hilbert, with whom he exchanged letters and met a couple of times in the late 1890s.

Dedekind's proofs of CBT were published only decades after they were conceived and so they did not have any impact on the development of set theory. Had Dedekind published his proof of CBT in 1887 or thereabout, or if the readers of Dedekind's *Zahlen* had noticed his cached proof, the Cantor-Bernstein Theorem would have been rightly named the Cantor-Dedekind Theorem. For it was 8 years before Bernstein and Schröder produced their proofs. And the name would have been very becoming to the theorem because it involves the two founders of set theory (on Dedekind's general contribution to set theory see, for example, Dugac 1981 pp 137, 138, 142) and it concerns a theorem that really stands at the border line between these two great mathematicians and their theories: Cantor, on the one side, with his theory of infinite numbers, in which the theorem finds its main application, and Dedekind, on the other side, with his primal theory of sets and mappings and his chain theory, in which the theorem finds its most elegant expression and proof. Even today it is perhaps not too late to make historical justice and rename the theorem as the Cantor-Dedekind-Bernstein Theorem.[1] The name of Bernstein ought to remain attached to the theorem

[1] A similar suggestion is made in Mańka-Wojciechowska 1984 p 195.

for it was his proof (in fact it was Borel's rendering of Bernstein's proof which in essence was probably different, see Sect. 11.2) that was there when set theory matured. Despite this suggestion, in order to keep within tradition until it changes by general consent, we stick with the 'Cantor-Bernstein Theorem' name.

9.1 Summary of the Theory of Chains

Dedekind's purpose in writing *Zahlen* was to reduce the truth of the laws of natural numbers "to others more simple", (*Zahlen* p 33) which seem to be the concepts of "infinite set", "mapping" and "one". The first two concepts are linked by the definition of an infinite set: it is a set equivalent to its own proper subset (*Zahlen* p 63 #64). Dedekind's "reduction" is a translation of arithmetic into a new dominant theory – the theory of sets (see Lakatos 1976 p 125).

If S is an infinite set and φ the equivalence that maps S into itself (we call it 'reflection' and S 'reflexive'), Dedekind denotes the image of any $K \subseteq S$ under φ by K'. A subset K of S for which $K' \subset K$ Dedekind calls a "chain" (*Kette*; *Zahlen* p 56). The origin for this name is unknown to us.[2] Dedekind defined chains even if φ is many-one but the main results we are interested in are for chains obtained by 1–1 mapping. It is to be noted that the property of being a chain is established by φ and is not a property of K alone. Obviously a chain is an infinite set and the chains are all the infinite subsets of S that are infinite by way of φ. S is a chain and so are the union and intersection of chains; thus the domain of chains (of a given set and reflection) is closed under union and intersection. The algebraic nature of the collection of all chains fits Dedekind's taste for abstract algebraic structures. By φ, S is endowed with a structure: the collection of its chains and their images (see below). A similar point of view, that when a certain set and an operation on it is given, a whole structure unfolds, is pronounced by Cantor (1872 p 97; cf. Hallett 1984 p 4) with regard to point-sets and the derivation operation: all derived sets from the given set are given with it.

If A is a subset of S then the intersection of all the chains that contain A is the minimal chain that contains A; it is denoted by A_0 and called the "chain of A" (*Zahlen* pp 57, 58). The minimality of A_0 is by the subsets relation which is a partial order; thus A_0 is a subset of any chain that contains A. Dedekind notes that $A_0 = A + {A_0}'$ and ${A_0}' = A'_0$; the latter set he calls the "image-chain" (*Bildkette*). A theorem that corresponds to complete induction Dedekind established by showing that to prove that a chain A_0 is a subset of a set P one has only to prove that $A \subseteq P$ and that for every member of A_0 if it is in P its image is also in P (*Zahlen* p 60 #59).

Dedekind noted that every infinite set has an infinite subset, say N, that is mapped into itself (so that it is a chain) by the reflection mapping φ that establishes the infinity of the set, and is such that there is only one element, denoted by 1, in N,

[2] Kekulé used '*kette*' in his structure of benzene (1865).

which is not in φ(N) (*Zahlen* pp 67, 68); in fact N is $\{1\}_0$ (this is where the concept "one" enters).[3] Such a chain as N Dedekind called a simple chain. Dedekind then shows, and this is the main content of *Zahlen*, that the elements of N satisfy all the properties of the natural numbers, namely, N is a model of the natural numbers.

If n is a natural number from one such model of the natural numbers, Dedekind uses φ^n in its standard meaning and he shows that $A_0 = \Sigma_n \varphi^n(A)$ (*Zahlen* p 92). We call $\varphi^n(K)$-$\varphi^{n+1}(K)$, when K is a chain, the frames of K. Frames are pairwise disjoint. If $U \subseteq S\text{-}S'$ then the chain U_0 generated by U is composed only of its frames; otherwise there may be a residue to the chain after all its frames are removed. Obviously, a simple chain can be generated from any element of any frame of a chain.

Dedekind's process of generating the natural numbers differs in more than one aspect from the one we are more familiar with, whereby the natural numbers are generated by the process of adding a unity. Under the latter point of view numbers are created as necessary and there are never really infinitely many numbers. Even in Cantor's construction by way of the generation principles, the finite and infinite numbers are created as necessary, or by an ad-hoc assumption, and not all at once, as Cantor stressed in *Grundlagen* §11, 12. In Dedekind's construction, however, all the numbers are created at once.

The fundamental observation of Dedekind is that when you say that S is an infinite set you are implying the existence of a reflection φ of S which maps it on a proper subset S′. Then, not only is S mapped onto S′; also, and by the same reflection, S′ is mapped onto S″, and S″ onto S‴, and so on. In particular, all the $S^n = \varphi^n(S)$ are infinite and are chains and all the φ^n are reflections of S and of any of its images. We can take as the image of S to be S′ or S″ or $S^{(n)}$ or, by CBT and using a combination of φ and the identity as will be described below, any T between S and any of these images. All these possibilities come with the notion of a chain, which comes with the notion of infinite set. The gestalt that φ imposes on S is more elaborate than what seems at first glance.

Moreover, the reflection process, in spite of its step-by-step description (on this favorite Dedekindian term see Sect. 7.4), happens with one chute: when we set off φ to run, a cascade of images is created and we do not need to apply φ again to get $\varphi(\varphi(S))$, $\varphi(\varphi(\varphi(S)))$ and all $\varphi(\ldots(\varphi(S))\ldots)$. The whole cascade chutes down its infinite (denumerable, to be sure) run at once. If only we could see and hear the chute in slow motion, accelerating to converge in finite time, what a formidable show of colorful lights and sounds it could be! However, in reality (reality?!), the chute does not happen: we simply become conscious, when we realize what it means that a set is infinite, of a structure that is immanently there![4]

[3] Dedekind writes 1_0 though he is aware of the distinction between a thing and the set that contains only it (see Sect. 8.4).

[4] Cavailles (1962 p 123) and Pla i Carrera (1993 p 278) are wrong when they see in Dedekind's chain an iterative concept. But one can certainly operate with this imaginary gestalt.

However, under the chain gestalt, the individuality of the members of S is blurred. Ideologies that strive for uniform social structures operate under the chain gestalt. J. Kőnig's proof (see Sect. 21.2) will restore the centrality of the individual in the gestalt of proofs of CBT.

That a definition by reflection can replace the traditional way of generating the natural numbers through the operation of sequent, is the core of Dedekind's theory of *Zahlen*. The impredicative nature of this definition became the subject of Poincaré's criticism of logicism in 1906 (see Sect. 19.3). Dedekind did not mention this criticism even in his third preface (1911) to *Zahlen*.

9.2 Dedekind's Proofs

We reproduce here all three proofs Dedekind gave to CBT.

9.2.1 The First Proof

Theorem[5]. If S is equivalent into itself, so that the image $\varphi(\mathrm{S}) = \mathrm{S}' \subseteq \mathrm{S}$,[6] and further $\mathrm{S}' \subseteq \mathrm{T} \subseteq \mathrm{S}$, then T is also equivalent to S.[7]
Proof: Clear in case $\mathrm{T} = \mathrm{S}$.[8] Otherwise let U be the system of all elements of S that are not contained in T[9] and let U_0 be the image-chain [*Bildkette*][10] (§4)[11] corresponding to it by φ. Let now s be any element of S so we set: $\psi(\mathrm{s}) = \varphi(\mathrm{s})$ or $= \mathrm{s}$,[12] if s is contained in U_0 or not.[13] Then ψ is a 1–1 mapping, and indeed $\psi(\mathrm{S}) = \mathrm{T}$. To this end [*hierzu*] we have to show that:

1. ψ is a 1–1 mapping. (a) When a and b are different and contained in U_0 then $\psi(a) = \varphi(a)$ and $\psi(b) = \varphi(b)$ are also different, since φ is a 1–1 mapping. (b) When a is contained in U_0 but b is not, then $\psi(a) = \varphi(a)$ and $\psi(b) = b$ are also different, since U_0 is a chain and

[5] In our translation, from Dedekind 1930–2 vol 3 p 447–8. We use "set" for Dedekind's "system", "mapping" for "*Abbildung*", "equivalence" or "1–1 mapping" for "*ähnlich abbildung*"; the translation in *Zahlen* is "transformation" and "similar transformation" or "similarity", respectively. We use + for union instead of Dedekind's $\mathfrak{M}$.

[6] φ is not explicitly introduced.

[7] This is the single-set formulation.

[8] Actually differentiating this case is not necessary because the proof that follows holds even if U is empty, in which case U_0 is also empty, but in *Zahlen* (p 45) Dedekind avoided the empty set.

[9] Namely, $\mathrm{U} = \mathrm{S}\text{-}\mathrm{T}$ only Dedekind does not use the difference between sets operation.

[10] It seems that Dedekind made a mistake here in his use of his own terminology: U_0 is the "chain of U" (*Zahlen* p 58 #4) while the image-chain refers to the chain generated by the image of U, i.e., U'_0 (*Zahlen* p 59 #57 where "transform" is used for our "image", Ferreirós 1993 p 355, Hessenberg 1906 p 689).

[11] Reference is to *Zahlen*.

[12] Namely, $\psi(\mathrm{s}) = \mathrm{s}$.

[13] This is a definition of a new mapping ψ.

$\varphi(U_0) \subseteq U_0$, so that $\psi(a)$ is contained in U_0, but b is not. (c) When a and b are not contained in U_0 and are different then $\psi(a) = a$, $\psi(b) = b$ are also different.

2. $\psi(S) \subseteq T$. Denote by V the set of all elements of S which are not contained in U_0, so that $S = U_0 + V$ and $\psi(S) = \varphi(U_0) + V$. Now since $U_0 \subseteq S$ it is that also $\varphi(U_0) \subseteq \varphi(S)$ and as $\varphi(S) \subseteq T$ also $\varphi(U_0) \subseteq T$; as in addition $V \subseteq T$ {for if an element v of V is not contained in T it is in U thus (§4) also in U_0, against the definition of V; better to advance[14]}, so it follows (§1) $\psi(S) \subseteq T$.
3. $T \subseteq \psi(S)$. Let t be any element from T. If t is contained in V t is contained (as a result of $\psi(S) = \varphi(U_0) + V$)[15] also in $\psi(S)$. But if t is not in V, then it is contained in U_0, then it must be, since (§4) $U_0 = U + \varphi(U_0)$ and by the definition of U, that t is contained in $\varphi(U_0)$, and therefore also in $\psi(S)$. Q.E.D.

Or right clear $T = \varphi(U_0) + V = \psi(S)$.
Theorem.[16] If A is equivalent to a subset of B and B is equivalent to a subset of A, then A and B are also equivalent.
Proof: [Given] are 1–1 mappings φ, ψ; and $\varphi(A) \subseteq B$, $\psi(B) \subseteq A$ so that
$\psi\varphi(A) \subseteq \psi(B) \subseteq A$; since the mapping $\psi\varphi$ is likewise an equivalence, $\psi\varphi(A)$ is equivalent to A, then (by the previous theorem) $\psi(B)$ is equivalent to A and as $\psi(B)$ is equivalent to B also A (by a simple theorem about equivalence) is equivalent to B. Q.E.D.

9.2.2 The Second Proof

Theorem 63 of *Zahlen* (p 63) runs as follows (we changed the signs to match the first proof and added comments in square brackets):

> If $S' \subseteq T \subseteq S$, and therefore S is a chain, T is also a chain [$T \subseteq S$ entails $T' \subseteq S' \subseteq T$]. If the same [T] is a proper subset of S, and U the set of all those elements of S which are not contained in T [U = S-T], and if further the chain U_0 is a proper subset of S [this condition is fulfilled when S' is a proper subset of T because then T-S' is not in U_0], and V the set of all those elements of S which are not contained in U_0 [V = S-U_0], then $S = U_0 + V$ and $T = U_0' + V$ [also $S' = U_0' + V'$]. If finally $T = S'$ then $V \subseteq V'$ [if φ is 1–1 then $V' \subseteq V$ so in this case and when $T = S'$ we have $V = V'$].

Ferreirós (1993 p 355, 1999 pp 239–40) noted that Theorem 63 does not assume that the mapping from S onto S' is 1–1 and that if this assumption is added then the conditions of the theorem are those of CBT.[17] Then, because $U_0 \sim U_0'$, the partitioning outlined in Theorem 63 give S ~ T which proves the theorem. Essentially this is Noether's remark mentioned earlier.

---

[14] This is an editing remark Dedekind made for himself. He meant that the argumentation regarding the sets and their images should be made prior to the definition of ψ. Indeed Dedekind implemented this remark as we will note below.

[15] There is a typo here in the original and V_0 is written instead of U_0.

[16] This is the two-set formulation.

[17] In its single-set formulation which Ferreirós describes as a "crucial lemma in the proof of the Cantor-Bernstein theorem" rather than an alternative formulation of it.

9.2.3 *The Third Proof*

We reproduce Dedekind's third proof of CBT, included in his letter to Cantor of August 29, 1899, (except the square brackets, the technical terms that we use after our convention and the signs for the sets that we changed to match those of the first proof):

> If the set S′ is a subset of the set T, and T a subset of the set S, and S is equivalent to S′, then S is also equivalent to T.
> Proof: The theorem is evidently trivial if T is identical with S or with S′. In the contrary case, if T is a proper subset of S, let U be the set of all those elements of S which are not contained in T – that is (in the notation of Dedekind[sic], Cantor, Schröder)
>
> S = $\mathfrak{W}$(U, T) = (U, T) = U + T.
>
> By assumption, S is equivalent to the (proper) subset S′ of T, so there is a 1–1 mapping φ of S into itself, by which S is mapped into [onto] $S' = \varphi(S)$; let U_0 be the 'chain of U' (§4 of my paper, *Was sind und was sollen die Zahlen?*), so[18]: $U_0 = U + U_0'$; since U_0 is a subset of S, $U_0' = \varphi(U_0)$ is a subset of $S' = \varphi(S)$, and therefore U_0' is also a proper subset of T; so U and U_0' have no common element, and U_0 is also a proper subset of S. Let V be the set of all those elements of S which are not contained in U_0, i.e., $S = U + T = U_0 + V$ [$= U + U_0' + V$ hence, since all the sets concerned are disjoint], $T = U_0' + V$, where U_0' as a subset of U_0 has no element in common with V. Now we define a mapping ψ of S by setting $\psi(s) = \varphi(s)$ or $\psi(s) = s$ according as the element s of S is contained in U_0 or in V. This mapping ψ of S is 1–1, for if s_1 and s_2 designate different elements of S, then they are[:]
>
> either contained in U_0; for $\psi(s_1) = \varphi(s_1)$ is different from $\psi(s_2) = \varphi(s_2)$, because φ is a 1–1 mapping of S (which is used here for the first time, and only here);
>
> or they are contained in V; for $\psi(s_1) = s_1$ is different from $\psi(s_2) = s_2$;
>
> or one element s_1 is contained in U_0 and the other element s_2 is contained in V; then $\psi(s_1) = \varphi(s_1)$ is different from $\psi(s_2) = s_2$, for $\psi(s_1)$ is contained in U_0' while s_2 is contained in V.
>
> This 1–1 mapping ψ maps $S = U_0 + V$ into [onto]
>
> $\psi(S) = \psi(U_0) + \psi(V) = T$ because $\psi(U_0) = \varphi(U_0) = U_0'$ and $\psi(V) = V$.
>
> Q.E.D.

9.2.4 *Comparing the Proofs*

Noether (Dedekind 1930–32 p 448) had already made a note comparing the proofs; she found that the third proof coincides with the first, except for notation changes (which we have ironed out), and that it is less tightly locked on *Zahlen*. She also remarked, as noted, that Theorem 63 contains the essential lemma for the proof. Let us make these comparisons in more detail.

The first and third proofs begin with settling the trivial cases. In the first proof only one trivial case is considered: that T = S; in the third proof the case T = S′ is

[18] Here in Ewald there is a reference to §8 which seems to be a typo; it is not in Cantor 1932 p 449.

also set aside. Dedekind is not explicit how the trivial cases are settled; in fact, they are not settled by the same argument: in the first case the similarity of S and T is provided by the identity mapping; in the second case it is given by φ. The two trivial cases thus suggest that in the solution of the general case the searched mapping should be a combination of the identity mapping and φ. This "mean value" heuristic is, unfortunately, never explicated when the proof of CBT is discussed in textbooks.

In the second proof the negations of the trivial cases are introduced as conditions for the result. But, as with the two other proofs, the results are not affected if these conditions are dropped or the trivial cases not differentiated; only then, use of the empty set is required, a use which Dedekind wanted to avoid as he explicitly stated in *Zahlen* p 45.

In all the proofs the chain generated by S-T is introduced first but while in the first proof the function ψ is defined next, in the other two proofs the second step is comprised of the partitioning of S and T by this chain, its image and its complement. These partitionings are presented in the first proof in (ii), (iii) and the concluding sentence. It thus appears that in the second and third proofs Dedekind implemented his editing remark from the first proof to change the order of the presentation. So we see that when Dedekind wrote down his third proof he was well aware of his first proof, contrary to Noether's remark that apparently he had forgotten about it.

The argument that $\mathrm{T} = \psi(\mathrm{S})$, in (ii) and (iii) of the first proof, by way of extensionality, is avoided in the third proof since the partitioning gives this result immediately. Dedekind introduced extensionality in #5 (p 46) of *Zahlen* and it is perhaps because the third proof made no reference to this theorem that Noether noted that it is less locked on *Zahlen*. In the second proof the stage is set for the definition of ψ as in the third proof but the definition itself is missing. To a modern reader the explicit definition of ψ may appear redundant because we are more used to operating with partitions and once it is clear that the partitions are pairwise equivalent and pairwise disjoint (note that the latter 'pairwise' is different from the previous one), the equivalence of the compounded sets is immediately perceived. Perhaps this is why Dedekind made the remark on the first and only use of φ in the third proof, to hint that the specific mention of φ is redundant. Cantor similarly in his 1878 *Beitrag* was satisfied in showing that if the partitions are equivalent so are the sets. Note that Dedekind in the third proof avoided referencing his #63 of *Zahlen* where the partitioning of S and T is provided, thus he avoided linking the proof to his earlier monograph.

Despite the differences noted, no differences in proof descriptors can be identified among the three proofs (see below) and all three can be considered as merely three variants of one proof, Dedekind's CBT proof. For comparison purposes we will consider the third proof as the basic one.

9.3 The Origin of Dedekind's Proof

It seems to us that Dedekind derived his proof of CBT in the single-set formulation from his proof of the following theorem (*Zahlen* p 64 #68, Hessenberg 1906 p 507):

> Every set S, which possesses an infinite part [subset] is likewise infinite; or, in other words, every part of a finite system is finite.

Dedekind needed Theorem 68 in order to assert that the scale of subsets does not mingle finite and infinite sets. Then, in the realm of finite sets, the relation of subsets provides for the transitivity of the 'greater than' relation between finite numbers. In his 1895 *Beiträge*, Cantor had his version of Theorem 68; it is Theorem E of §5 of that paper (Cantor 1915 p 98ff). The reason that Cantor needed Theorem E is that in his 1895 *Beiträge* Cantor dropped the ordinal definition of the finite numbers by way of the first principle which he used in *Grundlagen*. Instead he used a cardinal approach that is comprised of the following steps: First the finite cardinal numbers are introduced by abstraction from a sequence of specific sets that are constructed by adding a new thing at each step. No doubt, an inductive process, which is used, however, to construct finite sets not the finite numbers. Then (§6) a finite set is defined as a set with cardinal number a finite number; all other sets are called "'transfinite sets' and their cardinal numbers 'transfinite cardinal numbers'". Because of this approach Cantor needed a theorem that warrants that the subsets of finite sets are finite.

Cantor proved his Theorem E by complete induction. Dedekind proved Theorem 68 as follows:[19]

> If T is infinite and there is hence such a similar transformation ψ of T, that ψ(T) is a proper part of T, then, if T is part of S, we can extend this transformation ψ to a transformation φ of S in which, if s denotes any element of S, we put $\varphi(s) = \psi(s)$ or $\varphi(s) = s$ according as s is element of T or not. This transformation is a 1–1 mapping; for, if a, b denote different elements of S, then if both are contained in T, the transform $\varphi(a) = \psi(a)$ is different from the transform $\varphi(b) = \psi(b)$, because ψ is a 1–1 mapping; if further a is contained in T, but b not, then is $\varphi(a) = \psi(a)$ different from $\varphi(b) = b$, because ψ(a) is contained in T; if finally neither a nor b is contained in T then also is φ(a) = a different from φ(b) = b which was to be shown. Since further ψ(T) is part of T, [and thus] also part of S, it is clear that also $\varphi(S) \subseteq S$. Since finally ψ(T) is a proper part of T there exists in T and therefore also in S, an element t, not contained in $\psi(T) = \varphi(T)$; since then the transform φ(s) of every element s not contained in T is equal to s, and hence is different from t, t cannot be contained in φ(S); hence φ(S) is a proper part of S and consequently S is infinite, which was to be proved.

This proof is similar to Dedekind's first proof of CBT in the metaphor that the constructed mapping is a combination of the identity and the given mapping. Another similarity is in the way that the constructed mapping is shown to be 1–1 ((i) of the first proof) by differentiating the cases according to the partitions upon which the mapping has different definitions. The difference between the two proofs

[19] We quote from Dedekind 1963. The translation does have some Germanized English.

is that in the first proof of CBT the given mapping is reduced to U_0 and then extended with the identity while in Theorem 68 the given mapping is just extended with the identity. Still it seems that Dedekind proof-processed his CBT proof from his proof of Theorem 68.

Another outlook on the similarity between Theorem 68 and CBT can be gained by analyzing them in terms of sets. In both theorems we have three sets $R \subset T \subset S$; in CBT we have an additional condition $R \sim S$ and the conclusion is that $R \sim T$; in Theorem 68 we have the additional condition $R \sim T$ which is the conclusion of CBT, and the conclusion is $R + (S\text{-}T) \sim S$ – similar to but not quite, the condition of CBT. In other words, Theorem 68 is an almost reciprocal theorem to CBT.

Further linkage of CBT with *Zahlen* appears when we compare it with Theorem 41 of *Zahlen*, which is:

> If the transform A' is part of a chain L then there is a chain K which satisfies the condition that $A \subseteq K$, $K' \subseteq L$; $A + L$ is just such a K.

Theorem 41 describes (when the underlying mapping is 1–1) the conditions of the single-set formulation of CBT (because $K' \subseteq L \subseteq K$) though its conclusion, that $L \sim K$, is not mentioned. Theorem 41 describes one of the cases when extending a chain generates a chain. Another case is described in Theorem 42 which says that the union of chains is a chain. These two theorems are part of the closure theorems of the domain of chains. Theorem 68, like Theorem 41, also speaks of extending a chain (an infinite set is a chain) by the addition of a set, only the setting is not within an infinite set and its reflection but "at large", in the world of sets, and the extended set is a chain by a new reflection which combines the reflection of T and the identity.

Both Theorems 41 and 68 appear in the draft of *Zahlen*, though there Theorem 41 is introduced after Theorem 68 (Dugac 1976 pp 294, 299, 306) which comes to show that the methodic presentation of *Zahlen* does not reflect the order of discovery. Thus we believe that in 1882, when Dedekind first heard about CBT, he immediately conceived its proof; his use of similar language in the proofs of Theorem 68 and CBT reveals his awareness to their similarity. However, we do not think that Dedekind was aware of the general importance of CBT before he was introduced to it by Cantor. For in the first drafts of *Zahlen* a definition of the equivalence classes of (infinite) equivalent sets (*Zahlen* p 55 #34) is missing and it is for the introduction of those classes that CBT is necessary (otherwise the equivalence classes may "mingle" by the subset relation – see Sect. 3.1). #34 was probably added after Cantor's 1878 *Beitrag* paper (where a similar notion was first introduced by Cantor) and Dedekind's 1882 communication with Cantor. Thus Dedekind's 1887 CBT proof was necessary for the introduction of equivalence classes to *Zahlen* but Dedekind decided not to include it in the monograph perhaps again to distance his work from that of Cantor.

Here we can come back to a point raised earlier, concerning the source of Dedekind's awareness to the two-set formulation which was not mentioned, for all that we know, by Cantor in his communication with Dedekind in 1882. In the first drafts of *Zahlen* the two-set formulation for finite sets is mentioned without proof (Dugac 1976 p 303, after the introduction of the versions of Theorem 68 and Theorem 41):

> If each of the sets A, B can be 1–1 mapped in the other [obviously it is not known at the outset if the subset is proper or not and it turns out that it cannot be a proper subset] and if one of them is finite, so is the other always finite, and the number [*Anzahl*] of the things contained in A is equal to the number of things contained in B.

In *Zahlen* there is no similar theorem and its content is dispersed in other theorems: First there is Theorem 69 (a simple corollary of Theorem 68) which states that every set that is equivalent to a part of a finite set is finite. Then there is Theorem 165, which states that a proper part of a finite set has a smaller number. Theorem 165 rests on Theorem 68 and a bunch of theorems stating the order properties of numbers (*Zahl*). Now it is, we believe, inevitable that when Dedekind was reflecting over his two-set formulation, he glided into the single-set formulation, partly present in Theorem 41. Thus Dedekind was able to relate his first proof to both formulations of CBT.

With regard to Dedekind's use of the term *Zahl* we note that prior to his two-set formulation Dedekind proved (Dugac 1976 p 303) that for every finite set B there exists a number n (and only one) so that $[n]$ (Dedekind used this notation in the first drafts for the set of all numbers which are smaller or equal n; in *Zahlen* this set is denoted by Z_n) can be 1–1 mapped in B but $[n']$ (n' is $n + 1$) cannot be 1–1 mapped in B. This is Theorem 160 of *Zahlen*. Following this theorem it is defined that the *Anzahl* of a set B is the number n of that theorem. This is definition 161 of *Zahlen*. Thus we see that, for Dedekind, *Zahl* (defined in the first drafts in Dugac 1976 p 300 and in *Zahlen* p 68 #73) was number in its ordinal capacity and *Anzahl* is number in its cardinal capacity applied to a certain finite set. Cantor used the reverse meaning for these two terms since his 1878 *Beitrag* and he regarded ordinal and cardinal numbers as different entities that correspond for finite numbers. If this was no coincidence then the two mathematicians must have conversed on the ordinal and cardinal nature of numbers already in their 1872 meeting (compare Ferreirós 1993).

9.4 Descriptors for Dedekind's Proof

The chain bestowed on an infinite set by a reflection is the fundamental gestalt of Dedekind's proof. Yet Dedekind's proof uses a derived gestalt. The set S is partitioned in two: a chain U_0 and a complement V. It is perceived then that therewith the set T is partitioned in two as well: U_0' and V. The proof is obtained when it is realized that the given mapping maps U_0 into its image-chain. The metaphor that we thus associate with the proof is 'pushdown the chain', and for the gestalt and metaphor together we recoin the old Roman dictum: "partition and pushdown". The partitioning gestalt must have been basic with Dedekind for it

Fig. 9.1 Lighthouse at Honfleur

occurred already in the context of his cuts, though there with an ordering between the partitions.

Note that all three variants of Dedekind's proof use the same gestalt and metaphor and so should be regarded as one proof. When we compared these proofs it was through literal analysis and not through differences in gestalt and metaphor. The gestalt and metaphor that we attached to Dedekind's proof will figure when we compare it to other proofs.

Before we proceed to compare Dedekind's proof to that of Cantor, let us reference some graphic images that depict nicely the chain gestalt. The reflection that generates a chain can best be described as a mirror looking into itself, such as the mirror ball of a composition, by G. G. Massimei, Jr.,[20] obviously influenced by M. C. Escher's work "Hand with Reflecting Sphere".[21]

The frames of a chain (see Sect. 1.1) are like the ribs of a hand fan while the residue of the chain is like the fan's handle. One reason why a hand fan is beautiful is its mathematical content.

Frames, however, need not be continuous as the ribs of a hand fan; in general they are diffused. Consider a pointillist painting, here by Seurat, and imagine that the points in one hue are obtained by shifting the points of another hue (Fig. 9.1).[22]

The importance of the chain gestalt in expressing the difference between the finite and infinite is well depicted by comparing a composition by A. Yassin, and the graphic work of Escher "Drawing Hands".[23] In Yassin's composition (Fig. 9.2) the essence of the infinite is captured consistently by the nesting sequence of the frames, and their intersection is hinted by the square of four '2006' in the middle. In Escher's drawing, a paradox is drawn and not an aspect of the finite or the infinite. Escher's drawing necessarily leads to the assumption of an external creator while

[20] http://www.flickr.com/photos/76039389@N03/?saved=1.

[21] http://en.wikipedia.org/wiki/Hand_with_Reflecting_Sphere.

[22] http://kl.wikipedia.org/wiki/Fiileq:Lighthouse_at_Honfleur_1886_Georges_Seurat.jpg.

[23] http://en.wikipedia.org/wiki/Drawing_Hands.

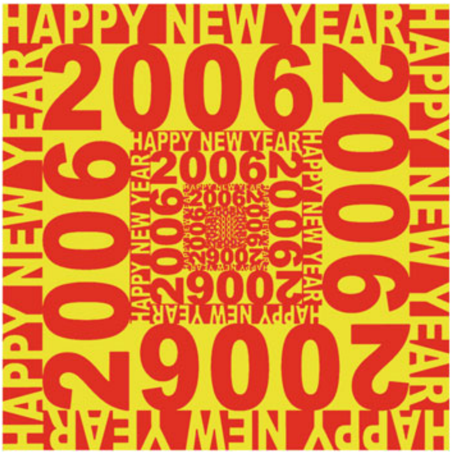


Fig. 9.2 Happy new year 2006

Yassin's forces no similar conclusion. Escher's paradox can be resolved by replacing the finite system of two hands by an infinite one: a hand drawing a hand drawing a hand – ad infinitum. This is how the theory of types resolves paradoxes. Reciprocally, it is mind boggling to imagine how a chain can turn into a paradoxical image.

9.5 Comparison to Cantor's Proof

The gestalt of Cantor's proof was that every set can be enumerated and his metaphor in the CBT proof was that the subset can be enumerated by the whole set. Clearly, there is nothing in common in the descriptors of the two proofs. In his letter to Dedekind of August 30, 1899, in which Cantor reacted to Dedekind's proof, Cantor described Dedekind's proof as simple. Surely he did not mean that it was trivial but perhaps he meant it in the sense that it was not as complex as his own proof. For the enumeration process of T by S, Cantor employed the inductive properties of his infinite numbers and the entire proof is by transfinite induction over the scale of number-classes (see Chap. 2). Dedekind's proof has no condition on the power of the set and no induction is employed. Further complication was added to Cantor's proof because of the need to comply with *Grundlagen*'s Limitation Principle. Cantor's proof also makes tacit use of the axiom of choice, in the proof of the Union Theorem, which Dedekind's proof avoids.

The idea of a chain and the pushdown metaphor used by Dedekind are absent from Cantor's proof. In fact, the given mapping between S and S′ is entirely disregarded in Cantor's proof and it is certainly not leveraged to construct the mapping between S and T. The equivalence of S and S′ is only used to assure that the enumeration process does not terminate too early.

Indeed, it is a lovely characteristic of Dedekind's proof that it follows the classic pattern of solving problems where the "given" is leveraged to construct that which is sought. Cantor was so immersed in his plan to create a scale by which all infinite (consistent) sets can be gauged that he saw, in every result obtained within this gestalt, a victory for the overall plan and he overlooked the complexity gap between the expression of CBT and the notions involved in his proof of it. Cantor disregarded Occam's razor.

Chapter 10
Schröder's Proof of CBT

Schröder's 1898 paper, "Over two definitions of finitude and G. Cantor's theorems", had two aims: (1) To transcribe to Schröder's pasigraphic system the basic notions of set theory, such as finite and infinite,[1] subset, equivalence, mapping and simple order (Cantor 1895 *Beiträge*). (2) To derive the basic relations among these notions by "analytic" considerations, that is, by manipulations of the pasigraphic formulas. The centerpiece of the paper is, however, a discussion around §2 of Cantor's 1895 *Beiträge*, in which the inequality relation between cardinal numbers is defined and the Comparability Theorem for cardinal numbers (Theorem A) is stated, along with its corollaries B-E (see Chap. 4). In this context Schröder suggested his proof of CBT, which we reproduce in detail.

The paper, which bears the date of January 1896, was presented by Schröder at the annual conference of the *Deutsche Mathematiker-Vereinigung* (DMV) held in September 1896 (see Sect. 6.5). At the time the proofs of Cantor and Dedekind, were not known and so Schröder's proof was the first presented.

In 1911 Korselt published a refutation of Schröder's proof, which we reproduce as well.[2] We also add our criticism of Schröder's proof and his style of writing. In his short paper, Korselt recounts an exchange of letters he had with Schröder in 1902. We know of the letters only from Korselt's paper. (On earlier correspondence between Korselt and Schröder see Peckhaus 2004a, footnotes 22). Korselt's letter to Schröder was sent on May 8, 1902 and Schröder responded on May 25, 1902. In his letter Korselt told Schröder about the mistake he had found; Schröder's response, partly quoted in Korselt's paper, was quite apologetic. Schröder admitted that Korselt's criticism was correct and said that he was aware of his mistaken proof

[1] Schröder compares the complementary notions 'finite' and 'infinite' as defined by Dedekind, in his *Zahlen* (1963), and Peirce, in his 1885 paper. Cantor's definition (no source is mentioned) is mistakenly identified by Schröder with that of Dedekind, perhaps because of Cantor's 1878 *Beitrag*, ignoring 1895 *Beiträge* §5, 6.

[2] Korselt's proof of CBT, provided in the same 1911 paper, we describe in Chap. 25, together with his solution to a problem raised by Schröder.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_10,

for some time, "without, indeed, having the time or opportunity to publicly retreat from it". Schröder added that about a year earlier, through an acquaintance of himself and Bernstein, Dr. Max Dehn of Münster, he conveyed a message to be passed on to Bernstein, that he "leaves him alone the honor of providing a proof of this theorem of Cantor". We do not know if the message was delivered, but Bernstein 1905 (p 121) mentions Schröder's proof without reservation so it is doubtful that he got the message from Schröder. Despite these known circumstances, the name of Schröder is, unfortunately, still linked to CBT, which is thus often called the Schröder-Bernstein Theorem or the Cantor-Schröder-Bernstein Theorem, as a quick search of the internet reveals. We believe this convention ought to stop.

Schröder further told Korselt in his letter that to close a matter of conscience he started to write Cantor a letter on the subject on September 1 (that is to say 9 months earlier!) but failed still to bring it to end. Then, 3 weeks after sending his letter to Korselt, Schröder died at the age of 62. The circumstances of his death are not clear. One suggestion is that he caught cold as he was cycling in a wintery day. Here it should be noted that Schröder was known to be engaged in various sports activities and especially he was a devout cyclist. There is a testimony from few weeks before his death that he appeared very fit. Another suggestion is that Schröder suffered from incurable illness, that he was depressed for a few months prior to his death and that he committed suicide as a result. We speculate that Schröder committed suicide because he felt dishonored by the found mistake in his proof of CBT. We support our speculation by Schröder's admittance to his inability to bring himself to admit his mistake publically, or just to Cantor. See Peckhaus 2004a, p 563f for details of these theories. Dipert 1990/1991 is another basic source on the biography of Schröder. Other papers on Schröder that should be mentioned are Peckhaus 1990/1991a, 1990/1991b, 1994, 1995a.

Though the papers of Schröder and Korselt are often mentioned in the literature on early set theory, it seems that their content was never thoroughly surveyed before. We will compare Schröder's proof to the preceding proofs of Cantor and Dedekind.

10.1 Schröder's Proof

Schröder, unlike Cantor or Dedekind, sets to prove CBT in its two-set formulation which he states as follows (p 337):

$$(a \sim_x b_1 \subseteq b \sim_y a_1 \subseteq a) \rightarrow (a \sim b \sim b_1 \sim a_1 \sim a)$$

Schröder uses $\Subset$ for both our $\subseteq$ and $\rightarrow$ (p 317). x and y are the bijections between a and b_1, b and a_1, respectively. Unlike Cantor, but like Dedekind, Schröder will attempt to construct out of x and y the bijection between a and b. Note the redundancy in the conclusion which is usually stated as a ~ b. This redundancy is typical of Schröder's style, upon which we will further comment below.

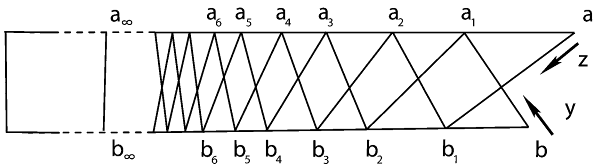


Fig. 10.1 Schröder's drawing

Some background on the environment in which Schröder operates is necessary.[3] For Schröder all elements belong to a "domain of thought" [*Denkbereich*] (Grattan-Guinness 2000, p 170, Peckhaus 2004a, p 591; compare *Zahlen* p 44 and Chap. 8). The basic notion for Schröder is "binary relation" [*binäres Relativ*] which consists of a set of pairs (i:j)[4] of elements of the domain. The pairs are considered ordered. The inverse of a given binary relation is naturally defined.[5] 1–1 mappings, such as x and y in the formulation of CBT, are introduced by (what we currently call) their graphs, the set of their ordered-pairs. In his 1895 book Schröder regarded sets as "unitary relations" (Grattan-Guinness 2000, p 171) but in the 1898 paper he speaks of sets as "binary relations" (p 307). We understand this to mean that in a Schröderian set we have all the pairs (i:j) where i ranges over the set in its regular, Cantor-Dedekind meaning,[6] and j ranges over the entire domain for every i. This understanding is necessary because Schröder speaks of the inverse of a set and the relative product (see below) of a set and a 1–1 mapping (p 323, 307).

To explicate the theorem Schröder (p 338) offers a drawing (Fig. 10.1) of which he said (p 337):

> The sets, as if they were point-sets, are presented by two parallel line segments whose endpoints connect to a trapeze. The sets can have elements in common. The latter are then (as points) just doubled: to be thought as inserted in the drawing both as objects and as images.[7]

Schröder here is seemingly conscious to the problem with the drawing in that it represents the sets, which can be of any power, by line segments that are of the power of the continuum. He continues with introducing the zigzagging lines between the line segments, saying that he now constructs a type of *Scheere* – a

[3] See Brady 2000, Chap. 7 and appendices, for an introduction on Schröder's logic and translation of a selection from Schröder 1895.

[4] For Schröder ':' is just a separator, as ',' is currently used.

[5] The inverse relation Schröder denotes by a small arc placed above the sign for the relation; we use the -1 exponent.

[6] For a Schröderian set a we will denote by $\underline{a}$ the corresponding Cantorian set.

[7] "*die Mengen so als ob sie Punktmengen wären durch parallele Strecken darstelle und deren Endponkte jeweils zu einem Trapez verbinde. Freilich können von vornherein die Mengen auch irgendwelche Elemente gemein haben. Die letzteren sind dann (wie Pnnkte) eben doppelt: als Objekte und als Bilder in der Figur eingetragen zu denken.*"

device that he describes as being made of wood and used in carnivals – [8]and he explains:

> I project forward up to infinity the projections b_1, a_1 of the sets a, b (on each other) and again their projections a_2, b_2, thereafter also the projections a_3, b_3, of them, etc. always by the same principles x, y by which the whole sets a, b were projected."[9]

Schröder failed to note that his presentation in his drawing of the subsets a_n and b_n as continuous segments is not essential and that these subsets can be diffused in a and b. Fortunately, this assumption does not affect his line of argument. The drawing did fail him but at another point (see below).

Schröder defines (p 338) the projections[10] analytically by: $a_{\lambda+1} = y;b_\lambda$, $b_{\lambda+1} = x;a_\lambda$, streamlining later (p 339) the definition by letting $a_0 = a$, $b_0 = b$. The ';' operation, which is Schröder's "relative product"[11] (Grattan-Guinness 2000, p 172; Peckhaus 2004a, p 592), takes two relations, say R and S, and composes a new one which consists of all pairs (i:j) for which there is an element h such that (i:h) is in R and (h:j) is in S. Thus, for a, x of CBT, since a is a set, x;a is also a set, it is the set b_1, because in x we have the pairs (i:j) with i belonging to $\underline{b}$ and j to $\underline{a}$ and x;a consists of all the pairs (i:k) where the i′s are the same as in the pairs of x (i.e., from $\underline{b}$) but k ranges over all members of the domain for each such i.

Schröder notes (p 338) that $a_{\lambda+1} \subseteq a_\lambda$, $b_{\lambda+1} \subseteq b_\lambda$, as a result of a theorem which he calls Dedekind's theorem (p 318), presumably after *Zahlen* p 55 #35, and we call "Dedekind's Lemma". Schröder calls (p 328) sequences of nested sets, such as the a_λ, by Dedekind's term *'Kette'* (chain, see Sect. 10.3). By substitutions he gets:

$$a_{2\lambda} = (y;x)^\lambda;a,\ b_{2\lambda} = (x;y)^\lambda;b,$$
$$a_{2\lambda+1} = (y;x)^\lambda;a_1 = y;(x;y)^\lambda;b = (y;x)^\lambda;y;b,$$
$$b_{2\lambda+1} = (x;y)^\lambda;b_1 = x;(x;y)^\lambda;a = (y;x)^\lambda;x;a.$$

For the second and third lines Schröder applies the "very easy to confirm" theorem (**) $x;(y;x)^\lambda = (x;y)^\lambda;x$ and the one with x, y interchanged (p 339, 54). The correct order in these lines, it seems, would have the third and fourth terms interchanged and then the lemmas applied.

To avoid exceptions in the use of these general formulas Schröder adds the convention that for an arbitrary relation z, z^0 is the identity relation, which he

---

[8] Perhaps he had in mind something that operates like scissors [*Schere*], and is used, for instance, to grip an object on a higher shelf. Another possibility is the toy called Jacob's ladder, see http://en.wikipedia.org/wiki/Jacob%27s_ladder_(toy).

[9] "*ich projizire bis ins Unendliche fort auch die Projetionen b_1, a_1 der Mengen a, b (aufeinander) und wiederum deren Projektionen a_2, b_2, sodann auch die Projektionen b_3, a_3 von diesen,* etc. *immerfort je [338] nach demselben Prinzipe x resp. y, nach welchem schon deren Ganzmengen a resp. b projizirt wurden*."

[10] Note that Schröder calls 'projections' the images of the sets; Borel will call projections the mappings between the sets (see the next chapter).

[11] "*relative Produkt*"; we bring the common translation though perhaps it would better be translated as 'product of relations'.

denotes by $1'$. It includes all pairs (i:i). 1 denotes for Schröder the universal relation which includes all pairs (i:j).

From the stated equalities Schröder then draws the equivalences:

$$a \sim_x b_1 \sim_y a_2 \sim_x b_3 \sim_y a_4 \sim_x b_5 \ldots \sim_y a_{2\lambda} \sim_x b_{2\lambda+1} \sim_y a_{2\lambda+2} \sim_x \ldots$$
$$b \sim_y a_1 \sim_x b_2 \sim_y a_3 \sim_x b_4 \sim_y a_5 \ldots \sim_x b_{2\lambda} \sim_y a_{2\lambda+1} \sim_x b_{2\lambda+2} \sim_y \ldots$$ [12]

Schröder now points out that to prove CBT it is enough to prove that any of the even and of the odd a's (or b's) are equivalent. Having obtained four nesting sequences of sets, represented by: (*) $a_{2\lambda+1}$, $a_{2\lambda}$, $b_{2\lambda+1}$, $b_{2\lambda}$, Schröder concludes that each sequence must converge for lim $\lambda = \infty$ and have a limit value.

He explains the notion of convergence of binary relations thus (p 340f):

> [I]t is sufficient by my research (vol. 3 [Schröder 1895], p 180ff) that either the empty places or the occupied (with an eye)[13] places of the matrix of u_λ to appear 'definitive', namely, that with increasing λ, they permanently hold as such. If generally $u_\lambda \subseteq u_{\lambda+1}$ the latter is rather the case [the occupied places are retained], as every eye of u_λ must be found in $u_{\lambda+1}$ and any higher u_λ; the limit u_∞ will then be achieved accumulatively (growing 'monotonously'). When in contrast generally $u_{\lambda+1} \subseteq u_\lambda$, must every empty place of u_λ be also an empty place of $u_{\lambda+1}$ and all following u_λ and the limit value will be achieved by ('monotone') decrease. This case applies to each of our four relations.[14]

Obviously Schröder's limits here are the union and intersections of binary relations sets which he introduced in his 1895 volume. An "empty place" seems to be a pair (i:j) of elements from the domain which is not in u_λ.

With regard to the four nesting sequences (*) this definition leads Schröder to four limits which he designates by $a_{2\infty}$, $a_{2\infty+1}$, $b_{2\infty}$, $b_{2\infty+1}$. From the relations $a_{2(\lambda+1)} \subseteq a_{2\lambda+1} \subseteq a_{2\lambda}$ (and respectively for the b's) Schröder concludes (no argument is supplied) that $a_{2\infty} \subseteq a_{2\infty+1} \subseteq a_{2\infty}$ and hence that $a_{2\infty} = a_{2\infty+1}$ which he therefore denotes by a_∞; similarly he derives b_∞. Schröder relies here (p 341) on the schema $(\alpha \subseteq \beta \subseteq \alpha) \leftrightarrow (\alpha = \beta)$ (Schröder uses '=' for our '$\leftrightarrow$').

---

[12] In the parts after the middle ellipsis there is seemingly a typo in the original: the x and y are interchanged.

[13] It seems that Schröder means by 'eye' a pair, and by 'matrix' the set of pairs of the binary relation.

[14] "*ist es nach meinen Untersuchungen (Bd. 3* [Schröder 1895] *p 180 sqq.) hinreichand, dass entweder die Leerstellen oder die (mit Auge) besetzten Stellen der Matrix von u_λ sich als "definitive" erweisen, die nämlich bei wachsendem λ sich permanent als ebensolche forterhalten. Ist etwa allgemein $u_\lambda \subseteq u_{\lambda+1}$, so findet das letztere statt, indem sich jedes Auge von u_λ, auch beim $u_{\lambda+1}$ und [341] bei allen höheren u_λ finden muss; die Grenze u_∞ wird dann zunehmend ("monoton" wachsend) erreicht. Wenn dagegen allgemein $u_{\lambda+1} \subseteq u_\lambda$ ist, so muss jede Leerstelle von u_λ auch eine solche von $u_{\lambda+1}$ und allem folgenden u_λ sein und der Grenzwert u_∞ wird ("monoton") abnehmend erreicht. Dieser Fall nun liegt bei jedem unsrer vier Relative 57)* [our (*)] *vor.*"

Now comes the punch line (p 341):

> But because all odd a, that is, all $a_{2\lambda+1}$, must be equivalent with $a_{2\infty+1}$, as must be all even $a_{2\lambda}$ with $a_{2\infty}$, from $a_{2\infty} = a_{2\infty+1} = a_\infty$ it is proved that those [the odd a′s] are equivalent with these [the even a′s] and thus all a_λ are equivalent.[15]

The fallacy of this conclusion will be discussed below.

Schröder goes on trying to obtain the binary relation that provides the equivalence of a and b. Because, for example, $a_{2\lambda} = (y;x)^\lambda;a$, it seems that Schröder was convinced that the sequence $(y;x)^\lambda$ must have a limit $(y;x)^\infty$ such that $a_\infty = (y;x)^\infty;a$. Schröder is cautious enough to say (p 341):

> We cannot simply maintain that the exponent relations $(y;x)^\lambda$ and $(x;y)^\lambda$ must converge. This needs to apply only to the parts of these relations, which operate internally on the sets a, b, etc., that is, more exactly: which falls into the columns of the converging y;b and a, respectively x;a and b, and likewise in the line of the other pairs of sets. Only by suitable restriction of the mapping principles x, y by the adventitious condition mentioned in p 323 can it be suggested that also those exponents converge.[16]

Making sense of this paragraph is not simple. The adventitious condition to which Schröder refers deals with the following: when we have that $a \sim_z b$, it may occur that z also relates elements from the complements of a and b (to the complements or to the sets, thus breaking the 1–1 outside a, b). In this case, Schröder asserts, it is possible to replace z by $a^{-1}bz$,[17] because we have $a \sim_{a^{-1}bz} b$; in the new relation only members of a and b, not of their complements, are related. What Schröder thus says is that for the exponent relations to converge we have for each λ to restrict the x and y which appear in $(y;x)^\lambda$ and $(x;y)^\lambda$ to certain subsets of a, b; but to which such subsets he does not clearly say. Still, he denotes by $(y;x)^\infty$ and $(x;y)^\infty$ the limits of the sequences $(y;x)^\lambda$ and $(x;y)^\lambda$, respectively, and he concludes (p 342):

$$(a \sim_x b_1 \sim_{(x;y)^\infty} b_{2\infty+1} \sim_{1'} b_{2\infty} \sim_{(y^{-1};x^{-1})^\infty} b) \rightarrow (a \sim_{(y^{-1};x^{-1})^\infty;(x;y)^\infty;x} b),$$

$$(b \sim_y a_1 \sim_{(y;x)^\infty} a_{2\infty+1} \sim_{1'} a_{2\infty} \sim_{(x^{-1};y^{-1})^\infty} a) \rightarrow (b \sim_{(x^{-1};y^{-1})^\infty;(y;x)^\infty;y} a).$$

Using the equality $x;(y;x)^\infty = (x;y)^\infty x$ which he obtains from (**) (see two pages above) by passing to the limit, Schröder asserts (p 342) that the equivalence $(y^{-1};x^{-1})^\infty;(x;y)^\infty;x$ (of a and b) can be presented by $(y^{-1};x^{-1})^\infty;x;(x;y)^\infty$ and the equivalence $(x^{-1};y^{-1})^\infty;(y;x)^\infty;y$ (of b and a) by $(x^{-1};y^{-1})^\infty;y;(y;x)^\infty$.[18] The inverse of the latter (when applying, also for the limit relations, the known rule that the

[15] "*aber alle ungeraden a, d. h. alle $a_{2\lambda+1}$, mit $a_{2\infty}$ alle geraden a, d. h. alle $a_{2\lambda}$, gleichmächtig seinmüssten, so ist mit $a_{2\infty+1} = a_\infty = a_{2\infty}$ auch die Gleichmächtikeit jener mit diesen, somit die sämtlicher a_λ erwiesen*".

[16] "*Man kann nicht ohne weitres behaupten, dass die Relativpotenzen $(y;x)^\lambda$ und $(x;y)^\lambda$ selbst konvergiren müssten. Dies braucht nur bei dem Teile dieser Relative zuzutreffen, der auf die Mengen a, b* etc.*, intern wirksam ist, d. h. genauer: der in die Kolonnen des Konversen von y;b und a resp. x;a und b und zugleuich in die Zeilen des andern Mengenpaars hinheinfält. Erst mittelst geelgneter Einschränkung der Abbildungsprinzipien x, y durch die S. 323 erwähnten Adventivbedingungen dürfte es hinzubringen sein, dass auch jene Potenzen gelbst konvergiren.*"

[17] For the intersection of sets Schröder just juxtaposes the sets.

[18] There is a typo here in the original and the middle y appears as y^{-1}.

inverse of a composite relation is the composite of the inverse relations in reverse order) is equal to $(y^{-1};x^{-1})^{\infty};y^{-1};(y;x)^{\infty}$. This, says Schröder, is not the inverse of the first equivalence, however, the difference is only formal: using the rule that $\infty + 1 = \infty$ we get from the first equivalence its second form $(y^{-1};x^{-1})^{\infty};y^{-1};x^{-1};x;(x;y)^{\infty}$ which is equal to $(y^{-1};x^{-1})^{\infty};y^{-1};(y;x)^{\infty}$ which is the inverse of the second equivalence. Hence Schröder arrived at showing that the equivalence of a and b is the inverse of the equivalence of b and a.

10.2 Criticism of Schröder's Proof

Korselt's first point is that the limit sets $a_{2\infty}$, $a_{2\infty+1}$, $b_{2\infty}$, $b_{2\infty+1}$, (on p 341) are introduced implicitly [*stillschweigend*]. This is not quite true because, as we have seen, Schröder explicitly discusses the conditions under which we can speak of the limit of certain sequences of sets, such as the sequences (*) (see two pages above). What Korselt may have meant is that Schröder did not point out explicitly that the limits of which he speaks are simply the intersection of the nesting sequences.

Korselt's second point is that, contrary to Schröder's claim, the limit set may be of smaller power than the members of the converging sequence. After bringing a short summary of Schröder's proof, mainly the nested sequences (*) and the passage to the limit, Korselt gives a simple example to disprove Schröder's claim that each limit is equivalent to the members of its sequence:

> Let a_0 and b_0 be the set {1, 2, 3, . . ., ad infinitum},
> $a_{2\lambda} = b_{2\lambda} = (1 + 6\lambda, 2 + 6\lambda, \ldots)$, $a_{2\lambda+1} = (3 + 6\lambda, 4 + 6\lambda, 5 + 6\lambda, \ldots)$,
> $b_{2\lambda+1} = (5 + 6\lambda, 6 + 6\lambda, \ldots)$. Then $a_{2\infty} = a_{2\infty+1} = b_{2\infty} = b_{2\infty+1} =$ Nothing,
> $a_{2(\lambda+1)} \subseteq a_{2\lambda+1} \subseteq a_{2\lambda}$, $b_{2(\lambda+1)} \subseteq b_{2\lambda+1} \subseteq b_{2\lambda}$, and hence for no finite λ
> $a_{2\lambda} \sim a_{2\infty}$, $a_{2\lambda+1} \sim a_{2\infty+1}$, as it should have been, according to Schröder.
> The 'transfinite induction' fails here.

Korselt's pun with 'transfinite induction' is directed at Schröder's incautious passage from reasoning on finite numbers to reasoning at the limiting case; here 'transfinite induction' does not have the meaning it acquired since Hausdorff 1914a (though already used, half voiced, by Cantor – see Sect. 1.1) and it simply means: inducing from the finite cases to a case that lies beyond the finite. A similar argument was raised by Harward against Jourdain in 1905 – see Sect. 17.5.

In his answer to Korselt, Schröder explains how he noticed the failure of his conclusion: when in his original drawing, the right side is symmetrically reproduced to the left of the points a_{∞}, b_{∞}, it becomes obvious that the sets degenerate to a point. Schröder reveals here what failed him: he used an argument based on his arbitrary drawing. A similar situation, suggested Schröder, this time with a new drawing (Fig. 10.2), would occur: "If a linear point-set on a line between

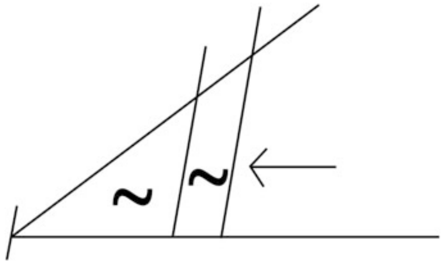


Fig. 10.2 Schröder's new drawing

the sides of an angle, is proven to have – by parallel displacement towards the vertex – the same power as the vertex."[19]

Schröder's new example seems to be: take the line segment between the sides of the angle to which the arrow points; any other segment between the first and the vertex, parallel to the first, is equivalent to the first by rays emanating from the vertex. However, the limit of all these segments is the vertex, which is surely not equivalent to any of them. Compare Schröder's new drawing to Cantor's drawing from his 1878 *Beitrag* (see 3.5).

It is interesting that Korselt did not criticize Schröder's argument for his conclusion that the residue is equivalent to the members of the converging sequence, namely the definition of the limit of the exponent relations, e.g., $(y;x)^{\infty}$, which is not only vague but, indeed, unacceptable.[20] The problem with trying to extrapolate this limit from $(y;x)^{\lambda}$, is this: (y;x) relates every member of $a_{2\lambda}$-$a_{2\lambda+1}$ to a member of $a_{2\lambda+2}$-$a_{2\lambda+3}$, while $(y;x)^2$ relates every member of $a_{2\lambda}$-$a_{2\lambda+1}$ to a member of $a_{2\lambda+4}$-$a_{2\lambda+5}$. Thus the two relations (y;x) and $(y;x)^2$, as sets of pairs of the corresponding elements, are disjoint! So there is no limit, according to Schröder's "adventitious condition" to the sequence of $(y;x)^{\lambda}$.

Schröder may have been aware of the problem when he suggested that to restrict the relations x, y in a way that obeys his adventitious condition. Thus we can construe $(y;x)^{\lambda}$ not by repeated application, λ times, of (y;x) but as the restriction of x, y to $a_{2\lambda}$ and $b_{2\lambda+1}$, respectively. The sequence of these relations does fulfill Schröder's adventitious condition and their limit are automorphisms of the residues from which one can obtain, using x or y, the equivalence of the residues and, it seems, even the formal relations: $x;(y;x)^{\infty} = (y;x)^{\infty};x$, $((y;x)^{\infty})^{-1} = (x^{-1};y^{-1})^{\infty}$, $(y;x)^{\infty} = (y;x)^{\infty+1} = (y;x)^{\infty};x$ which Schröder (p 342) was very proud to have obtained, but not the equivalence between the sets and the residues.

---

[19] "*wie wenn man für eine lineare Punktmenge auf einer Geraden zwischen den Schenkeln eines Winkels – durch Parallelverschiebung dieser Geraden bis in den Scheitel hin – die Gleichmachigkeit derselben mit, diesem Scheitelpunkte beweisen wollte*."

[20] Korselt (1906 p 218) criticizes as unacceptable a similar passage to infinity made in Schoenflies 1906 p 25, though, it seems to us, Schoenflies did not commit the error made by Schröder and his definition of S_{ω} there can be corrected. See our reference above to Harward.

In the opening of his paper, Korselt introduced his criticism of Schröder's proof by the following words: "Schröder's proof contains, however, a faulty conclusion, until now unnoticed by the mathematicians". It is indeed remarkable that the important mathematicians that dealt in early set theory, who often reference Schröder's proof, made no comment on its validity, even though the problems in the proof are quite obvious.

On the other hand, it is plausible that some unvoiced reservation from the proof was there, even before Korselt's paper, for no one ever reproduced the proof. Borel (1898) and Poincaré (1906b) did not even mention Schröder's name with respect to CBT, only that of Bernstein. Russell (1903 pp 306, 367) mentions both names but references only Borel and Zermelo 1901.

Interestingly, Schröder 1896 is often referenced as the location of Schröder's proof. So did even Bernstein (1905). Peano (1906, pp 136, 137f, 141) mentions both Bernstein's proof and that of Schröder, referencing for the latter besides the 1898 paper also the lecture of 1896 (which he dates to May instead of September). Hessenberg 1906 (p 522) describes Schröder's proof as operating through logical calculations and that it is only sketched, so that it seems that he referred to the 1896 paper and not to that of 1898.

In view of all these sources, which do not make much sense, our impression is that some of the references to Schröder's proof were made by simply quoting prior references without reading Schröder's papers. Others were perhaps familiar with the proofs of Borel or Schoenflies and were able to easily port rigorous arguments to cover Schröder's blunder.

In fact, it is possible that because the readers of Schröder 1898 were eminent mathematicians, they were not interested in stopping to comment on its defects. As most of them must have shared with Cantor the view that Schröder's pasigraphy is "superfluous to mathematics" (Grattan-Guinness 2000 p 175), they surely read the proof with a machete, cutting through Schröder's verbal foliage to its core mathematical gestalt and metaphor. Then, when they got the picture they moved on. It seems that we have here an example, unfortunately not documented, of the way mathematicians read texts that are only obliquely related to their field of interest. The circumstances of Schröder's sudden death could have been another reason why criticism of his CBT proof was delayed.

We do not think that the above reasons can explain why in the modern literature only passing criticism is launched against Schröder's style of writing. Dugac (1976 p 80) quotes from Dedekind his appraisal that Schröder's 1873 book on arithmetic and algebra "contains much that is good, but also much that is superfluous". Muddox (1991 p 423) quotes from Peirce his judgment of Schröder's vol 3: "its glaring defect of involving hundreds of merely formal theorems without any significance". Muddox balances this view on Schröder when he quotes (1991 p 449) Tarski saying that Schröder's book (1895) "contains a wealth of unsolved problems". The two views do not contradict and rather shed some light on the different approaches of the mentioned personalities, no doubt at least partly a result of their different time perspectives.[21] With these remarks we can perhaps understand more broadly Cantor's opinion quoted above.[22]

Still, these in-passing comments only hide more than they reveal; Schröder's style is typically obsessive, with much attention given to minute trivialities while intricacies of rigor are neglected. Here are some particularly striking examples:

- In p 317 Schröder presents the following three theorems:

  24) $(a \subseteq b \sim c \subseteq a) \rightarrow (b \sim a \sim c)$,
  24_1) C. $(a \subseteq b \subseteq c \sim a) \rightarrow (a \sim b \sim c)$, 24_2) $(a \sim b \subseteq c \subseteq a) \rightarrow (b \sim c \sim a)$

  These are all variations on the single-set formulation of CBT. This is indicated explicitly in 24_1 which references Corollary C of Cantor's 1895 *Beiträge* §2. The three variations differ in the position of ~ in the premises and in the order of the equivalent sets in the conclusions. But all three formulations say the same thing: that if a set is equivalent to a subset of its subset than it is equivalent to the latter. The only difference is in the use of different letters to name the sets. There is no difference in the mathematical content of the three variants! Still Schröder finds it important to present all three and to draw our attention (p 328) to the fact that they result from permutations of the letters; he says: "as schemata for the applications it is in no way redundant to have the theorem ready in all its three expressions".

- A simpler example of Schröder's obsessive style, is his need to state the same thing twice or thrice in different words when no clarification is actually gained by the repetition. For example, Schröder never gets fatigued of saying "a proper part-set or subset" ("*echte Teilemenge sive Untermenge*") when he refers to a proper subset (e.g., p 337).
- Here is an example of Schröder's ability to cloud the essence by the trivial. Says Schröder (p 337) with regard to CBT:

  > It is important to notice and this is the kernel of the theorem, what the theorem expresses when the subsets a_1, b_1 are proper subsets.

  It is a parody of mathematical style to say that the kernel of the theorem is in the non-trivial cases. Schröder even gives this "essential" case a special name: the '$\subset$ case', and he repeatedly takes pains, or perhaps pleasure, in providing his various expressions, regarding the relations between the sets in the nested sequences, separately for the $\subset$ case.[23] Finally, the reason for the distinction

[21] On Russell's view of Schröder see Anellis 1990/1991. On Wiener's view of Russell's view of Schröder, see Grattan-Guinness 1975 and Brady 2000. For Frege's view on Schröder see Frege 1895. On Peirce's view on Schröder see Dipert 1990/1991, p 15.

[22] Grattan-Guinness (2000, p 171) added his own opinion on Schröder's discussion that it was with "painstaking detail".

[23] Note, and Schröder was aware of this, that Cantor only presented the non-trivial case, as he considered only proper subsets.

of the $\subset$ case emerges: Schröder draws our attention to the fact that the $\subset$ case may seem to some like a paradox (p 343), because the limit in this case, of one of the nested sequences, is not a proper subset. A proper subset of what, you may ask; well, Schröder is not very explicit here, as was his custom when reaching points of rigorousity, but it seems safe enough to say that it is not a proper subset of the limit of the sequence when the $\subseteq$ is used; actually both limits are equal!

Schröder further continues to compare this potential paradox to other three famous "paradoxes" of mathematics: that of Achilles and the turtle, that of the change in the angle of a tangent to the graph of a function at a point of non-differentiability (if we understood him correctly) and especially the case of three monotonous descending sequences of numbers, which fulfill $a_i < c_i < b_i$ but converge to the same limit. Schröder concludes (p 344) that the reason why no paradox emerges in his proof of CBT is an achievement due to his well-executed theory of limits, and he praises his drawing (!) which enables seeing the by-pass of the paradoxes (whatever this means).

We don't know how to consolidate Schröder's importance in the history of the algebra of logic (Peckhaus 2004a p 557, Brady 2000) with his problematic style, so let us only point out that Schröder's style is not all bad: Schröder has this good pedagogic point of stating what he wants to do next before he does it formally, much in line with Lakatos' (1976) suggestion on how to improve the Euclidean style of presentation. Also it is noteworthy that despite his fervor for pasigraphy and redundant rhetoric, Schröder is using images as heuristic. For example, in the context of his lengthy discussion aimed to prove his theorem 21 (p 317) which says: $a \sim d \subseteq b \rightarrow \exists c(a \subseteq c \sim b)$,[24] Schröder wants to add to a the elements of b-d, only he has to make sure that the addition is disjoint from a so he suggests placing a mirror (p 325) in front of a∩(b-d) and replacing in b-d these elements of the intersection by their images in the mirror, or, alternatively, by the elements obtained from them after an accent is added; the new b-d obtained is then adjoined to a to obtain the required c. Since Schröder, like Dedekind, assumes that all elements of the sets considered are from the "domain of thought" [*Denkbereich*], the mirror images (or the "accented elements") are in no way inferior to the already available elements of b. Schröder disregards, however, the possibility that his new elements can still be elements of a (though not of a∩(b-d)); also he seems to hold that there is no difference between an element and its name, when suggesting the accent option. Such details of rigorousity Schröder typically disregards.[25]

[24] We are using $\exists$ instead of Schröder's Σ.

[25] See our note on Schröder's explication of his first drawing regarding the possible diffusion of the subsets in the set. For an example of Schröder's aptitude for pictorial descriptors, from 1890, see Peckhaus 2004a, p 582. Incidentally, theorem 21 had a nice future, see Sect. 25.2.

10.3 Comparison with Cantor and Dedekind

Schröder's gestalt is the *Scheere* of nesting odd and even sequences of subsets of the given sets. His metaphor was to port the equivalence of the members of the sequences to their limits (their intersection). 'Limit' was likely a common metaphor at the time in situations when focus shifted from an infinite sequence to what it culminates to. It obviously had its roots in analysis. Thus, Cantor used it in his definition of 'limit' infinite numbers. Schröder\s metaphor led him to the correct results that the limits are equivalent, but his argument was wrong, and his other conclusion, that the limits are equivalent to the members of the sequences, was mistaken.

Schröder's proof is clearly remote from the proof of Cantor. The latter had the gestalt of the scale of number-classes and the metaphor of enumerating the subset by the whole set. Schröder had no numbers and no numbering.

Schröder, like Dedekind, perceived CBT's proof in the language of sets and mappings (both were for him 'binary relations') – the notions used to articulate the theorem. It is perhaps by processing Dedekind's notion of chain (see Sect. 9.1), to which Schröder devoted considerable attention in his 1895 volume (Grattan-Guinness 2000 p 172, Brady 2000 §7.3.2), that Schröder uncovered the nesting sequences implicit in the conditions of CBT. But Schröder had not noticed the partitioning and pushdown the chain metaphors. Perhaps because he worked on the two sets formulation in which this metaphor is less obvious.

On the other hand, Dedekind never mentioned the residue of a chain (its intersection). Perhaps because he was mainly interested in chains generated by a set, and in particular 'simple chains' generated by one element, and such chains have an empty residue.

It is possible that it was because of its use of chains that Cantor regarded (in his letter to Dedekind of August 30, 1899) Dedekind's proof of CBT to be in accord (*stimmt überein*) with that of Schröder. Another similarity between Schröder's and Dedekind's proofs, which Cantor may have found, is that both named the given equivalences, the x, y in Schröder's proof, something which Cantor never did, and that in both, an attempt is made to define the required equivalence by means of the given ones.

Other points to note are these:

- Schröder used pasigraphy while Cantor and Dedekind used only natural mathematical language in their presentations of CBT.
- Schröder's attempted proof of CBT was for its two-set formulation; Cantor and Dedekind proved the theorem in its single-set formulation, with Dedekind noting how the two-set formulation can be obtained, a point that Cantor never explicitly made.
- Like Cantor and Dedekind, Schröder refers to 'equivalence' with the terms "law" or "principle" (Schröder 1898 p 337), but contrary to both, he perceived 'equivalence' as binary relation rather than as mapping. In this Schröder took part in the trend to distance the notion of function from its analytic origin to a set theoretic characterization and he may be seen as a forerunner to Zermelo's 1908b definition of mappings as sets of pairs.

Chapter 11
Bernstein, Borel and CBT

In the Dictionary of Scientific Biography (Gillespie 1970–1980 p 58), the circumstances of Bernstein's finding of his CBT proof are described: The date of the finding is given as "1895 or 1896 while [Bernstein was] a student in the Gymnasium" and it is said that:

> Cantor, who had been working on the equivalence problem, had left for a holiday and Bernstein had volunteered to correct proofs of his book on transcendental numbers. In that interval, the idea for a solution came to Bernstein one morning while shaving. Cantor then worked with this approach for several years before formulating it to his satisfaction.

The dictionary seems erroneous in this passage. First, there is no book of Cantor on transcendental numbers. Second, according to Cantor himself (letter to Dedekind of August 30, 1899), Bernstein first presented his proof of CBT, in a Halle seminar in Easter 1897 (which was in April, cf. Purkert-Ilgaud 1987, p 139). Third, Bernstein was born in February 1878 and so he was 19 years old at the time of his discovery, so most likely already a student at the university not in the gymnasium. Fourth, the suggestion that Cantor was working at the time on CBT conflicts the strong evidence that Cantor worked on CBT in 1882. Fifth, it is not conceivable that Cantor continued to work on the theorem for several years following Bernstein's approach for the theorem is simple and the proof simple too.

In Bernstein's internet biography[1] the story is slightly changed and the book is Cantor's 1895/7 *Beiträge*, though Bernstein is still depicted as a gymnasium student. In Ebbinghaus 2007 the "book" is the more reasonable 1897 *Beiträge* and the time is the winter semester of 1896/7,[2] which fits the letter to Dedekind. Then we can understand the reference to transcendental numbers to mean the ordinal numbers, which are the subject of Cantor's 1897 *Beiträge*.

It is, however, possible that Bernstein found the proof when he was reading the proofs of 1895 *Beiträge*, in early 1895, when he was still in the gymnasium. Then

---

[1] http://www-history.mcs.st-and.ac.uk/Biographies/Bernstein_Felix.html.

[2] It is the dating in Frewer 1981.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_11, © Springer Basel 2013

the cardinal numbers are the transcendental numbers mentioned in the dictionary. Under this explanation the presentation of the proof in the Halle seminar was 2 years after the discovery of the proof.

The anecdote of the Dictionary of Scientific Biography about the shaving is, nevertheless, intriguing because we can imagine Bernstein in front of a mirror and by placing another mirror in front of the first, experiments the conditions of CBT. Then must have occurred the inspiring moment of gestalt switch, when he concentrated on the frames of the images of the mirrors instead of concentrating on the images of the whole mirrors. We will get back to this thought experiment.

Strangely, Bernstein never published his original proof.[3] Instead, it was published by Borel, in an appendix on the notion of power to his book on the theory of functions (1898 pp 102–107). There Borel recounted (p 103 footnote 3) that the proof was communicated to him by Cantor, whom he approached during the first international congress of mathematicians held in Zurich in August 1897.

In the appendix, Borel thanked Cantor for the permission to publish the proof. Two things appear peculiar in Borel's story: That Cantor made a decision on Bernstein's behalf and that he ignored the benefit to Bernstein that could have resulted if Bernstein himself would have published the result. With regard to the first point, the answer could be that Cantor had a paternalistic attitude towards Bernstein, who was the son of his friend and colleague (Gillespie 1970–1980 p 58). With regard to the second point, perhaps Cantor did not regard the proof as very important because he himself already had a proof of CBT in 1882 (see Chap. 1). In addition, Cantor was probably aware, when he met with Borel, of Schröder's attempted proof of the theorem. As Cantor told Dedekind in the August 30, 1899, letter, he considered Bernstein's proof to be similar to that of Schröder. So, Cantor perhaps thought that Bernstein's proof was a nice finding by a young man but no great discovery in set theory research. Still, Cantor had high regard for the proof, perhaps because the proof was not leaning on Cantor's scale of number-classes but still used his idea of abstraction (see below). In the dictionary (p 58) it is said that Bernstein was influenced to forsake his studies in fine arts at the university of Pisa and return to mathematics, by two mathematics professors in Pisa who heard his praise from Cantor at a mathematical congress (perhaps the very same congress when Borel met Cantor).

As it happened, Bernstein was not negatively affected by Borel's publication. He still got his name attached to an important theorem though he never publicly produced his original proof for it. What is nevertheless strange, is that even in his doctorate dissertation of 1901, published in his 1905 paper, he only mentioned his proof in a by-the-way manner, referencing its presentation to Borel's book (Bernstein 1905 p 117 footnote). His reason could have been that by then, with the publication of the proofs of Schröder (1898) and Zermelo (1901) which he mentions in his paper, the importance of his own proof was diluted. We believe, however, that such an attitude is unnatural and that Bernstein had a deeper reason

[3] In 1906 Bernstein did publish a proof of CBT, in the context of the debate between Poincarè and the logicists, which resembled a 1906 proof of Peano (see Sect. 21.3).

for not presenting his proof: it seems to us that Bernstein's original proof was cruder than the one presented by Borel. So, Bernstein stood nothing to gain by repeating his original argument. We will first review the proof of CBT given by Borel in 1898. Then we will interpolate the original argument of Bernstein, never published. We will then compare both proofs to the earlier proofs of Cantor, Dedekind and Schröder.

11.1 Borel's Proof

We quote Borel's proof from his 1898 book (p 104ff) with our comments in the footnotes.

> There exists a proper subset[4] A_1, of A, that has same power[5] as B and there exists a proper subset B_1, of B, that has same power as A. It is required to prove that A has same power as B.[6]
>
> As B and A_1 have same power, there exists a projection[7] from B on A_1,[8] that is to say a law after which the elements of B and of A_1 correspond in a unique and reciprocal fashion.[9] There exists even an infinity of such projections;[10] but we choose from them one well determined.[11] It is clear that, by such a projection, to every proper subset of B correspond a proper subset of A_1; let A_2 be the proper subset of A_1 that thus corresponds to B_1; A_2 has, by the very definition of the power,[12] same power as B_1 and, as a result, same power as A.

---

[4] We use 'subset' for Borel's 'partie'.

[5] Borel uses 'same power' (*méme puissance*) for Cantor's 'equal power' (*gleiche Mächtigkeit*). Borel avoids using 'the' before 'same power' surely to avoid any interpretation that a power is an entity. Borel rejected this possibility, despite the respect which he expresses for Cantor's results, because he thought that Cantor had not resolved the question of comparability. Borel, however, thought that the notions of one set having a smaller power (Borel 1898 p 104) than another, or of two sets having the same power, notions introduced in Cantor's 1878 *Beitrag* (see Sect. 3.1), are clear. Hence, he was ready to discuss CBT. Still it is strange that he is using the term "same power" at all, for Cantor introduced the terminology "same power" in his 1878 *Beitrag* together with its synonym "equivalence" and Cantor himself, after 1878 *Beitrag*, preferred the latter (1887 *Mitteilungen* p 413, 1895 *Beiträge* §2).

[6] Cantor and Schröder also used for the subsets the same letter used for the whole-set, but Bernstein 1905 (p 121) used for the subsets the letter used for the sets to which they are equivalent.

[7] This is Borel's term for '1–1 mapping'.

[8] Borel is using here Cantor's definition of 'same power' (or 'equal power') from the first paragraph of Cantor's 1878 *Beitrag*.

[9] Like his predecessors, Cantor (1932 p 283), Dedekind (1963 p 50) and Schröder (1898 p 337), and some later mathematicians too (J. Kőnig 1906), Borel characterizes 'projection' as a law after which the elements of the two sets correspond. For some notes on the historical development of the notion of mapping from the analytic notion of function, see Lakatos 1976 pp 151–152.

[10] This observation is also taken from the first paragraph of Cantor 1878 *Beitrag*.

[11] This choice, which Borel repeats at every step of his construction, involves the axiom of choice, quite unnecessarily, in Borel's proof.

[12] In the original "*d'après la definition meme de la puissance*"; it seems that the definite article was inserted here by mistake.

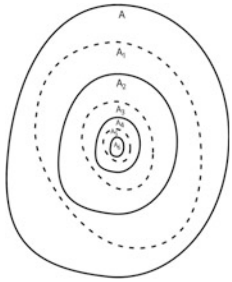


Fig. 11.1 Borel's drawing for his CBT proof

Moreover, A_2 is a proper subset of A_1, which is itself a proper subset of A. It all comes down to saying that, A_2 having same power as A, A_1 also has same power as A (because B has same power as A_1).[13]

By the hypothesis A_2 has same power as A; we choose a determined projection from A on A_2; A_1 which is a proper subset of A becomes a proper subset A_3 of A_2, and A_2 becomes a proper subset A_4 of A_3. This is what we indicate on the schematic figure below, where the projection from A on A_2 is an homothetic transformation.[14] Therefore A_3 has same power as A_1 and A_4 same power as A_2 and as A. If we project A on A_4, A_1 and A_2 will project upon A_5 and A_6; A_5 will be a proper subset of A_4 and will have same power as A_1 and A_3; A_6 will be a proper subset of A_5 and will have same power as A,[15] A_2, A_4. Thus continuing, we form a sequence of sets A, A_1, A_2, A_3, A_4, . . . such that each of them is a proper subset of its preceding set and in addition such that all the sets of even index have same power as A and all the sets of odd indices have same power as A_1 (Fig. 11.1).

This sequence can be continued indefinitely, that is to say that the set A_n is defined, whatever the integer n; it contains, incidentally, elements, because it has same power as A or A_1, accordingly if n is even or odd.

Consider now[16] the set D formed of the elements common to all the sets A, A_1, A_2, . . ., A_n, ... This set D can contain, incidentally, no element. It is clear that this set D can be

[13] At this point Borel shifts the proof from proof of the two-set formulation of CBT to proof of its single-set formulation.

[14] Namely, 'shrinkage'. Obviously, for the general case, the geometric aspects of the figure ought to be disregarded. It is interesting that in many of the early proofs of CBT a figure is used.

[15] There is here a typo in the original where it is written 'A_1' instead of 'A'.

[16] Here Borel inserted the following footnote (p 105): "By introducing the set D, I modify lightly the demonstration that Mr. G. Cantor had communicated me (see note for page 104), in order to avoid the introduction of transfinite numbers. This modification is, for that matter, without importance". We explained above Borel's reason for avoiding the cardinal numbers. Medvedev (1966 p 235) thought that Borel made here the first attempt to remove CBT from the context of the Comparability Theorem for cardinal numbers. Potter (2004 pp 165–6) says that Borel produced a simplified version of Bernstein's proof; perhaps because of this footnote (why 'simplified' rather than 'modified' we do not know). Incidentally, Potter's reference to the note in Borel 1898 pp 103–4 is mistakingly to p 105. We will get back to Borel's footnote in the next section.

obtained by successively removing from A the sets: A-A_1, A_1-A_2, A_2-A_3, . . ., A_n-A_{n+1}, ...[17] We can therefore write

A = D + (A-A_1) + (A_1-A_2) + (A_2-A_3) + (A_3-A_4) + (A_4-A_5)... and each of the symbols between parenthesis designates a determined set, since A_{n+1} is a proper subset of A_n. We can similarly write

A_1 = D + (A_1-A_2) + (A_2-A_3) + (A_3-A_4). . . .

It is now easy to demonstrate that the sets A and A_1 have same power; it is enough to remark that we can regard them as formed of a denumerable infinitude of sets having pairwise same power. It results, in effect, from the nature of the projection[18] by which we obtained the sets A_3, A_4,. . . that A-A_1 has same power as A_2-A_3, as A_4-A_5, as A_6-A_7, ... Similarly, the sets A_1-A_2, A_3-A_4, A_5-A_6, . . . all have same power. It suffices, hence, to write the expression of A_1 in the form

A_1 = D + (A_2-A_3) + (A_1-A_2) + (A_4-A_5) + (A_3-A_4) + (A_5-A_7) + . . ., to recognize that each of the terms of this series has same power as the term of the same place in the series which defines A.[19] The theorem is thus demonstrated.[20]

11.2 Bernstein's Original Proof

Borel's footnote to his definition of the set D suggests that the proof given by him is not exactly the proof given by Bernstein (as presented to Borel by Cantor). The original proof, according to Borel, contained reference to transfinite numbers. In Cantor's set theory, transfinite numbers, cardinal numbers for our context here, are introduced by abstraction. So, Bernstein's original argument probably was that when passing from the frames to their cardinal numbers, A and A_1 are partitioned to partitions with the same cardinal numbers. Thus, they too have the same cardinal number. Whatever differences there are in the frames they disappear after

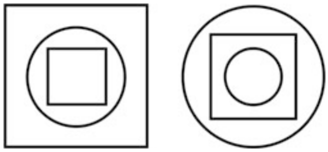
[17] Borel is using the difference operation between sets and + for the union operation. Cantor used (A, B) for the union of A and B (1895 *Beiträge* §3) and used A-B rarely. The notation A + B is attributed by Dedekind to Schröder (see Sect. 9.2.3).

[18] The original reads "*du mode de projection*". This strange expression is explained by a footnote below.

[19] A similar conclusion from the equivalence of the partitioning to the equivalence of the sets appears in Cantor's 1878 *Beitrag* (p 129), which is another evidence that Borel was influenced by this paper. Poincaré would later point out (see Sect. 26.7) that this step makes a special appeal to intuition.

[20] Here (p 106) Borel puts the following footnote: "It is important to point out a grave error that could be suggested by the preceding demonstration. We have seen that A-A_1 and A_2-A_3 have same power, because of the projection by which we have obtained A_3 [which carries the partitioning of A by A_1 and A-A_1 to the partitioning of A_2 by A_3 and A_2-A_3]; but this is not at all a consequence of the fact that A_2 has same power as A and A_3 same power as A_1. It is easy to be assured that, two sets A and B having same power, and their proper subsets A_1 and B_1 having also the same power, nothing can hence be concluded for A-A_1 and B-B_1. It is enough to take, for example, for A the set of points included between 0 and 2, for B the set of points included between 0 and 1, and for A_1, as well as for B_1, the set of points with incommensurable abscissas [the irrationals] included between 0 and 1. The set A-A_1 has now the power of the continuum, while the set B-B_1, is denumerable. Likewise, it could have been supposed, for that matter, that both A_1 and B_1 are identical with B."

Fig. 11.2 Mirrors drawing



abstraction. This metaphor Bernstein could have picked from Cantor who used it to prove the commutativity and associativity of the addition operation of powers (cardinal numbers) in his 1895 *Beiträge* §3.[21] We seem to have here a case of proof-processing. Cantor may have had his high regard for Bernstein's proof because of this use of abstraction as an operator, just as envisaged by Cantor.

Bernstein's argument comes out well from the figure that Borel gave. If we float the image of A_1 and its subsets above that of A, it is easy to gain with one coup the metaphor that the difference between the two sets in their sequence of even frames of A is irrelevant to their powers. Borel, however, uses the figure in an inessential way, when he points out the nesting character of the sequence of sets and that this sequence is in fact composed of two interchanging sequences of equivalent sets. These points are easy enough to understand from Borel's prose and Borel does not refer to the figure to switch the gestalt from the sequence of nesting sets to the sequence of frames. Thus, it seems that the figure was perhaps suggested to Borel by Cantor and it may have originated with Bernstein (being an art student).[22] Borel may have brought the figure to keep a trace of the proof presented to him by Cantor, from which he departed when he partitioned the sets into frames and residue and proved the equivalence of the sets as a result of the equivalence of the partitions.

This observation brings us back to the story of the discovery of Bernstein's proof while shaving. Let us note the following thought experiment suggested by a drawing provided in Abian 1965 p 254. Take a round mirror and a square mirror and place them facing each other. In the square mirror the image of the round mirror will be shown and in the round mirror the image of the square mirror (Fig. 11.2).

Moreover, in the image of the round mirror in the square mirror, the image of the image of the square mirror that is in the round mirror will be shown and likewise in the image of the square mirror in the round mirror the image of the image of the round mirror that is in the square mirror. This sequence of images continues for (countable) ever. Now this is nice but disturbing because even though the two mirrors seem to correlate they do not perfectly match. Or do they? To realize that they do, and establish the proof of CBT, one has only to shift focus and observe that in both mirrors we have two infinite sequences of frames: one sequence of frames is square on the outside and round on the inside and the other round on the outside and

[21] This linkage of Bernstein's proof to the 1895 *Beiträge*, supports the possibility that Bernstein found his proof in 1895 when reading the proofs of 1895 *Beiträge*.

[22] Bernstein used figures also in his 1905 paper, when he discussed situations similar to the proof of CBT (see Sect. 14.2), while Borel used only one other figure in the whole of his 1898 book.

square on the inside. Separating the second type of frames from both A and A_1, it is immediate to realize that the remaining part and the separated parts have the same power after abstraction. Of course, in the general case, the nesting sets can be diffused and are not limited to the plane, but the clue derived from the thought experiment holds. Our thought experiment and the shaving story lead us to believe that Bernstein's proof was for the two-set formulation, with no shift to the single-set formulation as in Borel's rendering of it.

Incidentally, the mirrors drawings leads us to the following ethical version of CBT: In the pupil of my eye there is an image of the whole of you; in the pupil of your eye there is an image of the whole of me; by CBT we are equivalent!

11.3 Comparison with Earlier Proofs

With Bernstein, a gestalt switch occurred, from the gestalt of the nesting sets identified by Schröder, to the gestalt of corresponding frames. Abstraction, in Cantor's sense, is the metaphor of Bernstein's proof,[23] and it differs from Schröder's metaphor of passage to the residue. Bernstein's descriptors differ also from Cantor's descriptors: the gestalt of number-classes and the metaphor of enumerating the subset by the whole set, which underlie his proof. Note that Cantor perhaps interpreted Bernstein's proof on his own terms whereby abstraction can be performed only on consistent sets. Thus, for Cantor, Bernstein's proof had no wider scope than his own *Grundlagen* proof.

With regard to Dedekind, there is some common ground between Bernstein's gestalt of the frames and Dedekind's chain gestalt but the similarity is not spelled out. If we grant Bernstein with separating one type of frames from the sets then Bernstein had obtained Dedekind's partitioning metaphor; however, Bernstein did not use Dedekind's pushdown the chain metaphor.

Bernstein's frames gestalt was adopted by Borel, and many others who later provided proofs of CBT, but the abstraction metaphor, that probably excited Cantor, did not. The idea of defining cardinal numbers by abstraction was rejected by the generation that followed Cantor. This is then an especially nice and important feature of Borel's proof, that it stays in the language of sets and mappings used to formulate the theorem. No appeal is made to extraneous notions such as enumeration-by, or abstraction, invoked in the proofs of Cantor and Bernstein. In this Borel is similar to Dedekind. For this reason the proofs provided by Borel and others were not limited to consistent sets.

Borel's metaphor lies in the rearrangement of the frames in equivalent pairs. It has nothing to do with Cantor's enumeration or Dedekind's pushdown, but it is clearly a translation of the abstraction metaphor to the language of sets and

[23] With regard to the metaphor behind this gestalt switch, it seems to be "passage to the sequence of difference", stemming from the theory of sequences and series of numbers.

mappings. Borel begins with an observation similar to Schröder's *Scheere* but switches to the frames gestalt and then to the single-set formulation. Since Borel does not mention Schröder it seems that Borel had obtained the *Scheere* construction of the nesting sets independently of Schröder, as well as his appreciation of the place of the residue in the decomposition of the sets involved. Perhaps because of their similar point of departure in the *Scheere* construction, their similar appreciation of the residue and their use of drawings, Schoenflies (1900 p 16 footnote 2) remarked that Bernstein's proof (he meant Borel's proof of course) is "essentially identical" with the proof of Schröder.

Let us recall again that in his letter to Dedekind of August 30, 1899, Cantor too said that Bernstein's proof is similar to the proof of Schröder and to that of Dedekind. We leave it to the reader to decide how much inaccurate this judgment is in terms of gestalt and metaphor. Cantor's view teaches us that mathematicians are quick to translate from one set of descriptors to another. Yet we maintain that not every set of alternative descriptors can provoke in a mathematician a solution to a problem he is occupied with. More study is necessary to establish this point of view.

Another metaphor of Borel is the use of complete induction. Schröder used complete induction too in defining his *Scheere*. Dedekind, however, consciously avoided complete induction, using instead an impredicative definition to define the chain of a set as the minimal chain that contains the first frame. Unlike Dedekind, Borel remains silent about the obvious fact that the sequence of odd frames and D (which compose Dedekind's complement to the chain pushed down) belong to the two sets and that therefore no argument on their equivalence is necessary.

An important peculiarity of Borel's proof is that he chooses at each step a new projection; thus he proves that the sets generated at each step are equivalent and then from the many mappings that can establish this equivalence he chooses one. In this way Borel unintentionally invoked the axiom of choice in his proof. Dedekind uses only the projection from A to A_2, which even if not assumed as given, it could be chosen from among the mappings that establish this equivalence, using only a single choice. (See in this regard Harward's criticism of Jourdain in Sect. 17.5.) Borel may have been aware that he is making infinitely many arbitrary choices and he may have been aware of Peano's 1890 paper where infinitely many choices were rejected. Of course, the axiom of choice was not yet stated in 1898 so Borel was not aware that he is employing it. After the axiom of choice was explicitly stated by Zermelo in 1904, Borel became one of its main antagonists.

Chapter 12
Schoenflies' 1900 Proof of CBT

In 1898, the German Association of Mathematicians (DMV) commissioned Schoenflies to write a report on "Curves and point manifolds". The first part of the report was published in 1900 under the title "The development of the science of point manifolds". The report was updated in 1908 and 1913 but the subject grew by that time out of the scope of any single report. In 1900, however, the report was the only text book on set theory and its applications, outside Cantor's 1895/7 *Beiträge*.

The report contained a proof of CBT (p 16ff). Because the report was widely read, its CBT proof became well known and was often repeated. Schoenflies proof which amalgamated the two proofs known at the time: the faulty proof of Schröder (1898) with its two-set formulation, and that of Borel (1898), attributed to Bernstein, which was rigorous and was conducted for the single-set formulation. However, since apparently Borel's 1898 book, where the proof of Bernstein was first published, was not available in Göttingen by 1899, Schoenflies must have asked Cantor to supply him the proof of Bernstein. Cantor did that in the attachment to his letter to Hilbert of June 28, 1899. The proof does not use the abstraction metaphor that we attributed to Bernstein's original proof. Instead it is a simplified version of Borel's proof with different notation. We will first bring Cantor's proof and then that of Schoenflies.

12.1 Cantor's 1899 Proof

We cite Cantor's proof from Meschkowski-Nilson 1991 p 401f[1]:

[1] Cantor notes that Schröder was first to have found a proof and that then Bernstein had independently found a proof which is independent of logical calculations.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_12, © Springer Basel 2013

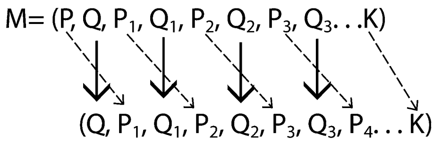


Fig. 12.1 Cantor's CBT drawing

We conduct easily the proof of Theorem U[2] from the following:

Given three sets M, N, S such that N is a subset of M, S a subset of N and it is given that M ~ S, then also M ~ N.

Proof: M is composed of N and P, N of S and Q. So we have M = (P,N),[3] N = (Q,S) M = (P,Q,S). Since M ~ S there is a 1–1 projection[4] from M onto S, by this projection the subsets P, Q, S of M correspond to the subsets P_1, Q_1, S_1 of S so that S = (P_1,Q_1,S_1) and P ~ P_1, Q ~ Q_1, S ~ S_1. By the same projection we have S_1 = (P_2,Q_2,S_2) and P_1 ~ P_2, Q_1 ~ Q_2, S_1 ~ S_2, and so on.

We thus have M = (P,Q,P_1,Q_1,P_2,Q_2, . . ., K). Here K is that part of N that by the projection of M onto S, is projected upon itself[5]; K can be empty. We have in addition:

N = (Q,P_1,Q_1,P_2,Q_2, . . ., K). Now we see that M and N can be 1–1 projected as follows: P of M onto P_1 of N, P_1 of M onto P_2 of N, P_2 of M onto P_3 of N, . . . Q of M onto Q of N, Q_1 of M onto Q_1 of N, Q_2 of M onto Q_2 of N, . . . K of M onto K of N. For greater clearness I represent the projection by the following way: (Fig. 12.1)

So we have M ~ N.

12.2 Schoenflies' Proof

If the sets M and N are in the relation that a subset of M is equivalent to N and a subset of N equivalent to M, then also M ~ N.[6]

Let P_1 be the subset of M equivalent to N and M_1 the residue set, so that (1) M = (M_1, P_1)[7] and analogously (2) N = (N_1, Q_1), so that P_1 ~ N and Q_1 ~ M. Because P_1 ~ N there is a 1–1 mapping from N onto P_1; thereby the subsets N_1 and Q_1 of N correspond to the subsets M_2 and P_2 of P_1, so that (3) P_1 = (M_2, P_2), M_2 ~ N_1 P_2 ~ Q_1 and

[2] Theorem U is CBT in its two-set formulation.

[3] Cantor is using his notation from 1895 *Beiträge* for union.

[4] We translate Cantor's "*eindeutige*" into 1–1. Cantor is using "projection" as did Borel.

[5] Cantor defines K by the property required for the proof rather than as the intersection of nesting sets or the residue after removal of frames.

[6] Schoenflies is using Cantor's sign '~' for equivalence which he defines (p 5) as Cantor by: "two sets M and N are called equivalent or of equal power (M ~ N), in case it is possible, according to some law, to put them in one-one correlation."

[7] Schoenflies follows Cantor's notation for union (§3 of Cantor's 1895 Beiträge). Note that Schoenflies notation for the sets generated in the proof is idiosyncratic for he uses signs derived from M, N for the frames and not for the nested sets, which are equivalent to M, N. In Schoenflies 1913 (p 34f) the notation was set straight.

(4) $M = (M_1, P_1) = (M_1, M_2, P_2)$. From $P_2 \sim Q_1$ and $Q_1 \sim M$ it follows additionally $P_2 \sim M$; we need then by (1) to set $P_2 = (M_3, P_3)$ for $M_3 \sim M_1$, $P_3 \sim P_1$. From here follows in addition $M = (M_1, M_2, M_3, P_3)$.

So we can continue. Now in addition it is to be noted that P_2 is a subset of P_1 and P_3 of P_2 and so on; we arrive thence finally to the representation (5) $M = (M_1, M_2, M_3, \ldots, P_\omega)$, where the set P_ω, in case such a set exists at all, contains those elements that are common to all the sets $P_1, P_2, P_3, \ldots$.

In the same manner we attain a partitioning of N, in which we begin from the equivalence $Q_1 \sim M$. By the mapping from M onto Q_1 the subsets M_1 and P_1 of M correspond to the subsets N_2 and Q_2 of Q_1, so that $Q_1 = (N_2, Q_2)$, $N_2 \sim M_1$, $Q_2 \sim P_1$. Thereupon it follows in addition that $N = (N_1, Q_1) = (N_1, N_2, Q_2)$.
Analogously we obtain out of $Q_2 \sim P_1 \sim N$, as above,
(6) $N = (N_1, N_2, N_3, Q_3)$ and finally
(7) $N = (N_1, N_2, N_3, \ldots, Q_\omega)$, where again Q_ω, in case it exists, contains the elements common to all $Q_1, Q_2, Q_3, \ldots$. But now it follows directly out of our designations that
$M_1 \sim N_2 \sim M_3 \sim N_4 \ldots$, $N_1 \sim M_2 \sim N_3 \sim M_4 \ldots$, and
$Q_1 \sim P_2 \sim Q_3 \sim P_4 \ldots$, $P_1 \sim Q_2 \sim P_3 \sim Q_4 \ldots$; According to the end of Chapter 2 [of Schoenflies' report] it is also that P_ω, the common set to all $P_1, P_2, P_3, \ldots$, is equivalent to Q_ω, contained in all $Q_1, Q_2, Q_3, \ldots$, and out of which it finally follows that (8) $M \sim N$.

The proof rests first upon the fact that whatever the powers of the sets M and N may be, the proof procedure still gets to its aim in a denumerable set of steps; it rests, secondly, on the possibility of the set M_ω established in p 14 [which is the end of Chapter 2]. This set provides one of the most important tools in set theory.

The reference to page 14, is to the unproved lemma: If $M_0 \sim N_0$ and M_i and N_i are two sequences of subsets of M_0, N_0, such that always M_i, N_i consist of corresponding [under the equivalence of M_0 and N_0] elements, then denoting by M_ω (N_ω) the intersections of M_i (N_i), if M_ω exists (is not empty) so does N_ω exist and $M_\omega \sim N_\omega$. Note that the lemma does not require the subsets to be nesting. The proof of the lemma employs complete induction, as Schoenflies hints when he talks about the "denumerable set of steps". The lemma provides the generation and equivalence of the residues P_ω, Q_ω. The final statement of the quoted passage emphasizes the rigorousity of the passage from the finite to ω. Perhaps Schoenflies hinted here at Schröder who obtained this correct conclusion by an erroneous argument, see Sect. 10.2.

12.3 Comparisons

Schoenflies used Cantor's notation in his letter to Schoenflies (see Sect. 12.1) and his convention to denote the frames directly and not as differences as was used by Borel. However, the letters which Cantor used to denote the frames, Schoenflies used to denote the nesting sets.

The gestalt of Schoenflies' proof is Schröder's *Scheere* gestalt but between the frames instead of the nesting sets as in Schröder gestalt. The metaphor is taken from Borel's proof, mediated through Cantor: the frames are pairwise equivalent but in this case, because the proof is for the two-set formulation, the metaphor is not that of rearranging the frames but rather reminiscent of finger-interlacing. This metaphor is complemented with the lemma about the residues which may correspond to the circle formed between the thumbs and fingers during finger interlacing. Like both Schröder and Cantor's rendering of Borel, Schoenflies uses complete induction to define the nesting sets and the frames. So we see that Schoenflies indeed merged the proofs of Schröder and Borel as provided to him by Cantor. Clearly, the proof of Schoenflies differs from Cantor's original proof in every aspect. Because of its resemblance to Borel's proof we will omit Cantor's 1899 proof from following comparisons.

Chapter 13
Zermelo's 1901 Proof of CBT

Zermelo published two proofs of CBT. The first in 1901, in his first paper on set theory, we review in detail in this chapter. The second in 1908, in the paper where Zermelo first presented his axiomatic set theory, we review in Chap. 23.

Both proofs proof-processed from Dedekind's chain theory from *Zahlen*. Still, when in Cantor's collected works, edited by him, Zermelo commented (Cantor 1932 p 451 [2]) on Dedekind's CBT proof (see Sect. 9.2.3), he noted its similarity only to his 1908 proof. The reason is perhaps that in the later proof Zermelo follows Dedekind more closely. In the 1901 proof an attempt is made to carry the discussion as much as possible in the language of cardinal numbers, taken from Cantor's 1895 *Beiträge*, instead of Dedekind's language of sets and mappings. Zermelo even differentiated, in the first sentence to the paper, between finite cardinal numbers and positive integers, presumably because the two are differently defined: the latter by induction and the first by Cantor's abstraction.

From this regard it may be argued that Zermelo 1901 demonstrates the limitations of Cantor's language of cardinal numbers: it is always necessary at certain points of a presentation of the theory of cardinal numbers to invoke arguments in the language of sets and mappings. Here it happens in the proof of the first part of Theorem I.[1] In Lakatosian terms (1976 p 125), Zermelo attempted in 1901 a translation of the dominant language of set theory from the language of sets and mappings to the language of cardinal numbers. This translation, however, did not enrich the discussion, as such a translation ought to according to Lakatos, and so it was deserted. It seems that for this reason Zermelo avoided the notion of cardinal number in his axiomatic set theory of 1908, where he presented CBT in the language of sets and mappings. In this Zermelo joined Dedekind who entirely ignored Cantor's transfinite numbers.

Zermelo's 1901 formulation of CBT is set in plain text, between theorems III and IV of the paper, without the title 'Theorem x', where x is a Roman numeral, and

[1] For the proof of this theorem Zermelo proof-processed certain elements from Schoenflies' 1901 rendering of Cantor's 1878 *Beitrag*; see below.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_13, © Springer Basel 2013

the spread letters which characterize all other statements of theorems of the paper, so as to have them stand-out. The reason is perhaps that the other theorems were original, as Zermelo points out (p 38), while CBT was proved by Schröder and Bernstein (p 37). Zermelo dates those proofs to 1896, which is correct for Schröder's proof but perhaps not for Bernstein's, who probably found his proof around Easter 1897 (see Chap. 11). Interestingly, Zermelo references only Borel 1898 (where Bernstein's proof was supposedly published) and Schoenflies 1900, not Schröder's 1898; perhaps this indicates criticism of the latter's proof.

Perhaps the crux of this chapter lies in our attempt to associate the convex-concave gestalt switch, which we identified in Zermelo's proof of CBT, with drawings of Escher and Vasarely, and with two poems of Emily Dickinson. We thus demonstrate how metaphoric descriptors of proofs can place mathematics inside general culture.

13.1 The Proof

Zermelo stated CBT as follows:

> When each of two sets is equivalent to a part of the other then the two sets are equivalent.

The theorem is stated without any symbols, as are its two bordering Theorems III and IV. We roll Zermelo's proof backwards, to its lemmas, contrary to the Euclidean order taken in Zermelo's paper:

> Is namely a) $\mathfrak{m} = \mathfrak{n} + \mathfrak{p}$ and b) $\mathfrak{n} = \mathfrak{m} + \mathfrak{q}$ it is also $\mathfrak{m} = \mathfrak{m} + \mathfrak{p} + \mathfrak{q}$ and therefore $\mathfrak{m} = \mathfrak{m} + \mathfrak{q} = \mathfrak{n}$.

Obviously Zermelo denotes by $\mathfrak{m}$ and $\mathfrak{n}$ the cardinals[2] of the two sets, by $\mathfrak{p}$ the cardinal of the difference between the first set and the image in it of the second, and by $\mathfrak{q}$ the cardinal of the difference between the second set and the image in it of the first. The third equality is obtained by substituting $\mathfrak{n}$ in the first equation by its value in the second and by the associativity and commutativity of cardinal addition, which Zermelo uses without mention. With this substitution Zermelo switches from the two-set formulation in which the theorem is presented to prove the theorem in its single-set formulation.

Following the proof of CBT, Zermelo notes that if (a) holds but (b) does not hold, then according to Cantor's 1895 *Beiträge* definition of the order relation between cardinal numbers (see Sect. 4.3), we have $\mathfrak{m} > \mathfrak{n}$. Hence Zermelo gave another, never before mentioned, equivalent formulation to CBT, in his Theorem IV: "The sum of two or more cardinal numbers is always greater or equal anyone of

---

[2] Sometimes we use 'cardinal' for 'cardinal number'; Zermelo sometimes uses 'number' for 'cardinal number'. Like Cantor in his 1895 *Beiträge*, Zermelo is using small Old English letters for cardinal numbers.

them". Cantor never used 'greater or equal' for cardinal numbers. Jourdain would later stress the point that $\geq$ rests on CBT (see Sect. 17.2). In fact, if (a) holds, we know that N (the set of cardinal number $\mathfrak{n}$) is equivalent to a subset of M (the set of cardinal number $\mathfrak{m}$). If also M is equivalent to a subset of N we have by CBT that the sets are equivalent and so $\mathfrak{m} = \mathfrak{n}$ according to Cantor's definition. Since we do not know if (b) holds or not we have $\mathfrak{m} \geq \mathfrak{n}$. On the other hand, under the conditions of CBT we have by Theorem IV that $\mathfrak{m} \geq \mathfrak{n}$ and $\mathfrak{n} \geq \mathfrak{m}$. But since under Cantor's definition of $>$, the two cases $\mathfrak{m}>\mathfrak{n}$ and $\mathfrak{n}>\mathfrak{m}$ are exclusive of each other, we obtain $\mathfrak{m} = \mathfrak{n}$, which is CBT. Hence the equivalence of CBT and Theorem IV.

The fourth equality in the proof of CBT is obtained by Theorem III of the paper that immediately precedes the statement of CBT:

> Theorem III. When the sum of two cardinal numbers added to a third leaves it unchanged, then the same happens for each of its summands.
> Is namely $\mathfrak{m} = \mathfrak{m} + \mathfrak{p} + \mathfrak{q}$ so is it also by Theorem II
> $\mathfrak{m} = \mathfrak{m} + \mathfrak{a}(\mathfrak{p} + \mathfrak{q})$[3]
> $= \mathfrak{m} + (\mathfrak{a} + 1)\mathfrak{p} + \mathfrak{a}\mathfrak{q} \qquad = \mathfrak{m} + \mathfrak{a}\mathfrak{p} + (\mathfrak{a} + 1)\mathfrak{q}$
> $= \mathfrak{m} + \mathfrak{p} \qquad = \mathfrak{m} + \mathfrak{q}$,[4]
> and naturally this equally applies also when the sum is composed of more than two summands.

Before Zermelo proved CBT from Theorem III, he said that the two have "the same kernel". This figurative language seems to draw our attention to the fact that not only is CBT implied by Theorem III but it also implies it. We realize this when we note that $M \sim M''$ implies not only that $M \sim M'' + (M'\text{-}M'')$ but also $M \sim M'' + (M\text{-}M')$ by a symmetrical argument. This point is not noticed in earlier discussions of CBT.

Theorem II, which Zermelo used in the proof of Theorem III, is:

> When a cardinal number $\mathfrak{m}$ remains unchanged by the addition of another $\mathfrak{p}$, it remains unchanged by infinitely many repeated additions of the same cardinal number, or by its addition $\mathfrak{a}$ times.

Zermelo proved Theorem II by using the first part of his Theorem I, while the second part of Theorem I is proved using Theorem II. Theorem I, the Denumerable Addition Theorem, is the following:

> Theorem I (main theorem). When a cardinal number $\mathfrak{m}$ remains unchanged by the addition of any of the cardinal numbers of the infinite sequence $\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \ldots$, it remains unchanged when all are added at once.[5]

[3] Zermelo denotes $\aleph_0$ by $\mathfrak{a}$.

[4] The two columns typographic arrangement leads us to read that the third (forth) expression implies the fifth (sixth).

[5] Zermelo remarked (p 38 towards the end of the paper) that Theorem I is an extension of CBT, which cannot be proved from it alone. He may have meant that to prove $(M, P_1, P_2, \ldots) \sim M$ by CBT it is necessary to prove that $(M, P_1, P_2, \ldots)$ is equivalent to a subset of M. Theorem I requires less, that each $(M, P_\lambda) \sim M$, and so it extends CBT.

Here is the proof of the first part of Theorem I, which proves Theorem II:

By our assumptions (1) $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}_1 = \mathfrak{m} + \mathfrak{p}_2 = \mathfrak{m} + \mathfrak{p}_3 = \ldots$ so a set M of power $\mathfrak{m}$ can be partitioned in the following way
(2) $M = (P_1, M_1)$[6] $= (P_2, M_2) = (P_3, M_3) = \ldots$,
where every subset P_λ, of the cardinal number $\mathfrak{p}_\lambda$, has no common element with its complementary M_λ and all M_λ are equivalent with the set M itself:
$M \sim M_1 \sim M_2 \sim M_3 \ldots$.
So there exist 1–1 mappings $\varphi_1, \varphi_2, \varphi_3, \ldots$[7] from the set M on its subsets $M_1, M_2, M_3, \ldots$, so that it can be written
(3) $M_1 = \varphi_1 M, M_2 = \varphi_2 M, M_3 = \varphi_3 M, \ldots$. For an arbitrary such mapping φ it is also always that
(2a) $\varphi M = (\varphi P_1, \varphi M_1) = (\varphi P_2, \varphi M_2) = (\varphi P_3, \varphi M_3)$[8]$\ldots$,
that is, for all equivalent sets hold the analog partitioning.
Therefore successively also

$$\begin{array}{lll} M & = (P_1, M_1) & \\ M_1 = \varphi_1 M & = (\varphi_1 P_2, \varphi_1 M_2) & = (P'_2, M'_2) \\ M'_2 = \varphi_1\varphi_2 M^{9} & = (\varphi_1\varphi_2 P_3, \varphi_1\varphi_2 M_3) & = (P'_3, M'_3) \\ M'_3 = \varphi_1\varphi_2\varphi_3 M & = (\varphi_1\varphi_2\varphi_3 P_4, \varphi_1\varphi_2\varphi_3 M_4) & = (P'_4, M'_4) \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ M'_{\lambda-1} = \varphi_1\varphi_2\ldots\varphi_{\lambda-1} M & = (\varphi_1\varphi_2\ldots\varphi_{\lambda-1} P_\lambda, \varphi_1\varphi_2\ldots\varphi_{\lambda-1} M_\lambda) & = (P'_\lambda, M'_\lambda) \end{array}$$

Therefore (4) $M = (P_1, P'_2, P'_3, \ldots P'_\lambda; M'_\lambda)$,[10] when the mappings $\varphi_1, \varphi_2, \varphi_3, \ldots$ are executed one after the other and use is made of the abbreviations:

$$\varphi_1\varphi_2\ldots\varphi_{\lambda-1} P_\lambda = P'_\lambda \sim P_\lambda$$
$$\varphi_1\varphi_2\ldots\varphi_{\lambda-1} M_\lambda = \varphi_1\varphi_2\ldots\varphi_\lambda M = M'_\lambda \sim M$$

Also here are each two of the subsets P'_λ of the original set M always without any common element, and it is finally.

(5) $M = (P_1, P'_2, P'_3, \ldots; M')$, that is M contains all subsets P'_λ and therefore their union set. For the corresponding cardinal numbers we have first (6) $\mathfrak{m} = \mathfrak{p}_1 + \mathfrak{p}_2 + \mathfrak{p}_3 + \ldots + \mathfrak{m}'$, and it is left only to show that here the cardinal number $\mathfrak{m}'$ can be substituted by $\mathfrak{m}$ itself.[11]

---

[6] Like Schoenflies, Zermelo uses Cantor's notation (M, N) for the union of M, N.

[7] The axiom of choice is necessary to provide this sequence of mappings from assumption (1). Moore (1982 p 90) further notes that already for the definition of infinite sum of cardinals the axiom of choice is necessary. He refers probably to the choice of the P_λ.

[8] Zermelo has here ',' which is clearly a typo.

[9] With Zermelo the rightmost mapping is applied first, unlike Bernstein 1905 (see the next chapter).

[10] The use of ';' in this notation seems to be a variant of Zermelo.

[11] If all the P_λ are disjoint from the start, we can skip the construction of the P'_λ and begin the proof at (6). The construction can be simplified (eliminating the need to define the P'_λ), if we apply $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}_2$ to M_1 directly, and so on. Note that the P_λ need not be different.

But first we apply the formula (6) to the case where $\mathfrak{p}_1 = \mathfrak{p}_2 = \mathfrak{p}_3 \ldots = \mathfrak{p}$, so that $\mathfrak{p}_1 + \mathfrak{p}_2 + \mathfrak{p}_3 + \ldots = \mathfrak{a}\mathfrak{p}$ and then it follows from the single equality $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}$ immediately
(7) $\mathfrak{m} = \mathfrak{a}\mathfrak{p} + \mathfrak{m}' = 2\mathfrak{a}\mathfrak{p} + \mathfrak{m}' = \mathfrak{m} + \mathfrak{a}\mathfrak{p}$.[12]

Hence Zermelo obtained Theorem II. Note, however, that $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}$ does not imply (7); it is (6) that implies (7). That $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}$ implies (6) follows from the first part of the proof of Theorem I as the P_λ are not required to be different! But then, as Zermelo noted (p 38), a direct proof of Theorem II emerges by taking the images $\varphi^\lambda P$ as the disjoint copies of P. Under the direct proof, Theorem II makes no appeal to the axiom of choice. Indeed it seems that instead of bundling Theorem II within Theorem I, Zermelo's presentation would have profited from establishing Theorem II directly and placing the first part of Theorem I in a lemma: If a set M can be partitioned in denumerably many ways $M = (P_1, M_1) = (P_2, M_2) = (P_3, M_3) = \ldots$, where all the M_λ are equivalent with the set M itself, then M contains disjoint copies of the P_λ. Zermelo's presentation evidences perhaps his flowing rather than formal style.

The union of all the copies of P is Dedekind's chain of P (see Sect. 9.1) and the copies of P are the frames of this chain. No doubt for this reason Zermelo called Theorem II the 'Chain Theorem'. Zermelo's application of Dedekind's chain gestalt in the proof of the first part of Theorem I is an example of proof-processing. The pushdown metaphor, however, is not applied.

Zermelo continued with the proof of Theorem I as follows:

Thereby we can widen our condition (1) in the following way:
(1a) $\mathfrak{m} = \mathfrak{m} + \mathfrak{a}\mathfrak{p}_1 = \mathfrak{m} + \mathfrak{a}\mathfrak{p}_2 = \mathfrak{m} + \mathfrak{a}\mathfrak{p}_3 \ldots$ and we have by (6):
(8) $\mathfrak{m} = \mathfrak{a}\mathfrak{p}_1 + \mathfrak{a}\mathfrak{p}_2 + \ldots + \mathfrak{m}''$

$$= 2\mathfrak{a}(\mathfrak{p}_1 + \mathfrak{p}_2 + \mathfrak{p}_3 + \ldots) + \mathfrak{m}'' \qquad = (\mathfrak{a} + 1)(\mathfrak{p}_1 + \mathfrak{p}_2 + \mathfrak{p}_3 + \ldots) + \mathfrak{m}''$$ [13]

$$= \mathfrak{m} + \mathfrak{a}(\mathfrak{p}_1 + \mathfrak{p}_2 + \mathfrak{p}_3 + \ldots) \qquad = \mathfrak{m} + \mathfrak{p}_1 + \mathfrak{p}_2 + \mathfrak{p}_3 + \ldots$$

whereby Theorem I, even in extended form,[14] is demonstrated.

Zermelo introduced the background for his Theorem I at the beginning of his paper through the following considerations. The sum of two finite cardinal numbers is always greater than each of the summands while the same is not true for infinite cardinal numbers: because an infinite set is always equivalent to one of its proper subsets every cardinal number is the sum of it and another cardinal number.[15] Zermelo then brings the equalities regarding $\mathfrak{a}$ as examples of this general property

[12] Zermelo assumes here properties of $\mathfrak{a}$ and the distributivity of cardinal multiplication over finite sum, besides the commutativity and associativity of cardinal addition.

[13] Zermelo uses here distributivity over denumerable sum.

[14] The extended form is unnecessarily emphasized; it results by theorem II applied on the conclusion of theorem I.

[15] Zermelo references Cantor 1878 *Beitrag* with regard to the mentioned property of infinite sets, pointing out that Dedekind used it as a definition of infinitude. He does not stress that Dedekind had devised his definition independently of Cantor (see Sect. 8.1).

of infinite sets: $\mathfrak{a} = \mathfrak{a} + 1 = \mathfrak{a} + \mathfrak{a} = 2\mathfrak{a}$. Then Zermelo adds that because every infinite set has a denumerable subset[16] the following sequence of equalities holds: $\mathfrak{m} = \mathfrak{m}' + \mathfrak{a} = \mathfrak{m}' + 2\mathfrak{a} = \mathfrak{m} + \mathfrak{a}$ and $\mathfrak{m} = \mathfrak{m}' + \mathfrak{a} = \mathfrak{m}' + (\mathfrak{a} + 1) = \mathfrak{m} + 1$. Thus 1 and $\mathfrak{a}$ are in the set of all cardinal numbers that do not change an infinite cardinal number $\mathfrak{m}$ under addition. Let us denote this set by $C_{\mathfrak{m}}$. Zermelo then points out that $C_{\mathfrak{m}}$ is closed under finite addition, which is easy to verify using substitution, complete induction and the associativity of cardinal numbers addition. Theorem I comes to show that $C_{\mathfrak{m}}$ is closed also under denumerable addition. Towards the end of the paper (p 38) Zermelo also shows, based on his Theorem III, that $C_{\mathfrak{m}}$ is closed under diminution (*Verkleinerung*), that is: if $\mathfrak{p}' < \mathfrak{p}$ and $\mathfrak{p} \in C_{\mathfrak{m}}$ then $p' \in C_{\mathfrak{m}}$, and that if $\mathfrak{m} < \mathfrak{m}'$ then $C_{\mathfrak{m}} \subseteq C_{\mathfrak{m}'}$. It is perhaps because of the diminution property that follows from Theorem III that Zermelo characterized it (p 37) as the reciprocal of Theorem I. Having presented these closure properties of $C_{\mathfrak{m}}$, Zermelo says that it forms "a certain 'group' of numbers". The attribute 'group' reveals the budding structuralistic attitude of the period.

It is interesting to note that Zermelo did not raise the question whether $\mathfrak{m}$ belongs to $C_{\mathfrak{m}}$. Most likely it was not because the question did not occur to him. We believe that the reason was that Zermelo, already in 1901, and like Russell in 1902 (see Chap. 15, Grattan-Guinness 1977 p 80), realized that a new postulate was necessary to establish this result, namely, he must have been on the tracks of the axiom of choice. The common view, let us note, is that Russell discovered the multiplicative axiom in the summer of 1904, "slightly before the conception and writing of Zermelo's [1904] paper" (Grattan-Guinness 2000 p 341). But Russell was on the tracks of the axiom already in February 1901 (Grattan-Guinness 2000 §6.5.2) and Zermelo lectured on the content of his 1901 paper in March 1901. So it appears that both Zermelo and Russell sensed the coming of AC at the same time! And it was 3 years before they formulated it.

13.2 The Reemergence Argument

In the proofs of Theorems I, II, III, Zermelo leveraged on the pre-Cantorian paradoxes of the infinite, namely, the equalities $\mathfrak{a} = \mathfrak{a} + 1$ and $\mathfrak{a} = \mathfrak{a} + \mathfrak{a} = 2\mathfrak{a}$, which he applied to products with other cardinal numbers. We name this Zermelian proof metaphor the 'reemergence argument', because of what happens when it is used in (7): $\mathfrak{m}$ seems to be reduced to a smaller $\mathfrak{m}'$ after partitioning out $\mathfrak{a}\mathfrak{p}$ only to be found that $\mathfrak{m}'$ can in fact be taken to be $\mathfrak{m}$. To us, the reemergence argument has captivating mystical beauty; we think that it wonderfully captures Cantor's innovative gestalt of infinite sets.

[16] Zermelo was yet unaware that the axiom of choice is necessary for this assertion, which is equivalent to the reflexivity of an infinite set mentioned earlier.

In 1901 the reemergence argument was used by both Zermelo and Bernstein (in his doctorate thesis published as Bernstein 1905, see the next chapter). Zermelo introduced the properties of $\mathfrak{a}$ while Bernstein simply assumed them tacitly.

Zermelo linked the reemergence argument with Dedekind's chain theory, when in Theorem II he (tacitly) saw in $\mathfrak{ap}$ the chain generated by P. But the reemergence argument can be traced back to Cantor's proof, in his 1878 *Beitrag* paper, that the irrationals are of the power of the continuum (though there the discussion is held in the language of sets and mappings). It is in Schoenflies (1900) presentation (p 21f) of that result (in a footnote Schoenflies indeed references Cantor's 1878 *Beitrag*) that the proof was transferred to the language of cardinal numbers:

> The set of all irrational numbers [...] in a given interval δ has the power $\mathfrak{c}$ [of the continuum].
>
> Let first J be the set of irrational numbers, and $\mathfrak{i}$ its power, then, the rational numbers of the interval having power $\mathfrak{a}$ [$\aleph_0$], $\mathfrak{c} = \mathfrak{i} + \mathfrak{a}$. Partition now J by separating a denumerable set A_1 into the sets A_1 and J_1, so that $\mathfrak{i} = \mathfrak{i}_1 + \mathfrak{a}$, and hereby it follows by the addition of $\mathfrak{a}$ on both sides $\mathfrak{i} + \mathfrak{a} = \mathfrak{i}_1 + \mathfrak{a} + \mathfrak{a}$ or, as $\mathfrak{i} + \mathfrak{a} = \mathfrak{c}$ and $\mathfrak{a} + \mathfrak{a} = \mathfrak{a}$, that $\mathfrak{c} = \mathfrak{i}_1 + \mathfrak{a} = \mathfrak{i}$.

It is perhaps to the reemergence argument that Schoenflies addressed his comment on the residue (see the end of Sect. 12.1). Schoenflies provided also a proof of the theorem in the language of sets and mappings instead of the "calculative [*rechnerische*]" proof quoted. We borrow from Schoenflies and assign to Zermelo's 1901 CBT proof the metaphor 'calculative'.

Affinity exists also between Zermelo's proof of the first part of Theorem I and the typography of Schoenflies' proof of CBT, except that Zermelo's notation is more reasonable. We see that there are proof-processing links from Zermelo 1901 to both Schoenflies 1900 and Dedekind's *Zahlen*.

Zermelo himself remarked (p 38) that with the direct proof of Theorem II, his proof of CBT is only different from Bernstein's proof (Zermelo surely means Borel's proof) in "arrangement and phrasing". This statement, which reveals Zermelo's essentialism (see Lakatos 1976 pp 53, 122),[17] cannot be upheld. Borel avoided explicitly any mention of cardinal numbers and he was not using the reemergence argument but the pairwise equivalence of the rearranged frames metaphor.

13.3 Convex-Concave

Following his proof on the power of the irrationals Schoenflies remarked:

> The above proof shows in addition, that the power of C [the continuum] is not changed in case a denumerable set is deleted from it or appended to it, that is $\mathfrak{c} + \mathfrak{a} = \mathfrak{c}$ and $\mathfrak{c} - \mathfrak{a} = \mathfrak{c}$.

---

[17] On Zermelo's philosophical position see Moore 1982 p 146ff (Moore's quotation on p 148 which reveals Zermelo's essentialism, appears in van Heijenoort 1967 p 194).

And he adds in a footnote that this assertion applies to any set the power of which is greater than $\mathfrak{a}$.

This remark led us to realize that while in the above discussion of Zermelo's proof the equation $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}$ was taken to mean that there is a set M and a subset P such that M is of cardinal $\mathfrak{m}$ and P of cardinal $\mathfrak{p}$ and M-P is of cardinal $\mathfrak{m}$, we could alternatively, represent the situation by saying that there are two disjoint sets M, P, with the same denotation of the cardinals as above, and such that M + P is of cardinal $\mathfrak{m}$. This equivalence of M + P and M-P reminds us of famous convex-concave *trompe d'oeil*, which is demonstrated in such drawings as the "Convex and Concave" drawing of Escher[18] and the "Expansive-Regressive" of Vasarely.[19] The staircases of Escher show that the same ambiguity is possible for any finite number of such P; the picture of Vasarely suggests how the same applies to chains and chain of chains. Note, however, that the convex-concave analogy is not a full analogy, because the tromp d'oeil does not apply to the flat body itself, before it receives its concave or convex bent. In the set context we have also that M is equivalent to its convex and concave variations.

From Professor Nancy Mayer,[20] with whom I discussed the convex-concave aspects of infinity, I received the following poem by Emily Dickinson (Franklin 1999 # 830):

The Admirations – and Contempts – of time –
Show justest – through an Open Tomb –
The Dying – as it were a Hight[21]
Reorganizes Estimate
And what We saw not
We distinguish clear –
And mostly – see not
What We saw before –
'Tis Compound Vision –
Light – enabling Light –
The Finite – furnished
With the Infinite –
Convex – and Concave Witness –
Back – toward Time –
And forward –
Toward the God of Him –

The poem is surely interesting in our context because of its mention of finite and infinite and of convex and concave. On the edge of mortality, time is convex and finite while God is convex and infinite. Death reorganizes our appreciation:

[18] http://en.wikipedia.org/wiki/Convex_and_Concave.

[19] http://arslonga.dk/BUDAPEST/3-vv15.jpg

[20] Of the English department at Northwest Missouri State University.

[21] Literally 'hight' means 'named'. Professor Mayer tends to think there were words Emily Dickinson simply misspelled, so that we should read 'height'. But then, we can read "a Hight" as making the dying a kind of standard, a measuring rod.

now our view of earthly things becomes blurred while our view by divine light becomes clear.

Another related poem of Emily Dickinson where the ‘finite’ and the ‘infinite’ are invoked in the explicit context of mathematics appears as the motto for Tymoczko 1986. It bears the dedication:

For Alice Dickinson
Mathematician, Teacher, Ringer of Change[22]

And the poem runs as follows:

There is a solitude of space
A solitude of sea
A solitude of Death, but these
Society shall be
Compared with that profounder site
That polar privacy[23]
A soul admitted to itself –
Finite infinity.

The second poem, taken as given, we interpret thus: ‘Finite’ is our measure – us, humans, in the real world; ‘infinite’ is our inner world of imagery, which we encounter by reflection. Just as the infinite riches of Dedekind’s structures of chains are generated by the reflection of an infinite set into itself, our inner reflection generates infinite diversity within the finite space that we occupy. The pole to society is individual creativity. Other forms of solitude face the individual with magnitudes of far greater measure: the sea; space; the abyss of death. In all these situations anxiety is triggered. Not so in reflection. “But” in the third row should be read as ‘except’; except these solitudes of anxiety, against which society compares favorably, contemplation is another pole, against which society can hardly withstand comparison.

Back to the first poem: The finite, we have seen, is us, mortals – the Dying; we are furnished with infinity – our ability for introspection which generates descriptions – the admirations (and contempts) of time which are the convex (concave) judgments of time, and they are most just – accurate – when expressed with regard to the open tomb, at the very edge of personal extinction when we estimate time as if it is the height of an open tomb (the depth is presented as height – the convex as concave). Our cognition is characterized by our ability to understand the new and overlook the known, otherwise we would not be able to switch between convex and concave vision. Our ability for poetic conceptualization is the light that enables our epistemic vision. Upon the open tomb we realize that back is the

[22] According to Professor Mayer, the dedication does not appear in Dickinson’s books and the identity of Alice Dickinson is unknown. Also the last line is considered an editorial addition. Check the web for ‘ringer of change’.

[23] Check out Winnicott 1958.

(convex) region of time and forward is the (concave) region of God. Naturally, feminine-masculine undertones (Tomb – Womb – Him) resonate in the finale.

Against the above artistic background we attribute to Zermelo's proof of the Denumerable Addition Theorem, from which he derived his calculative proof of CBT, the additional gestalt 'convex-concave', to expand on the chain gestalt upon which we remarked earlier. Zermelo saw the partitioning of the set into a chain and a residue that is the gestalt of Dedekind's CBT proof; he then multiplied the chain because he saw in the convex residue the original set concaved by the chain so that repeated repartitioning can generate more and more chains.

We leave to the reader the further comparison of Zermelo's proof with the earlier proofs.

Chapter 14
Bernstein's Division Theorem

Bernstein's Division Theorem (BDT) states that if an infinite cardinal number is divisible by a finite number then the quotient is unique, namely, if $k\mathfrak{m} = k\mathfrak{n}$ then $\mathfrak{m} = \mathfrak{n}$, where k is a natural number, $\mathfrak{m}$, $\mathfrak{n}$ cardinal numbers.[1] The theorem is included (p 122) in Bernstein's doctorate dissertation of 1901 (published in 1905). It was reproduced in Hobson 1907 pp 159–162. BDT is sometimes called Bernstein's theorem but since there are other results that bear Bernstein's name we use 'Bernstein's Division Theorem'.

Bernstein gave detailed proof only for the case $k = 2$. His proof is a great example of proof-processing, probably from Schoenflies' proof of CBT, wetted, perhaps, with insight gained from Dedekind's *Zahlen*. The proof is notorious as complex mainly due to its careless notation. We reproduce the proof, ironing out in footnotes its exposition wrinkles. We precede the exposition with an outline of the proof's plan and its CBT roots. Following the proof for $k = 2$, we present Bernstein's sketched proof of the general case. This proof is incomplete and public opinion has it that Bernstein actually never proved the general case. Strangely, even Tarski (1930 p 248f, 1949a p 78) held that Bernstein only proved BDT for $k = 2$ and that it was only Lindenbaum (Lindenbaum-Tarski 1926 p 305) who first proved the general case, without AC. We believe otherwise and show how Bernstein's proof can be completed, within the mathematics Bernstein employed for the special case. We then discuss briefly an inequality result proved by Bernstein with the same technique he used to prove BDT.

BDT turned out to be the source of several research projects that had considerable consequences. D. Kőnig's attempts (1914–1926) to prove BDT by implementing his father's (J. Kőnig) strings gestalt, which was devised for J. Kőnig's 1906 proof of CBT (see Chap. 21), led him to important results in

[1] It should be observed that despite the use of the notion of cardinal number here and below, BDT (and the inequality-BDT) are theorems about sets and mappings, just as CBT. Thus the proofs below do not seem to require that the sets involved be consistent.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_14,

graph theory, generally known as D. Kőnig's Theorem and D. Kőnig's Infinity Lemma (see Chap. 22). D. Kőnig's proof used, however, the axiom of choice, which Bernstein's proof avoids, a point missed by D. Kőnig and other commentators.

In 1914, Hausdorff, probably by proof-processing J. Kőnig's string gestalt against the backdrop of BDT, was led to the so called 'Hausdorff's paradox' (see Chap. 27) which in 1924 led to the 'Banach-Tarski paradox'. Similar proof-processing may have led to the earlier 'Mazurkiewicz-Sierpiński paradox' of 1914.

Within the Polish school of logic, in the 1920s, interest regarding the theorem grew because of several reasons. Sierpiński wanted to improve on D. Kőnig's proof by avoiding AC. Kuratowski, in preparation for the Banach-Tarski paper, wanted to specify how the given partitions correlate by the given mappings. Both Sierpiński and Kuratowski discussed the theorem for $k = 2$. Lindenbaum-Tarski touched on BDT in their attempt to catalog all the results in cardinal and ordinal arithmetic that do not apply AC. They also generalized Bernstein's inequality result to what we name the inequality-BDT. Later (late 1940s) Sierpiński linked BDT with the Lebesgue problem and proved the inequality-BDT for $k = 2$. Tarski then provided a proof for the general inequality-BDT. All these results will be surveyed in part IV below.

14.1 The Proof's Plan

By the conditions of the theorem there is a set S which is partitioned into two equivalent partitions x_1, x_2, and then again into two other equivalent partitions x_3, x_4. Let us call the x's halves. It is required to show that the halves of the different partitionings are equivalent.

Obviously each half of one partitioning is partitioned by the halves of the other partitioning into not necessarily equivalent partitions. Let us call the partitions of the halves 'quarters' and denote them, by x_{13}, x_{14}, x_{23}, x_{24}. In this notation two quarters with a common index are the partitions of the half with that index and each quarter is a quarter of two halves – one from each partitioning. Thus far we described the opening gestalt of the conditions of the theorem. The opening metaphor, which immediately associate the problem situation here with that of CBT, is that two mappings are given: between the halves. Proving that a half is equivalent to one of its quarters would entail that this half is equivalent to a subset of each of the halves of the other partitioning. Proving the same also for a half from that other partitioning, provides the conditions of CBT and hence, by that theorem, BDT follows. Thus we have outlined the metaphors that can provide the required result.

BDT is a partial reciprocal to CBT. This comes out clearly if we describe both theorems as follows: Let M, N be two sets, each partitioned into two partitions M_1, M_2, N_1, N_2. If we assume that $M \sim N_1$ and $N \sim M_1$ then we obtain that

M ~ N (this is CBT) and $M_1 \sim N_1$ but not necessarily that $M_1 \sim M_2$ or $N_1 \sim N_2$ or $M_2 \sim N_2$. If, however, we assume that M ~ N, $M_1 \sim M_2$ and $N_1 \sim N_2$, then we obtain that $M_1 \sim N_1$ and $M_2 \sim N_2$ (this is BDT) but not necessarily that $M \sim N_1$ or $N \sim M_1$.[2]

To prove that a half is equivalent to one of its quarters (the target quarter) it is enough to demonstrate that there is a 1–1 mapping that generates a chain of the 'other quarter' in the target quarter. Then the half can be 1–1 mapped onto the target quarter by pushing the other quarter down its chain, mapping all the other members of the half (if any, they belong to the target quarter), by the identity mapping. The partition and pushdown the chain metaphor is applied.

Bernstein's proof demonstrates that such a mapping exists, though it is not feasible to construct it, even though the proof does not use the axiom of choice. It should be noted that when Bernstein found his proof the axiom of choice had not been suggested yet, and the proof's avoidance of the axiom was not "on purpose".[3] But when Bernstein published his dissertation, the axiom was already known and widely discussed, mostly critically. Still, Bernstein did not comment on the avoidance of the axiom in his proof.

The question arises whence Bernstein obtained the partition and pushdown metaphor that he used in his proof of BDT. He could not have obtained it from Dedekind's proof of CBT (see 9.2) because it was only published in 1932. Since the same metaphor appears also in Zermelo's 1901 paper, in the proof of Theorem I, it seems that there was one source that influenced both mathematicians, who resided at the time in Göttingen. It could have been Dedekind's *Zahlen* directly, or it could have been Schoenflies' CBT proof (1900) which, like Borel's, emphasized the generation of a sequence of copies of a subset.

With regard to Dedekind, it is noteworthy that Bernstein (1905) did not mention him, unlike Zermelo in his 1901 paper. The reason could have been Dedekind's remark to Bernstein, when they met in 1897, that CBT can be proved by means of his chain theory (see Sect. 7.4). It is possible that as a result of Dedekind's remark Bernstein studied *Zahlen* carefully, but he surely did not want to demonstrate too close a link between that monograph and his own work.

Anyway, it is quite obvious that Bernstein had an original addition to the pushdown metaphor: whereas for Dedekind and Schoenflies (and Zermelo 1901) a reflection is provided by the context (of the single-set formulation of CBT), in Bernstein's context of BDT, no such reflection is given. Thus Bernstein needed to construct the images of the other quarter in the target quarter from the two given 1–1 mappings of the halves, which we designate by φ and ψ. The crucial observation that led to his proof seems to have been this: The mapping that takes x_1 onto x_2

---

[2] In this paragraph we give an example of proof-processing performed with no intention for its application at sight. The quest is for a certain symmetry (gestalt) obtained by reversal of roles (metaphor). Here the gestalt and metaphor are at a higher level above the context than in our other examples. Cf. 9.3 for a similar remark concerning Theorem 68 and CBT.

[3] But Bernstein wanted, no doubt, to avoid the well-ordering doctrine.

takes x_{13} into x_{23} or x_{24} and the mapping that takes x_3 onto x_4 takes x_{13} and x_{23} into x_{14} or x_{24}. Then again the mapping that takes x_1 onto x_2 takes x_{13} and x_{14} into x_{23} or x_{24}. If we assume the gestalt that no element of x_{13} is mapped by any composition of φ and ψ into x_{24}, we obtain by the metaphor of repeated application of these mappings the gestalt of a sequence of copies of x_{13} in x_{23} and x_{14}. At each round, x_{13} and its copies in either x_{23} or x_{14}, are mapped to x_{14} or x_{23}, respectively, and as the copies defined before each round are disjoint so are the copies produced at each round. Note that at each round exactly two new copies are generated from the two generated in the previous round and there are concrete mappings from each copy to the next one so that AC is avoided.[4] Note further that in the construction described here, no nesting sets are generated because members of x_{23} can be mapped into x_{24}; still a sequence of frames does cascade (see Sect. 9.1), and Bernstein was tuned to watch for frames through his own proof of CBT in its Borel rendering.

To establish the gestalt assumption that members of x_{13} or their copies, are never mapped to x_{24}, the halves, when this is not the case, can be replaced by others, equivalent, halves, for which the assumption holds. The new halves are obtained from the given ones by transferring a subset of x_{13} to x_{23} and an equivalent subset of x_{24} to x_{14}. We do not know whence Bernstein proof-processed this interchange metaphor.

14.2 The Proof

In the introduction to the theorem (Theorem 2)[5] Bernstein defines the union of sets, M + N, the multiplication of sets, MN, and the exponentiation of sets, M^N. Bernstein does not follow Cantor's notations (1895 *Beiträge* §3, 4), which were, respectively, (M,N), (M.N), (N|M). Bernstein does not limit the union operation to disjoint sets as did Cantor (implicitly also for the multiplication of sets). Bernstein notes that all the operations are associative and commutative. This seems to be a slip of tongue regarding exponentiation. He does not mention distributivity of multiplication over addition, which he uses below.

Theorem 2. Out of the equation 2M = 2N follows M = N.

Bernstein uses the same notation for sets and their powers,[6] and one needs the context to decide which meaning is intended. Thus, if we take here M, N to

[4] Carrying out the here described scheme without introducing some notation seems almost impossible, leading to the well-known conclusion that even for a sketch of mathematical reasoning, some notation is necessary.

[5] Theorem 1 (of the first chapter of the paper, p 121) is CBT which is given in its two-set formulation taken from 1895 *Beiträge*.

[6] Bernstein does not use the notion 'power', except once with regard to $\aleph_0$.

be sets, as the introduction to the theorem implies, M = N must mean M ~ N. Then we have to interpret 2M (and 2N) to mean $(M.M')((N.N'))$ and 2M = 2N to mean $(M.M') \sim (N.N')$, where M′ (N′) is equivalent but disjoint from M (N). But if we take M, N to be the powers of the sets M, N, then 2M (2N) must be taken as the power of $(M, M')((N, N'))$. Under this interpretation the equality signs need not be replaced by ~, and no sign is used for the multiplication operation of the cardinal numbers; Bernstein just juxtaposes the multiplicands.

We continue with Bernstein's presentation.[7]

> Before the proof we set out a series of easy to prove lemmas [*Hilfsätze*] about 1–1 mappings [*umkehrbar eindeutige Abbildungen*].
>
> Lemma 1. The 1–1 mappings of a set [*System*] S in itself form a group Φ_S.[8]
>
> Lemma 2. Be it given a series of 1–1 mappings (1) 1, χ_1, χ_2, χ_3, ... which [*es mögen*] forms a group, that is, (2) $\chi_\mu \cdot \chi_\nu = \chi_\rho$ and to every χ_μ there is [*möge*] one and only one χ'_μ such that (3) $\chi_\mu \cdot \chi'_\mu$[9] $= 1$(where 1 denotes the identity mapping). If now s is an element of S, for which (4) $s \neq \chi_\nu(s)$ $(\nu = 1, 2, 3, ...)$, then
> (5) $\chi_\mu(s) \neq \chi_\nu(s)$ $(\nu \neq \mu = 1, 2, 3, \ldots)$.[10]
>
> Lemma 3. If $s \neq s'$[11] and (6) $s \neq \chi_\nu(s')$ for ν=1, 2, 3, ..., then
> (7) $\chi_\mu(s) \neq \chi_\nu(s')(\mu, \nu = 1, 2, 3, ...)$.
>
> Because from (8) $\chi_\mu(s) = \chi_\nu(s')$[12] it follows, by multiplication by χ'_μ,
> $s = \chi_\nu \cdot \chi'_\mu(s') = \chi_\rho(s')$.[13]
>
> Definition [*Erklärung*]. If T_1, T_2, T_3, ... is a series of subsets of S, from which none has a common element with another, I call them a *disjoint system* of subsets.
>
> Lemma 4. If $T = \{t\}$[14] is a subset of S and always
> (9) $t \neq \chi_\nu(t')(\nu = 1, 2, 3, ...; t \neq t')$, then the equivalent subsets
> (10) $T, \chi_1(T), \chi_2(T), \ldots$ form a *disjoint system*.[15]

---

[7] Our comments to the proof, rather numerous due mostly to Bernstein's rather loose style, will be given in footnotes. Cantor must have been aware of this shortcoming of Bernstein when he remarked, after hearing J. Kőnig's 1904 Heidelberg lecture, which relied on a theorem from Bernstein's dissertation, that he suspects the king less than the king's ministers (Dauben 1979 p 249).

[8] Bernstein must be talking here about 1–1 mappings from S onto S. Lemma 1 is not used in the proof.

[9] Apparently the ′, here and below, belongs to μ not to the χ.

[10] Clearly Bernstein is talking about mappings from S onto S. Bernstein does not define the composition of mapping which he denotes by the same sign he sometimes uses for the multiplication of cardinal numbers. Condition (5) is rather: $\nu \neq \mu$, ν, $\mu = 1, 2, 3, \ldots$.

[11] Identifying 1 with χ_0, as Bernstein does below, makes this premise redundant.

[12] There is a typo in the original and '$\neq$' is printed instead of '$=$'.

[13] The sentence should have ended: 'a contradiction to the premise'. Note that contrary to current use (and Zermelo's) Bernstein writes $\chi_\nu \cdot \chi_\mu'(s')$ for $\chi_\mu'(\chi_\nu(s'))$. We stay here with Bernstein's convention.

[14] Bernstein is making use here of Cantor's convenient convention of denoting by $\{t\}$ (only they both use () instead of our { }) a set whose members are denoted by t with various subscripts or superscripts. Thus t' also belongs to T.

[15] Bernstein tacitly adopts the convention that χ(T) is the set of all $\chi(t), t \in T$.

Lemma 5. Under the premises of the previous lemma [*Satzes*] we have
(11) S = S + T.[16]

Because, denoting by $\aleph_0$ the power of the set of natural numbers, by repeated application of the equality (11)[17] we obtain a partitioning in the form:
S = T·$\aleph_0$ + R[18] so that T + S = T($\aleph_0$ + 1)[19] + R = T$\aleph_0$[20] + R therefore
S = S + T.

Proof of Theorem 2. We write the premises in the form:
(12) (a) S = $x_1 + x_2 = x_3 + x_4$,[21] (b) $x_1 = x_2$, (c) $x_3 = x_4$.[22] The mappings that correspond to the equality signs can be construed as 1–1 mappings of S into [should be 'onto'] itself, which we denote by φ_a,[23] φ_b, φ_c. As is directly obvious
(13) $\varphi_a^2 = \varphi_b^2 = \varphi_c^2 = 1$.[24]

As the figure below illustrates, the x's are partitioned according to (12a) in the following way: (14) $x_1 = x_{13} + x_{14}$, $x_2 = x_{23} + x_{24}$, $x_3 = x_{31} + x_{32}$, $x_4 = x_{41} + x_{42}$, where $x_{ij} = x_{ji}$.[25] (Fig. 14.1)

Let T_1 be any subset of x_1 and T_2 an equivalent subset of x_2, then one can transform x_1 and x_2, so that the elements of T_1 are interchanged with those of T_2. For the transformed x_1^* and x_2^* the relation (12) holds as well as the equations

[16] In this equation, S, T are the powers of the corresponding sets, not the sets themselves. Interpreting them as sets trivializes (11). Lemma 5 requires rather the conclusion of lemma 4 as premise not its premises, namely: T, χ_1(T), χ_2(T), ... form a disjoint system, then S = S + T.

[17] Bernstein's use of (11) for the proof of (11) seems to be a mistake and it is not necessary for under the premises of lemma 4, or assuming the conclusion of lemma 4, there exists a disjoint system of equivalents of T in S.

[18] R is the residue obtained from S after removing the disjoint system. Note that the equation S = T·$\aleph_0$ + R is an equation of powers and Bernstein uses the sign · for the multiplication of powers operation which he never introduced.

[19] Bernstein drops here and henceforth the use of '·' to denote the multiplication operation between sets, powers or mappings, and instead uses juxtaposition.

[20] Bernstein is using here the equality $\aleph_0 + 1 = \aleph_0$. Bernstein applies here the reemergence argument we encountered in Zermelo (1901) (see Sect. 13.2). Bernstein uses the reemergence argument also in other parts of his paper, e.g., p 129f. While Bernstein deduced from S = T·$\aleph_0$ + R that S = S + T, Zermelo (1901 theorem II) obtained S = T·$\aleph_0$ + R from S = S + T; thus we can say that Bernstein and Zermelo walked the same road in opposite directions.

[21] Bernstein takes 2M = 2N to imply that there is a set S of power 2M (2N) which has two partitionings: the partitions of the first partitioning are x_1, x_2, and those of the second are x_3, x_4. See the previous section.

[22] In b), c), '=' should be read as equivalence, '~'. While it is possible to read the x_i here as powers, later they surely mean sets.

[23] The introduction of φ_a is pointless; it can be taken as the identity. Indeed, φ_a is not used in the proof.

[24] Since x_1 is disjoint from x_2, given φ_b from x_1 onto x_2, we can extend it over x_2 by defining $\varphi_b(s)$ for every $s \in x_2$ to be equal to φ_b^{-1}(s). A similar argument holds for x_3, x_4 and φ_c. Thus property (13) for (the extended) φ_b and φ_c is established.

[25] The given proof can be conducted for each of the quarters, in the role of x_{13}, and its diagonally opposing quarter, instead of x_{24}.

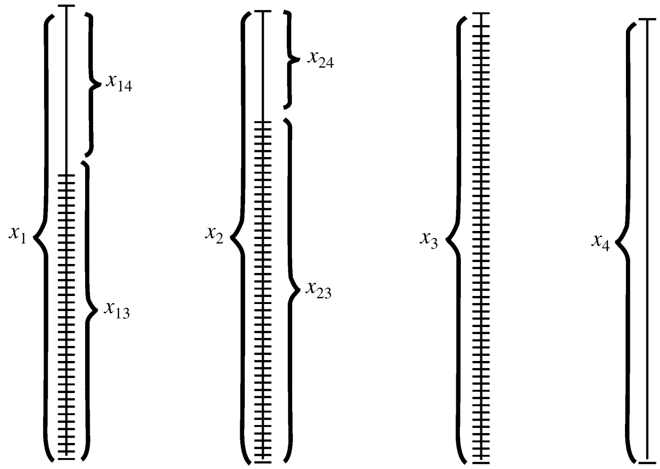


Fig. 14.1 The partitioning drawing for BDT

(14*) $x_1{}^* \sim x_1, x_2{}^* \sim x_2$; and likewise $x_3{}^* \sim x_3, x_4{}^* \sim x_4$. It is therefore enough to demonstrate Theorem 2 for $x_1{}^*$, $x_2{}^*$, to be able to make the conclusion for x_1, x_2. In a similar way allow also the sets x_3 and x_4 such interchange transformation.[26]

The objective of the proof is to show that we can obtain by suitable transformations such a partitioning in the form (14), that x_{13} can be neglected against x_{14} as well as against x_{23}, namely that $x_1 = x_{13} + x_{14} \sim x_{14}$ and $x_3 = x_{13} + x_{23} \sim x_{23}$. Because from this follows the equations $x_1 = x_2 = x_{14}$, $x_3 = x_4 = x_{23}$.[27] Thereby are given the premises of the Equivalence Theorem [namely, CBT] for the sets x_2 and x_4. Hence the conclusion can be drawn that $x_2 = x_4$ which entails the assertion [Theorem 2].

We search for our aim to establish the set x_{13} with x_{14} and x_{23} in suitable relation. For this purpose we construct out of the mappings φ_b and φ_c all possible compositions, which we arrange in two sequences:

(15) $1 = \chi_0, \varphi_\mathrm{b} = \chi_2, \varphi_\mathrm{b}\varphi_\mathrm{c} = \chi_4, \varphi_\mathrm{b}\varphi_\mathrm{c}\varphi_\mathrm{b} = \chi_6, \ldots;$[28] $\varphi_\mathrm{c} = \chi_3, \varphi_\mathrm{c}\varphi_\mathrm{b} = \chi_5, \varphi_\mathrm{c}\varphi_\mathrm{b}\varphi_\mathrm{c} = \chi_7, \ldots.$

To every mapping χ there is one and only one inverse, because by considerations from (13): (16) $\chi_{4n+2}\chi_{4n+2} = 1, \chi_{4n}\chi_{4n+1} = 1, \chi_{4n+3}\chi_{4n+3} = 1[, \chi_{4n+1}\chi_{4n} = 1]$.

[26] Bernstein's way of expression in this passage is confusing: he suddenly uses the equivalence sign ' ~ ' which he avoided before; then he designates by (14*) something that is not related to (14); he continues as if theorem 2 provides a conclusion with regard to x_1 and x_2 only and an analogous case needs to be considered for x_3 and x_4. Obviously interchanging T_1 and T_2 affects all four halves. Note that once φ_b and φ_c are regarded as transformations of S they are not affected by an interchange. Bernstein is correct, of course, in maintaining that if theorem 2 is proved for the transformed halves it is also proved for the original halves, and this is the main point for it is through interchange transformation that Bernstein intends to change the original partitionings of S if they do not come under the conditions of his proof-plan.

[27] Here the equality sign should be replaced by the equivalence sign, which Bernstein briefly used in the preceding equations.

[28] Bernstein did not define χ_1 and it caused some blunder below.

The composition of any other mappings produces again a mapping from the same sequence,[29] which is different from 1.

As a result of their formation the χ form a group of 1–1 mappings from $S = x_1 + x_2$ into [onto] itself.

There are two possible cases with regard to the mapping of an element e_{13} of x_{13}:

(1) e_{13} is transferred by some χ mapping with *finite* index[30] into an element of x_{24}.
(2) e_{13} is never transferred by a mapping from the sequence (15) to an element of x_{24}.[31]

The essential idea consists now in that by interchange of elements from x_{13} against elements from x_{24}, it is achieved that solely the second case takes place.

Assuming namely, that indeed the second case holds for all elements e_{13}, then I claim, that the image $\chi_{2\nu}(x_{13})[\nu > 0]$ (respectively $\chi_{2\nu+1}(x_{13})[\nu > 1]$) alternatively lies in x_{23} and x_{14}, and there a simply infinite disjoint system of subsets is formed.[32] But hereby it follows from Lemma 5 that $x_{13} + x_{14} = x_{14}, x_{13} + x_{23} = x_{23}$,[33] from which, as it was demonstrated above, our theorem results.

In fact,[34] in the second case x_{13} is fully transferred by the mapping χ_2 to x_{23}. By χ_4[35] the elements of x_{23} are transferred into such of x_{14} or x_{24}, hence by (2) $\chi_4(x_{13})$ is transferred entirely in x_{14}.

By complete induction we can proceed to conclude, that by each mapping with an index $4n$ the set x_{13} is fully transferred in x_{23}, and with each mapping with index $4n + 2$ the set x_{13} is transferred in x_{14}. The analog occurs with the mappings χ_{4n+1} and χ_{4n+3} instead.[36]

Finally, the mappings χ of the series fulfill the requirement that they are 1–1 mappings of S in [on] itself, besides they build this part of a group of such mappings, where for each element there is one and only one inverse element of the group.[37] Considering in addition,

[29] From the sequence of χ_ν; the subsequences of mappings with odd or even indices are not closed under multiplication.

[30] We don't know why Bernstein thought it necessary to stress the finitude of the index when no other possibility was discussed or even seems possible in the context.

[31] In focusing on the mappings of an element, Bernstein introduces for a moment the string gestalt that will appear in the CBT proof of J. Kőnig from 1906 (see Chap. 21). Perhaps J. Kőnig proof-processed this gestalt from Bernstein here.

[32] In each of x_{23}, x_{14} a disjoint system is formed. The attribute 'simply infinite' (it must mean 'denumerable') is redundant.

[33] The equality signs should be read as equivalence' $\sim$ ' or the x_{ij} as powers. Lemma 5, however, cannot be applied as given to obtain the desired results $x_{13} + x_{14} \sim x_{14}, x_{13} + x_{23} \sim x_{23}$. Instead the following lemma 5* is necessary: if $S', T \subseteq S$ and S' has a disjoint system of subsets all equivalent to T, then $S' + T \sim S'$. The proof of lemma 5* is the same as the proof of lemma 5. Figuratively speaking, lemma 5* (5) is the convex (concave) version of lemma 5 (5*).

[34] This and the next two paragraphs, justify the claim with regard to case 2).

[35] There is a typo here in the original: x_4 is printed instead of χ_4 and actually it should be χ_3 as explained in the next footnote.

[36] There is a typo here in the original: x_{24} is printed instead of x_{14}. Besides, Bernstein confuses the cases and does not point out that for the odd cases the analogy is not straight but crossed. The correct assignment is: $\chi_{4n},\ n > 0$, maps x_{13} into x_{14} and likewise χ_{4n+3}; χ_{4n+2} maps x_{13} into x_{23} and likewise $\chi_{4n+1},\ n > 0$. So the alternation occurs every third (not every second) step. Why all the images of x_{13} are disjoint is explicated below.

[37] "*ausserdem bilden sie den Teil einer Gruppe von solchen Abbildungen, wo es zu jedem Element ein und nur ein inverse Element der Gruppe gibt*." Bernstein seems to say that the series of χ's forms a sub-group of the group of all 1–1 mappings from S onto itself, denoted in lemma 1 by Φ_S. Why he mentions only the property of having an inverse and not the property of closure – we don't know. Anyway, the group nature of the χ's, the existence of an inverse, is apparently applied

that the image of x_{13} never falls back over x_{13}, it is thus implied, that the application of Lemmas 1–5 is permitted, which leads to the desired conclusion.[38]

Now we move on to the general case (1). We understand by x'_{13} those elements[39] of x_{13} which are transferred into x_{24} by χ_2.[40] Further, by x''_{13} we understand those elements different from x'_{13},[41] which are transferred by χ_3 into such elements of x_{24}, which were not hit by χ_2.[42] In the same way we define $x_{13}{}^{(3)}, \ldots, x_{13}{}^{(\nu)}, \ldots$.[43] We obtained the following scheme (17):[44]

$$\begin{array}{lll} x'_{13}, & \chi_2(x'_{13}) & \text{in } x_{24}, \\ x'_{13} \neq x''_{13}, & \chi_2(x'_{13}) \neq \chi_3(x''_{13}) & \textit{in } x_{24}, \\ x'_{13} \neq x''_{13} \neq x'''_{13}, & \chi_2(x'_{13}) \neq \chi_3(x''_{13}) \neq \chi_4(x'''_{13}) & \text{in } x_{24}, \\ x'_{13} \neq x''_{13} \neq \ldots \neq x^{(\nu)}{}_{13}, & \chi_2(x'_{13}) \neq \chi_3(x''_{13}) \neq \ldots \neq \chi_{\nu+1}(x^{(\nu)}{}_{13}) & \text{in } x_{24}, \end{array}$$

We form now the equivalent sums:

$\bar{x}_{13} = \Sigma_{\nu=1\ldots\infty} x^{(\nu)}_{13}$ and $\bar{x}_{24} = \Sigma_{\nu=1\ldots\infty} \chi_{\nu+1}\left(x^{(\nu)}{}_{13}\right)$.

We now consummate an interchange of $\bar{x}_{13}$ with $\bar{x}_{24}$; we set

(18) $x_{13} = \bar{x}_{13} + \bar{\bar{x}}_{13}$, $x_{24} = \bar{x}_{24} + \bar{\bar{x}}_{24}$[45] and we denote the transformed x with x^*, so that

(19) $x^*_{13} = \bar{\bar{x}}_{13}$, $x^*_{14} = x_{14} + \bar{x}_{24}$,[46] $x^*_{24} = \bar{\bar{x}}_{24}$, $x^*_{23} = x_{23} + \bar{x}_{13}$.[47]

---

to establish that the χ's are 1–1, though this seems to follow from the fact that the χ's are compositions of 1–1 mappings.

[38] Lemma 1 is not mentioned in the proof. Lemmas 2–4 are not necessary for lemma 5 or 5*, which only require that the images of x_{13} under the χ's be different. This point can be established directly: By φ_b, namely χ_2, x_{13} is mapped into x_{23} and its copy is thus disjoint from x_{13}. φ_c then maps the copy of x_{13} in x_{23} into x_{14} and this copy is thus disjoint from x_{13}. This copy then is a result of applying χ_4 to x_{13}. φ_b then maps the copy from x_{14} into x_{23} and this new copy is disjoint from the previous one because they are both obtained from disjoint subsets of x_1 by the same 1–1 mapping. This new copy is the result of applying χ_6 to x_{13}. Continuing in this way we see that each χ_{4n+2} produces a new copy of x_{13} in x_{23}, disjoint from the previous ones, and χ_{4n} likewise produces a new copy of x_{13} in x_{14} which is again disjoint from the previous ones. In Bernstein's proof this procedure is unnecessarily doubled by the oddly indexed sequence of χ's. In fact it is enough for the proof to take the sequence of even (or odd) χ's. $\varphi_b\varphi_c$ ($\varphi_c\varphi_b$) map each copy of x_{13} in x_{23} (x_{14}) to the next one.

[39] We should read: the set of those elements.

[40] Bernstein writes χ_1, which he did not define. Therefore, here and below, we increase the index of the χ's mentioned by 1. This is the blunder we mentioned earlier.

[41] Namely: the set of those elements [of x_{13}] different from the elements in x'_{13}.

[42] Bernstein is not worried about elements of x_{13} transferred by χ_3 to elements of x_{24} already hit by χ_2 because after the interchange consummated below these elements of x_{24} will be moved and so χ_3 will no longer take members of x_{13} to x_{24}.

[43] Thus, for $\nu > 1$, $x^{(\nu)}{}_{13} = \chi_{\nu+1}{}^{-1}\left(\chi_{\nu+1}(x_{13}) \cap x_{24} - \Sigma_{\mu=1\ldots(\nu-1)}\chi_{\mu+1}\left(x^{(\mu)}{}_{13}\right)\right) - \Sigma_{\mu=1\ldots(\nu-1)}x^{(\mu)}{}_{13}$.

[44] Scheme 17 is graphically similar to the scheme in Zermelo 1901 proof of theorem I (see 0). Apparently, both mathematicians were influenced by a similar scheme used in Schoenflies 1900 proof of CBT. We wonder if in circumstances around BDT begins the bitter feelings that Zermelo developed towards Bernstein (Peckhaus 1990 p 48; Ebbinghaus 2007 §2.8.4).

[45] In (18) Bernstein defines $\bar{\bar{x}}_{13}$ and $\bar{\bar{x}}_{24}$.

[46] There is a typo here in the original and $\bar{x}_{14}$ is written instead of $\bar{\bar{x}}_{24}$.

[47] Symmetrically, $\bar{x}_{13}$ could have been transferred to x_{14} and $\bar{x}_{24}$ to x_{23}.

The new $x_1^*, x_2^*, x_3^*, x_4^*$ emerge then after formula (14).[48] In accordance with the remarks made in the beginning we have now only to occupy ourselves with them. Examining now the content of $x^*{}_{13} = \overline{\overline{x}}_{13}$ against the mappings (15), we find that now only the second case can occur. Because assume some element of $x_{13}^*{=}\overline{\overline{x}}_{13}$ is transferred by $\chi_{\nu+1}$ $[\nu > 0]$ into $x^*{}_{24}{=}\overline{\overline{x}}_{24}$; this will contradict with the definition of $x^{(\nu)}{}_{13}$, because $x^{(\nu)}{}_{13}$ should contain all elements of x_{13}, which are not in
$x'_{13}...x^{(\nu\text{-}1)}{}_{13}$ and are transferred [by $\chi_{\nu+1}$] to elements of x_{24}, which do not lie in $\Sigma_{\mu=1...(\nu-1)}$ $\chi_{\mu+1}(x^{(\mu)}{}_{13})$.[49]

If the considered sets are finite, this must entail, by the provision of the consummated interchange specified, that x_{13} and at that also x_{24} vanish entirely. This one can easily visualize to oneself.[50] In the general case we proceed, as already specified.

It results that $x_1^* = x_3^*$ and by (14*) $x_1^* \sim x_1, x_3^* \sim x_3$ so that it follows
(20) $x_1 = x_3 = x_2 = x_4$.[51]

14.3 Generalizations of the Theorem

The proof of BDT for $k > 2$ was only sketched by Bernstein. First Bernstein changed the notation of the proof for $k = 2$ to ease the generalization. Instead of halves he denoted what might be called rows x_i and columns y_i. These intersect in cells (instead of quarters) $x_{ij} (= y_{ji})$ (as suggested by the matrix (2) below).[52] To prove that the row x_1 is equivalent to its x_{11} cell, it is proved that $x_{11} \sim x_{11} + x_{1i}$ for every $i > 1$. The proof idea is the same as for $k = 2$, except that the interchange step for $i > 2$ requires a special treatment to assure that it will not damage the previous cells x_{1j}, $1 < j < i$. The generalization given by Bernstein runs as follows:

Theorem 3. From $n{\cdot}\mathrm{M} \sim n{\cdot}\mathrm{N}$[53] follows M ~ N when n signifies an arbitrary finite number.

The proof is an exact generalization of the above conducted proof. The only difficulty consists in the correct choice of the sequence of mappings χ.[54]

Setting accordingly:
(1)[55]

$$x_1 + x_2 + ... + x_n = y_1 + y_2 + ... + y_n$$
$$x_1 = x_2 = x_3 = ... = x_n$$
$$y_1 = y_2 = y_3 = ... = y_n$$

[48] By combining the x_{ij}^* according to the equations in (14).

[49] In the original ℵ is written instead of μ, which is surely a typo.

[50] "*Sind die betrachteten Mengen endlich, so muss der nach der angegebenen Vorschrift-vollzogene Austausch zur Folge haben, dass x_{13} und dann aber auch x_{24} völlig verschwenden sind. Dies kann man sich leicht veranschaulichen.*" It seems that Bernstein wanted to emphasize here that if the exchange empties the sets, which can happen even if the sets are infinite, the result still holds.

[51] Surely Bernstein means $x_1 \sim x_3 \sim x_2 \sim x_4$; likewise it should be $x_1^* \sim x_3^*$, as implied by lemma 5*.

[52] For the generalized theorem Bernstein changed his notation calling by y the partitions of the second partitioning. This change suggests that the generalized theorem was proved after the case $k = 2$.

[53] Bernstein uses 'n' for our 'k' and he now uses ' ~ ' where in theorem 2 he used ' = '. Here, then, we can interpret $n{\cdot}$M as {0, 1, . . ., n-1}·M, etc.

[54] Actually the problem is with the interchange not the choice of the χ's.

[55] In the second and third rows, ' = ' should be read as equivalence, ' ~ '.

We Denote the mappings of the sequence of x with φ_{12}, φ_{13},..., the mappings of the y with ψ_{12}, ψ_{13},...,[56] and we generate in addition the partitioning
(2)[57]

$$x_1 = x_{11} + x_{12} + \ldots + x_{1n}$$
$$x_2 = x_{21} + x_{22} + \ldots + x_{2n}$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$x_n = x_{n1} + x_{n2} + \ldots + x_{nn}$$

So we have to show that, with proper interchange (3)[58] $x_{11} = x_{11} + x_{12}$, and with additional interchange $x_{11} = x_{11} + x_{13}$, and so on. Thereby, taking all the interchanges at once, it then becomes that
(4) $x_{11} = x_{11} + x_{12} = x_{11} + x_{13} = x_{11} + x_{14}\ldots$. But this has, as it is directly seen, the consequence that $x_1 = x_{11}$ that is $\leq y_1$.[59] The symmetry of the presuppositions permits in the same way a development for the y. We arrive thereby to the relation (5) $y_1 = y_{11}$, that is $\leq x_1$, and it follows by the Equivalence Theorem [CBT] that (6) $x_1 = y_1$.

In order to find the interchanges which lead to the equation (3), we have to set up $\chi_0 = 1$, $\chi_1 = \varphi_{12}$, $\chi_2 = \varphi_{12}\psi_{12}$, $\chi_3 = \varphi_{12}\psi_{12}\varphi_{12}$,[60] ... and then according to schema (17) [of Theorem 2] do the interchange so that the images of x_{12} are either in x_{22} or in x_{11}.

Here Bernstein left the proof, obviously incomplete; we will attempt to complete it for the case $n = 4$, which represents the general case.

Let χ_n^i, $1 < i$, denote the sequence of mappings constructed as in Bernstein's proof from φ_{1i}, ψ_{1i}. Thus χ_n^2 is Bernstein's χ_n. Remember that the mappings in χ_n^i are applied from left to right. Thus, the odd indexed χ_n^i map x_1 to x_i and vice versa, while the even indexed χ_n^i map y_i to y_1 and vice versa.

Let A_j be the set of all members of x_{12} that are mapped by χ_1^2 to x_{2j}, $j \neq 2$. Let $B_j = \chi_1^2(A_j)$. We move the members of A_j to x_{1j} and the members of B_j to x_{22}. The power of the rows and columns will not change by this interchange, but we

[56] The φ's (ψ's) are between the first row (column) and each of the others. Like in theorem 2 we assume that φ_{1i} (ψ_{1i}) are defined on x_i (y_i) by φ_{1i}^{-1} (ψ_{1i}^{-1}). Expanding the mappings to S (the union of the rows (columns)) is not necessary.

[57] Bernstein does not say explicitly that $x_{ij} = x_i \cap y_j = y_{ji}$. All the ' $=$ ' signs in this paragraph are to be read as ' $\sim$ '. Obviously the proof is symmetrical with respect to which cell is chosen to be established as equivalent to the row. Bernstein takes x_{11}.

[58] In this paragraph too ' $=$ ' should be read as ' $\sim$ '.

[59] The inequality sign requires regarding x_1, y_1 as powers

[60] The rightmost φ in the definition of χ_3 in the original is φ_{13}, which seems a typo for there is no reason to use φ_{13} at this stage of the proof. The definition of the χ_i here is simpler than in the proof of theorem 2 and corresponds to the even χ_i there. We have noted previously that the same simplification could be applied to theorem 2. Note that the definition of χ_0 is redundant as the χ_i do not form a group ($\varphi_{12}\psi_{12}$ does not have an inverse). The 1–1 character of the χ_i is guaranteed directly by the 1–1 character of the φ's and ψ's. The simplified definition of the χ's indicates, again, that the generalized theorem was found and proved after the case $k = 2$. We learn here something about Bernstein's character: he did not bother to go back to his proof of theorem 2 to streamline it with the insight gained in the proof of theorem 3. Bernstein appears indifferent to perfecting his public results.

need to modify the definition of the ψ's. χ_1^2 will now map any member of x_{12} only to x_{22}.[61]

Let now A_j be the set of all members of x_{12} that are mapped by χ_2^2 to $x_{j1}, j \neq 1$. Let now $B_j = \chi_2^2(A_j)$. We move the members of A_j to x_{j2} and the members of B_j to x_{11}. No element will be moved into $x_{j1}, j \neq 1$. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the φ's. χ_2^2 will now map any member of x_{12} only to x_{11}.

Let now A_j be the set of all members of x_{12} that are mapped by χ_3^2 to $x_{2j}, j \neq 2$. Let $B_j = \chi_3^2(A_j)$. We move the members of A_j to x_{1j} and the members of B_j to x_{22}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the ψ's. χ_3^2 will now map any member of x_{12} only to x_{22}.

Moreover, the images of any e from x_{12} under χ_1^2 and χ_3^2, both in x_{22}, will be different. The reason is this: $\chi_2^2(e)$ is in x_{11}, so it is different from e and thus φ_{12}, the last component of χ_3^2, takes it to a an element different from $\varphi_{12}(e) = \chi_1^2(e)$, since φ_{12} is 1–1. Also the images of different e's from x_{12} will be different because they are generated by 1–1 mappings from different origins.

Let now A_j be the set of all members of x_{12} that are mapped by χ_4^2 to $x_{j1}, j \neq 1$. Let now $B_j = \chi_4^2(A_j)$. We move the members of A_j to x_{j2} and the members of B_j to x_{11}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the φ's. No element will be moved into $x_{j1}, j \neq 1$. χ_4^2 will now map any member of x_{12} only to x_{11}. By a similar argument to the one given above, the images of any e in x_{12} under χ_2^2 and χ_4^2, which are necessarily in x_{11}, will be different and different e's will have different images.

If we continue the process of interchange in a similar fashion[62] as outlined above for every χ_n^2, we will get that the members of x_{12} will be mapped for odd n only to x_{22} and for even n to x_{11}. All the images of the members of x_{12} will be different.

Hence we have obtained the possibility to prove that $x_{11} \sim x_{11} + x_{12}$. For every element e of x_{12} there is a simple chain in x_{11}. We can apply the pushdown metaphor and push every such element e down its simple chain. The construction yields also that $x_{22} \sim x_{22} + x_{12}$ but it seems that this result is not applied in the proof. The simple chains in x_{22} were constructed only to obtain the simple chains in x_{11}.

We can now repeat the algorithm symmetrically for x_{21}. In the process no element will be moved to x_{12} and so the result of the algorithm when applied to x_{12} will not change. Thus we will obtain that $x_{11} \sim x_{11} + x_{21}$. Note that if x_{12} or x_{21} become empty at any point in the process outlined above, the required equivalence is trivially obtained.

---

[61] To avoid cumbersome notation, we maintain the notations of the rows, columns, cells and their subsets after the interchanges.

[62] We have in fact to define by induction a sequence of nesting subsets of x_{12}, denoted by $x_{12}(n)$, which contains what is left in x_{12} after the n first interchanges. The intersection of $x_{12}(n)$, which we denote by x_{12} again, has the desired property.

We now turn to prove that $x_{11} \sim x_{11} + x_{13}$. Let A_j be the set of all members of x_{13} that are mapped by χ_1^3 to $x_{3j}, j \neq 2, 3$. Let $B_j = \chi_1^3(A_j)$. We move the members of A_j to x_{1j} and the members of B_j to x_{33}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the ψ's. χ_1^3 will now map any member of x_{13} only to x_{33} or x_{32}.

Let C be the set of all members of x_{13} that are mapped by χ_1^3 to x_{32}. Let A_j be the set of all members of C that are mapped by χ_1^2 to x_{2j}, $j \neq 2$. Let $B_j = \chi_1^2(A_j)$. We move the members of A_j, for $j \neq 3$, into x_{1j} and the members of B_j into x_{23}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the ψ's. Let us now move the members of $\chi_1^3(A_3)$, which are in x_{32}, into x_{33} and the members of B_3, which are in x_{23}, into x_{22}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the ψ's. χ_{12} will now map any member of C into x_{22}.

Let A_j be the set of all members of C that are mapped by χ_2^2 to $x_{j1}, j \neq 1$. Let $B_j = \chi_2^2(A_j)$. Because χ_1^2 maps all members of x_{13} into x_{22}, A_j is well defined : every member of C has an image under χ_2^2. We move the members of A_j into x_{j3} and the members of B_j into x_{11}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the φ's. χ_2^2 will now map any member of x_{13} into x_{11}.

If we continue the process in a similar fashion for every χ_n^2, we will get that the members of C will be mapped for odd n, only to x_{22} and for even n to x_{11}. All the images of the members of C will be different. We now move C to x_{12} and $\chi_1^3(C)$ to x_{33}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the ψ's. The new addition to x_{12} has the same properties of the previous x_{12} of having for every member a simple chain in x_{11} that can be pushed down. So we have not ruined the result obtained that $x_{11} \sim x_{11} + x_{12}$.

Let A_j be the set of all members of x_{13} that are mapped by χ_2^3 to $x_{j1}, j \neq 1$. Let $B_j = \chi_2^3(A_j)$. We move the members of A_j to x_{j3} and the members of B_j to x_{11}. The power of the rows and columns will not change by this interchange, but we need to modify the definition of the φ's. χ_2^3 will now map any member of x_{13} only to x_{11}.

Repeating the above process we get that members of x_{13} are mapped by χ_n^3, for odd n, to x_{33} and for even n to x_{11}. Thus we obtain the means to establish $x_{11} \sim x_{11} + x_{13}$. By a symmetrical argument for x_{31} we will get that $x_{11} \sim x_{11} + x_{31}$, without distorting already obtained results. The proof can be continued in a similar fashion for x_{14}. A similar proof can be constructed for any n in the premises of the theorem[63].

[63] It will be interesting to find an algorithm with only one inductive process.

14.4 The Inequality-BDT

Following his sketch of a proof for the generalized BDT, Bernstein introduced the following as Theorem 4: "Out of $2x = x + a$ it follows that $x \geq a$."[64] Bernstein described it as a theorem that handles subtraction. For its proof Bernstein used the same procedure he applied to prove his Theorem 2: Let x_1, x_2, x_3, x_4 be sets such that $x = x_1 \sim x_2$[65] by φ (which is defined on x_2 to be equal to φ^{-1}) and $x = x_2 \sim x_3$ by ψ, which is similarly extended to x_3. a is the power of x_4. The proof of Theorem 2 provides that x_{14} can be neglected so that $a = x_4 \sim x_{24} \subseteq x_2 \sim x$ and thus it follows that $x \geq a$.

Bernstein did not mention the obvious generalization that $2x \geq 2a$ entails $x \geq a$. This is the inequality-BDT for $k = 2$. The general theorem is: $k\mathfrak{m} \geq k\mathfrak{n} \rightarrow \mathfrak{m} \geq \mathfrak{n}$, when k is finite and $\mathfrak{m}$, $\mathfrak{n}$ cardinal numbers. The inequality-BDT was first announced in Lindenbaum-Tarski 1926. Its proof was published only in the late 1940s. We will cover these future developments of the inequality-BDT in part IV.

---

[64] x, a are powers not sets.

[65] $x = x_1$ means that x_1 is of power x, and naturally so is x_2. Similarly for x_3.

Chapter 15
Russell's 1902 Proof of CBT

Whitehead's 1902 paper "On cardinal numbers" was aimed at establishing Cantor's theory of cardinal numbers on concepts of logic and in the symbolism of Peano and Russell developed by that time. For background on logicism, to which Whitehead 1902 belongs, and its relation to set theory, see Grattan-Guinness 1997 or 2000 Chap. 6. Whitehead's paper includes, in its third section, a proof of CBT, which we review below. Whitehead said that section III is "entirely due to Russell and written by him throughout". We therefore refer to the writer of our aspect of the paper as Whitehead-Russell or WR, or sometimes just as Russell. CBT was not mentioned in Russell's earlier papers.

Russell's contribution to Whitehead 1902 was written between January and late June 1901 (Grattan-Guinness 2000 p 307). It is thus unlikely that Russell saw Zermelo's paper (see Chap. 13), to which it somewhat resembles, when he wrote section III. Zermelo presented his paper to the Göttingen Academy of Sciences on March 9, 1901 (Moore 1982 p 89) but the paper did not actually appear in print until 1902. Zermelo's paper is referenced in "The Principles of Mathematics" (POM) part V p 306. That part was written in late 1900 (Grattan-Guinness 1977 p 133), so we must assume that Russell had updated it with the reference to Zermelo in 1902 (probably June – Grattan-Guinness 2000 p 323, Russell 1993 p 422). Whitehead's 1902 paper was published in October 1902, so, apparently, Russell did not update it in view of his reading of Zermelo's proof.

Russell's proof is the first formal proof of CBT. By 'formal' we mean that the proof consists of a sequence of sentences written in pasigraphy, where each sentence is a theorem either proved earlier or obtained from earlier sentences in the list by a stated logical rule. Schröder too employed pasigraphy in his CBT proof (see Sect. 10.1) but his reasoning was in natural language. WR call CBT "the Bernstein's and Schröder's theorem" (pp 369, 383). They were thus the first to name the theorem after mathematicians.

Russell's proof of CBT turns out to be defective on two counts: First, it relies on a postulate that requires the axiom of choice. Second, it is circular, in both its argument and its methodology. For these reasons we will not compare it to previously obtained CBT proofs.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_15,

15.1 The Core Arguments

CBT was presented in Whitehead's 1902 paper in its two-set formulation[1]:

$*4.5\ u, v \in \text{cls}\ .\ \exists \mu u \cap \text{cls}`v\ .\ \exists \mu v \cap \text{cls}`u\ .\supset .\ u \text{ sim } v$

'cls' is the class of all classes and $u \in$ cls is a way of saying that u is a class. In 1902 'class' was just the corresponding term to Cantor's 'set'. Later, with the emergence of Russell's no-class theory, 'class' lost its ontological commitment so the two terms were differentiated. WR's pasigraphy is basically that of Peano: '.' is Peano's sign for 'and'; '$.\supset.$' is his sign for implication. We may use instead $\rightarrow$. The existential quantifier ($\exists$) signifies that the expression on its right side is not empty (cf. Grattan-Guinness 1977 p 71f). WR denote by μu the cardinal of u which they identify with the collection of all classes equivalent to u (*1.2); they denote by 'cls'v' (read: 'class of v') the class of all subclasses of v. The second (respectively, third) expression of *4.5 means then that there is a subclass of u (v) in the cardinal number of v (u) [equivalent to v (u)]. WR's 'similar' ('sim', *1.1) is Cantor's 'equivalent', which he denoted by '~'.

The proposition *4.5 then reads: if [(u, v are classes) and (the intersection of (the cardinal of u and the class of subclasses of v) is not empty) and (the intersection of (the cardinal of v and the class of subclasses of u) is not empty)] then (u is equivalent to v).

Note that for the definition of cardinal numbers, CBT is necessary for otherwise, if $A \subset B \subset C$ are classes and A, B are in the same cardinal, nothing can warrant that B is of the same cardinal. In addition, cardinal inequality defined in *3.1 is not transitive. These are major lacunae of the 1902 paper, which are not mentioned in Russell 1993. This is the methodological circularity we mentioned above. It was corrected in Principia Mathematica (PM), perhaps under Poincaré's emphasis (1906a p 27) that CBT is the fundamental theorem of the theory of infinite cardinal numbers. Of a similar nature was Bernstein's note (1906 p 953) that CBT must be established before induction is introduced, unlike the order of presentation in Whitehead 1902.

For proof of the theorem, WR bring the following sequence of assertions:

(1) $S \in 1\rightarrow 1\ .\ \sigma = u\ .\ \breve{\sigma} \supset v\ .\ S' \in 1\rightarrow 1\ .\ \sigma' = v\ .\ \breve{\sigma}' \supset u\ .\supset .\ \sigma(\sigma') = u\ .\ \breve{\sigma}'(\breve{\sigma}) \supset \breve{\sigma}' \supset u$
(2) (1) $.\ \mu u = \alpha\ .\ \mu v = \beta\ .\ \mu u\text{-}\breve{\sigma}' = \gamma\ .\ \mu \breve{\sigma}'\text{-}\breve{\sigma}'(\breve{\sigma}) = \delta\ .\ \alpha \geq \beta\ .\ \beta \geq \alpha\ .\ \alpha = \beta + \gamma\ .\ \beta = \alpha + \delta.$
$\supset .\ \alpha = \alpha + \gamma + \delta$
(3) (2) $.\ *4.38\ .\supset .\ \gamma + \delta \leq \alpha\ .\supset .\ \delta \leq \alpha\ .\supset .\ \alpha + \delta = \alpha\ .\supset .\ \beta = \alpha\ .\supset .$ the theorem

*4.38 is: $\delta < \alpha\ .\supset .\ \alpha + \delta = \alpha.$

WR use, after Peano, a similar sign for subsumption (inclusion, our $\subseteq$) and implication (perhaps because $x \in P\ .\supset .\ x \in Q$ is equivalent to $P \supset Q$, Ferreirós 1999 p 305). WR denote by $x\,S\,y$ that the elements x and y are in relation S (§II 1.00).

[1] We reference section I of the paper by page number and other sections by the section number and item number. Default section is III.

By the lowercase Greek letter that corresponds to the Latin letter that denotes the relation, WR denote the domain of the relation, i.e., the collection of all things that are related to other things by the relation (appear on its left); the range of the relation, i.e., the collection of all things that are related to by the relation (appear on its right), WR denote by the same Greek letter as the domain with an inverted bow above (p 375). The signs $\sigma(\sigma')$ and $\breve{\sigma}'(\breve{\sigma})$ are not introduced in the paper; they surely stand for the domain and range of the composite relation SS′, respectively.

Proposition (1) reads: If [(S is a 1–1 relation (p 377) with domain *u* and range a subclass of *v*) and (S′ is a 1–1 relation with domain *v* and range a subclass of *u*)] then (the domain of the relation SS′ is *u* and its range is a subclass of the range of S′ which is a subclass of *u*).

Proposition (1) gives the domain and range of the composite relation SS′ (WR's 'relational product' (p 376)), expressed in terms of the domain of S and the range of S′. It further shifts *4.5 to the single-set CBT formulation and introduces the language of relations which lies under the language of cardinal numbers used in the expression of the theorem.

The use of relations instead of the usual 'mappings', is characteristic of Russell's writings. At the time Russell's logic of relations was considered a great discovery and was one of the touchstones of WR's formal language, including in PM. The difference between the two notions is now of little importance, but originally the use of relations assisted in distancing the comparability of classes from the analytic concept of function that then prevailed. On the debt of Russell to Peirce and Schröder with regard to his theory of relations, pointed out by Wiener, see Brady 2000.

Proposition (2), which takes the thesis of (1) as one of the conjuncts of its hypothesis, switches back to the language of cardinal numbers, in which the proof remains. The thesis of (2), which is part of the hypothesis of (3), reflects the conditions of the single-set formulation in cardinal language. Note that the first four equalities of the hypothesis of (2) are definitions and not assertions, and that in the definitions of γ and δ brackets are implied on the expression following μ.[2] Note further that the equalities and inequalities in the second part of the hypothesis of (2) do not rely on the intuition acquired from (1) but on the formal definitions of addition and inequality of cardinal numbers defined by WR in *1.92, *3.

WR did not define in 1902 equality of cardinal numbers explicitly but it appears that they take two cardinal numbers to be equal when they are equal as classes (extensionality). Obviously when cardinal numbers have a common member they are equal. Inequality, $\beta \leq \alpha$, is defined (*3.12, in the second of the two lines numbered with the same number by mistake; cf. Russell 1993 p 787) to hold when there is a class in β that is a subclass of a class in α. This definition fails if CBT is not already established (the circular methodology failure). For strict inequality, $\beta < \alpha$ (*3.1), it is further required that α and β be disjoint, namely,

[2] This shows that WR's use of pasigraphy at the time was far from being polished.

they are unequal (which is where WR lean on the missing definition of equality). WR's definition of strict inequality of cardinals corresponds to Cantor's definition of inequality between powers from his 1878 Beitrag and not to his definition of his 1895 *Beiträge*, which avoids the need for CBT for its transitivity. Once CBT was established Russell concluded (*4.51) that if $\alpha > \beta$ then $\neg(\beta \geq \alpha)$.[3] Grattan-Guinness (2000 p 308) says that thereby Russell established "trichotomy"; actually *4.51 only asserts that the three cases, with the above definition of strict inequality, are mutually exclusive and not trichotomy, which requires a proof that at least one of the cases occurs.

WR define the sum of α and β (*1.92) as the cardinal number of a class that is the union of two disjoint classes one from α and the other from β. Extensionality, and closure under equivalence of the cardinal number of a class, imply that the sum is unique. In proofs of the lemmas described here, symmetry and substitution rules for equality, as well as the commutativity and associativity of cardinal addition, are required, a point which WR ignore. The substitution and transitivity rules are also required to establish the conclusion in the implication of (2). These points are not mentioned in Russell 1993.

Proposition (3), in which the arguments to establish the theorem really begin, is best understood when it is decomposed to the following propositions:

- $\alpha = \alpha + \gamma + \delta \rightarrow \gamma + \delta \leq \alpha$; $\gamma + \delta \leq \alpha \rightarrow \delta \leq \alpha$ which result from the definitions of addition and inequality between cardinal numbers (*1.92, *3);
- $\alpha + \delta = \alpha \rightarrow \alpha = \beta$ which uses the given $\beta = \alpha + \delta$ to obtain the required result;
- $\delta < \alpha \rightarrow \alpha + \delta = \alpha$ which is *4.38;
- $\alpha = \alpha + \alpha$ which is *4.32, a reference that WR forgot to make. This proposition covers the case $\delta = \alpha$.

As proof of *4.38 WR give the following:
$(\alpha \leq \alpha + \delta \leq \alpha + \alpha \;.\; \alpha = \alpha + \alpha) \rightarrow \alpha = \alpha + \delta$. The first conjunct is derived (it too) from the definitions of addition and inequality between cardinal numbers (*1.92, *3); the second conjunct is *4.32 again. So we see that *4.32 is fundamental to the proof.

However, before we turn to the proof of *4.32 let us note that there is a mistake in the proof of *4.38. Substituting the second conjunct in the first brings the implication into $\alpha \leq \alpha + \delta \leq \alpha \rightarrow \alpha = \alpha + \delta$, but this implication is again CBT in its single-set formulation – the very theorem that WR set out to prove! WR's argument turns out to be a circular argument, a point not mentioned in Russell 1993. Interestingly, WR had at their fingertips a valid argument (metaphor) based on *4.32 (which they use anyway): from (2), $\alpha = \alpha + \gamma + \delta$ and using *4.32 for δ we get $\alpha = \alpha + \gamma + \delta + \delta = \alpha + \delta$.

It is possible that the pasigraphy clouded the identification of the problem, a shortcoming of the method which Poincaré will emphasize (1905) with regard

[3] WR denote negation not by $\neg$ but by – used in Whitehead 1902.

to Burali-Forti paradox, and it is perhaps with regard to such errors that Russell exclaimed (Grattan-Guinness 1977 p 74): "I would publish my views as I get them if they were not so liable to change and so full of errors which I keep on discovering". Of a similar circular argument that when spotted led to the discovery of the multiplicative axiom, Russell told Jourdain in 1906 (Grattan-Guinness 1977 p 80).

*4.32 is proved by two other lemmas: *4.31 which states, given a cardinal number α, that there is a cardinal number β such that $\alpha = \aleph_0\beta$[4] and *4.2 $\aleph_0 = \aleph_0 + \aleph_0$. To obtain *4.32 substitute $\aleph_0$ in *4.31 with $\aleph_0 + \aleph_0$: $\alpha = (\aleph_0 + \aleph_0)\beta = \aleph_0\beta + \aleph_0\beta = \alpha + \alpha$. The proof of *4.32 requires the definition of cardinal multiplication and proof of its distributivity and substitution laws. Multiplication and distribution are introduced in section IV of the paper, written by Whitehead (compare Grattan-Guinness 2000 p 308); substitution is, as already noted, ignored.

As for *4.31, Russell presents it as an immediate conclusion from *4.3 which says that every infinite class[5] can be partitioned into denumerable partitions, a proposition which WR assumed as a postulate, or "primitive proposition" (Pp) in their terminology. In his explication of *4.3 Russell did say that the reason he believes in *4.3 is that every set possess a denumerable partition (*2.6, note on p 381). So if the residue is infinite the process can be repeated until the residue is finite and can be stashed into one of the denumerable partitions. The number of the denumerable partitions is β of *4.31. The use of *4.3 to prove *4.32 is the new gestalt that WR brought in 1902 to CBT, corrected as suggested above.

Russell noted that Borel's proof does not make use of *4.3 (p 383) but he justified its postulation by the explication given above and because it provides nice results in cardinal arithmetic (*30–32). He added that its use in the proof of CBT was made just to demonstrate its importance. *4.31 can be established, of course, using the axiom of choice (cf. Moore 1982 p 9f).

In passing, let us note our understanding as to what Russell's attitude to axioms was. Russell was a realist (cf. Gödel 1944); while he accepted tentative postulation of certain propositions that describe the relations between entities of the world, he rejected the addition of entities by postulation. Thus he avoided Cantor's definition by abstraction and Dedekind's postulation of the irrational numbers to fill the gap in a cut, and instead took the class of all equivalent classes to represent a cardinal number and Dedekind's cuts (as classes of classes) to represent irrational numbers (as suggested by Weber in 1888 and rejected by Dedekind, see Ewald 1996 vol 2 p 835,

[4] WR use α_0 instead of $\aleph_0$. Russell told Jourdain in 1904 (Grattan-Guinness 1977 p 26) that he cannot draw the aleph sign; snobbism? Antisemitism? It seems that he did learn it finally. But then, Zermelo used $\mathfrak{a}$ for $\aleph_0$ (see Chap. 13) and we made no similar comment.

[5] The class of all infinite classes (cls_infin *1.82) is the union of (∪') the class of all infinite cardinal numbers (Nc_infin, *1.6) which is obtained from the class of all cardinal numbers (Nc *1.3) after removing the finite cardinal numbers (Nc_fin *1.6, *1.8). The class of finite classes (cls_fin) is the union of the class of finite cardinal numbers (*1,81). Note: WR use space for our underscore.

Ferreirós 1999 p 134 footnote 2, McCarty 1995 p 63 and passim, Weber 1906 173). Thus it seems that his dictum (1919 p 71) that "the method of 'postulating' what we want has many advantages; these are the advantages of theft over honest toil" must be understood to be directed towards postulating entities and not towards the description of their relations.

This dictum was indeed originally directed against Dedekind's abstraction of irrational numbers from cuts of the rationals. It seems that Russell saw Dedekind's abstraction as an act of postulation. Interestingly, Dedekind too was against postulation in the context of the natural numbers (introduction to 1963), but it seems that he did not consider abstraction as postulation. Russell's view differed from that of Peano who was a mentalist and said (Peano 1906 p 365) "We think of number therefore number exists" and he said there that postulates are satisfied by our ideas. Russell expressed a mentalist view with respect to models of physics (Grattan-Guinness 1977 p 124) but with regard to mathematical entities he seems to require that they exist not because of our thinking about them (Grattan-Guinness 2000 p 299).

15.2 The Definition of $\aleph_0$

The denumerable cardinal entered WR's proof through *4.2 $\aleph_0 = \aleph_0 + \aleph_0$ and *2.6 that every infinite set has a denumerable subset. Both theorems were considered as easy to prove in Cantor's 1878 *Beitrag* (1932 p 120).[6] But for WR, who refused to rely on intuitive considerations, the theorems needed grounding within their logicist system. For sake of completeness of our review and for later reference, we follow WR's presentation.

WR proved *4.2 $\aleph_0 = \aleph_0 + \aleph_0$ by taking as a class of cardinal number $\aleph_0$ the class of all finite cardinal numbers Nc_fin (*1.6). Then, using the sequent relation R (*2.2) which is defined in this class, WR define (*4.2) a relation R′ that holds between two finite numbers when the $2n$ power of the sequent relation holds between them for some finite n.[7] Then they partition Nc_fin into two classes, one containing the even finite numbers (the range of 0 under R′) and the other the odd finite numbers (the range of 1 under R′). The partitions are denumerable because each number in one of the partitions can be assigned uniquely to the n for which it relates to the generating number of the partition (0 or 1) by R^{2n}.

As for *2.6, first, relying on a proof by induction that if n is finite so is $n + 1$, it is proved that the class obtained from an infinite class u by removal of a finite number of elements, is still infinite. By another application of induction it is derived that u has a subclass of cardinal number m for every finite m. Then WR define a class K that contains such classes v of finite subclasses of u that have one element u_m of

[6] In fact, *4.2 is not explicitly stated in 1878 *Beitrag* but implied in the union theorem of p 120 and in (2) of p 130.

[7] The nth power of a relation (with range a subclass of its domain) is the relation composed with itself n times.

cardinal number m, which is a subset of the element u_{m+1}, for every finite m, so that all members of v are compatible. The union of v ($\cup$'v) of any member of K is thus a subclass of u and it belongs to $\aleph_0$ because every member of v can be assigned the first m to the u_m of which it belongs.[8] As expected, the axiom of choice is needed in the proof: to provide that K is not empty.[9]

The introduction of $\aleph_0$ is made in (*2.4):

$$\aleph_0 = \text{cls}\cap s\ni\{\exists 1\to 1\cap \text{S}\ni(\breve{\sigma} \supset \sigma \,.\, \exists\sigma\text{-}\breve{\sigma} : s \supset \sigma \,.\, \breve{\sigma}s \ \supset s \,.\, \exists s\text{-}\breve{\sigma}s \,.\supset.\, s = \sigma)\}$$ [10]

This definition, however, does not make sense: the variable s of the implication ought to be different from the first s (this is not noted in Russell 1993). Indeed, in Russell's 1901 paper (1956 *1.1 p 15, 1993 p 325), this definition is better worded:

$$\aleph_0 = \text{cls}\cap u\ni\{\exists 1\to 1\cap \text{S}\ni(u \supset \sigma \,.\, \breve{\sigma}u \supset u \,.\, \exists\, u\text{-}\breve{\sigma}u : s \in \text{cls} \,.\, \exists su\text{-}\breve{\sigma}u \,.\, \breve{\sigma}(su) \supset s. \supset_{s}. u \supset s)\}$$ [11]

Russell was kind enough to provide us with a transcription of the latter pasigraphy in natural language[12]:

> $\aleph_0$ is the class of the classes u such that there is a one-one relation S such that u is contained in the domain of S, and that the class of terms to which the different u's have the relation S is contained in u, without being identical with u, and which, if s is any class whatsoever to which belongs at least one of the terms of u to which any u does not have the relation S, and to which belongs all terms of u to which a term of the common portion of u and s has the relation S, then the class u is contained in the class s.

To use $\aleph_0$ in *4.2 one has to prove that $\aleph_0$ is a cardinal number (*2.41) and that it includes Nc_fin (*2.42).[13] For *2.41 WR refer to *1.9 of Russell's article of 1901 (1956 §3) which affirms that any two progressions (members of $\aleph_0$) are equivalent: Since a

---

[8] The formal proof in the paper is a bit difficult to read because in its line (4) the definition of K is embedded within a proposition about it and in (5) the '....' should be omitted and the subsumption is in fact an implication, missing its dots. These remarks on (5) are not noted in Russell 1993.

[9] A similar use was made by Dedekind (1963 #159).

[10] The reverse epsilon $\ni$ is Peano's notation. s$\ni\{\varphi(s)\}$ stands for what is currently defined by $\{s \mid \varphi(s)\}$. The clause 'cls$\cap$', so typical of the 1902 paper, is redundant here for the reverse epsilon operator produces classes (p 373). $\breve{\sigma}$s is the class of all images under S of members of s (§II 1.05).

[11] In the 1901 paper (Russell 1956) the definition is for ω, the first infinite ordinal, but in his 1902 paper, in a note to 1.35 (1993 p 391), Russell remarks that it is actually the definition of $\aleph_0$ (cf. Grattan-Guinness 1977 p 21 footnote 2, Grattan-Guinness 2000 p 306 (cf. p 309), Russell 1993 p 727 325:6). In the 1901 formulation R is used for S (thus also ρ, $\breve{\rho}$ for σ, $\breve{\sigma}$). Here, 'su' is used for the intersection of s and u and the subscript s on the implication sign stands for a universal quantifier on s over the implication the antecedent of which begins after the ':'. There is a typo in the 1901 formulation in that the')' before the '}' is missing (notes in Russell 1993 p 784). The formulation of 1901 is unnecessarily generalized (typical to Russell); it can be limited to the cases where s is a subclass of u (and S has u as domain); then the conclusion is $s = u$. Perhaps it was Russell's rejection of definition under hypothesis (Grattan-Guinness 1977 p 29) that motivated the generalization. In $\breve{\sigma}(su)$ the brackets were added in the original, contrary to the convention, no doubt to avoid ambiguity.

[12] The 1901 paper (Russell 1956) was originally written in French.

[13] It is not quite clear why WR did not define $\aleph_0$ as μNc_fin. Perhaps they wanted to define an infinite class independently of any notion of number, finite or infinite, to match Dedekind's achievement in *Zahlen* (1963). The spirit of Dedekind is felt in the definition of $\aleph_0$ also in the requirement that the image of u under S be a subclass of u, which is the qualifying requirement for

progression has a first element (u-$\breve{\sigma}u$ has only one member – there *1.2), given the sequent relations for u, v in $\aleph_0$ (the S in the definition of the members of $\aleph_0$) an equivalence between u and v (actually it is a similarity as it keeps the sequent relation) can be constructed by correlating the zeros of the two progressions and their nth sequents (obtained by the nth powers of the sequent relations), to form single pair-relations for each of the correlates, then take the union of these single pair-relations.[14] Next, (§3 *1.91) it is proved that any set equivalent to a progression is a progression: using the equivalence between $u \in \aleph_0$ and v, a sequent relation S′ on v can be defined from the sequent relation of u, so that v is in $\aleph_0$ too.

Finally, the proof of *2.42 Nc_fin $\in \aleph_0$ is easy: the sequent relation of Nc_fin (*2.2, *2.31, *2.32) is the required S of the definition of $\aleph_0$. The definition of Nc_fin, given in *1.6, is, however, more subtle:

$$\text{Nc\_fin} = \text{Nc} \cap n\ni\{s \in \text{cls} \,.\, 0 \in s : m \in \text{Nc} \cap s \,.\supset_m .\, m+1 \in s :\supset .\, n \in s\}$$

It translates to: Nc_fin is the collection of all cardinal numbers (Nc, *1.3) which are included in every collection s that has the following properties: 0 (*1.4) belongs to s and if m is a cardinal number that belongs to s then $m + 1$ also belongs to s. 0 is defined as the class whose only member is the empty class; $m + 1$ is the collection of all classes u such that if any member of u is removed from it the remaining class belongs to m (*1.5).[15]

The proposition inside the curled brackets in the definition of Nc_fin,[16] WR named 'Induct' and used it in proofs as a metaphor that provides complete induction. WR use Induct in a peculiar way: they simply put it among the propositions that comprise a formal proof without bothering to state which is the class s upon which the induction argument is intended, or demonstrating that this class fulfills the required induction conditions.

Because of the involvement of Induct in the definition of Nc_fin, WR stated (p 368) that "the finite cardinals are defined as those cardinal numbers which can be

---

Dedekind's chains; and note that by the relation of sequent defined in *2.2, Nc_fin is a simple chain.

[14] Russell assumes (1901 §1 *1.8) as primitive proposition the existence of a single pair-relation for every two elements, similar to Zermelo's pairing in his axiom of elementary sets.

[15] There are two typos in *1.5 that render it meaningless: the first minus sign should be replaced by the equality sign and the second epsilon sign should be replaced by reverse epsilon as in *1.43. See Russell 1993 p 787. Notice that as *1.5 stands, nothing prevents the empty class from belonging to every $m + 1$. This is not noted in Russell 1993 but noted in Grattan-Guinness 2000 p 308. In passing let us note that in Russell 1993 it is not noted that §V *23.1 (p 392) is also due to Russell, according to Whitehead's remark. It is, however, noted in Grattan-Guinness 2000 p 307. *23.1 is a proof of Cantor's theorem that is reminiscent of Russell's famous pile of socks example. Russell begins with a class a of pairs, say the class of the pairs (yes γ, no γ) for all $\gamma < \beta$, and an assumption that there is a 1–1 mapping, say φ, from a to its multiplicative class $a^{\times}$ of selectors from its pairs, assigning to every pair p the selector $\varphi(p)$. The member of the intersection of p and $\varphi(p)$ is denoted by p_1 and the other member of p by p_2. The selector which contains all the p_2 cannot be in the range of φ.

[16] Notice its resemblance to the proposition regarding s in the definition of $\aleph_0$.

obtained by induction". Hence the finite numbers in WR's system are called 'inductive', a term introduced by Poincaré (1906b p 308; cf. Fraenkel 1966 p 28). WR showed that an inductive number is different from its sequent (*3.1) while the addition of an element to Nc_fin does not change its cardinal number (*2.43), hence Nc_fin is non-inductive and reflexive, namely, equivalent to one of its proper subsets.[17] In order to generalize this result to every non-inductive class WR needed *2.6, that every infinite class has a subclass in $\aleph_0$, which requires AC. Interestingly, while in 1901 (1956 p 24), Russell was aware of the incompatibility of the two approaches to the definition of finite and infinite classes, his own and that of Dedekind, in 1902 he seems comfortable with his proof through *2.6 that the two definitions are equivalent.

[17] This result is achieved without AC. Compare to *Zahlen* # 65 where it is proved that a set with a single member is non-reflexive.

Chapter 16
The Role of CBT in Russell's Paradox

Russell tells in his 1903 "The principles of mathematics" (POM p 101) that he "was led to it [his paradox][1] in the endeavor to reconcile Cantor's proof that there can be no greatest cardinal number with the very plausible supposition that the class of all terms[2] ... has necessarily the greatest possible number of members".[3] Cantor's proof mentioned here is the proof of Cantor's Theorem (1892) which, Russell says (p 362), "is found to state that, if *u* be a class, the number of classes contained in *u* is greater than the number of terms of *u*". The class of all terms thus appeared to Russell as a refutation to Cantor's Theorem, so that Russell's Paradox was obtained within a Lakatosian proof-analysis.

Russell's refutation of Cantor's Theorem is mentioned in his draft of POM from November 1900 and his letter to Couturat of December that year (Coffa 1979 p 33ff; Moore 1995 p 230f). There it is claimed that since Cls contains Cls'Cls (the power-set of Cls), the power of Cls is the maximal power possible. In his 1901 paper (1917 p 69) Russell said with regard to U, that it has maximal power because

[1] On the chronology of Russell's Paradox see Anellis 1991. Regarding the good reasons to call this paradox the Zermelo-Russell paradox, see Ebbinghaus 2007 p 45 ff. Regarding Zermelo's preference of the term 'antinomy' to 'paradox' see *ibid*. footnote 62.

[2] 'Class' is Russell's term for Cantor's 'set' and is explicated in Chap. 6 of POM. 'Term' is a technical term with Russell, rather loosely defined to stand for anything that may occur in thought, language or reality ("can be counted as one") (POM p 43, 55); it corresponds to Dedekind's 'thing' (1963 p 44). The class of all terms we denote by U. Occasionally Russell juggles with other collections in place of U such as the class of all objects (Grattan-Guinness 2000 p 320) or the class of all propositions (POM p 367; cf. Russell 1919 p 136, Russell 1944 p 388). The class of all classes, which is clearly the class of all subclasses of U, we denote, following Whitehead 1902, by Cls.

[3] Russell describes here a clash of two gestalts: under one gestalt there is no maximal set; under the other there is such a set.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_16,

"there is nothing left to add".[4] In POM Russell proved by CBT that the power of U is equal to the power of Cls, which is its power-set. CBT is stated in POM (p 306) without proof as follows (where u, v are classes): "if there is a part of u which is similar to v, and a part of v which is similar to u, then u, v are similar". Russell refers to CBT as the Bernstein and Schröder Theorem (p 306) or the Theorem of Schröder and Bernstein (p 367). For its proof Russell referenced Borel 1898 and (p 306) Zermelo 1901.

16.1 Russell's Proof of Cantor's Theorem

Russell's formulation of Cantor's Theorem differs from Cantor's original in its use of subsets (or subclasses) instead of Cantor's characteristic functions (Ewald 1996 vol 2 p 920). The language of subsets and the notion of power-set, implicit in the above formulation of Cantor's Theorem, were alien to Cantor (Ferreirós 1999 p 264). The translation of Cantor's Theorem to the language of subsets may have begun in Borel 1898 (p 110) who pointed out that to each characteristic function correspond two complementary subsets of the function's domain. But Russell seems to have been the first to formulate Cantor's Theorem wholly in the language of subsets and to provide a proof of the translated version. Indeed, the translation is not noted in Schoenflies 1900 but appears fully articulated in Hessenberg 1906 (§24–25). We see here a Lakatosian shift in the dominant theory.

For the proof of his version of Cantor's Theorem Russell asserts (p 366 §347 second paragraph) the lemma that for a relation R, with equal domain and range, say u, the class "ω which consists of all terms of the domain which do not have the relation R to themselves" cannot be a relata[5] to a member of u. For if $y \in u$ has ω as its relata, then if yRy then $y \in \omega$ contrary to the assumption that ω contains only members of u that are not related to themselves by R[6]; if not yRy then y ought to belong to ω, which entails that yRy, another contradiction. This lemma provided Russell with a metaphor to prove Cantor's Theorem in its following formulation: There is no 1–1 mapping from u to its power-class. For otherwise, if φ is such a mapping, take R to be the relation: xRy iff $y \in \varphi(x)$, namely, y belongs to the relata of x iff y belongs to the image of x under φ. Then every φ-image of a member of u is the R-relata of that member. Since the above ω cannot be the R-relata of a member

[4] The name 'Cantor's paradox' is used for this contradiction (as well as for the contradiction regarding the class of all cardinal numbers). Moore-Garciadiego (1981 p 331; cf. Peckhaus 2002 p 5) reference POM p 367 as the place of origin of the name, but there, only the expression 'Cantor's argument' appears. The issue is clarified in Grattan-Guinness 2000 p 310f, 312f.

[5] The relata of x (in u) is the class of all members y of u such that xRy – p 24.

[6] It seems that if ω is empty the contradiction should emerge from $y \in \omega$.

of *u*, it cannot be the φ-image of any member of *u* and the theorem is proved. Obviously, Russell adapted to his language of relations, Cantor's original argument of his theorem.

In his argument Russell assumed that R is such that no two elements of *u* have the same relata. But this assumption seems unnecessary; Coffa (1979 p 35 footnote 10) took φ to be a function, thus not necessarily 1–1. Indeed, in this case, R is not 1–1; still, the above proof stands, and it gives a proof of a generalization of Cantor's Theorem: There is no many-1 mapping from *u* to its power-class. It is the generalized theorem that Russell used, perhaps without notice, in the derivation of his paradox.[7]

16.2 Derivation of Russell's Paradox

To prove that U has the same power as Cls Russell argued that "the number of classes should be the same as the number of terms" (POM #348 p 367). Russell contended that "classes will be only some among the terms" and "there is for every term, a class consisting of that term only". The first contention, which holds by the definition of 'term' – a notion that covers classes, means that the identity relation provides a similarity between Cls and a part of U.[8] By the second contention U is similar to a part of Cls – the class of all singletons in Cls, by relating to each term its singleton class. As the two contentions provide the conditions of CBT for Cls and U, these two classes are similar,[9] contrary to Cantor's Theorem.

Having refuted Cantor's Theorem Russell wanted to check closely what goes wrong with Cantor's argument in the case of U and Cls. He said (POM p 367 §349) that "it is instructive to examine in detail the application of Cantor's [diagonal] argument to such cases [as the class U] by means of an actual attempted correlation". So Russell needed the 1–1 mapping between U and Cls to test it under the diagonal argument. From a later perspective one expects that Russell would use the 1–1 mapping (or its relation counterpart) provided in the proof of CBT for U and Cls. However, contrary to prevailing belief (Banaschewski-Brümmer – 1986 p 1), the proofs of CBT published by that time (Borel 1898; Schoenflies 1900; Zermelo 1901) did not provide a 1–1 mapping between the two sets satisfying the conditions

[7] We remind the reader that Dedekind originally defined his chains for many-one mappings (see Sect. 9.1).

[8] Two classes are similar (POM p 113) when there is a "one-one relation whose domain is the one class and whose converse-domain [or 'range' in prevailing terminology] is the other class". 'Similar' corresponds to Cantor's 'equivalent' but it implies a 1–1 relation whereas equivalence implies a 1–1 mapping.

[9] We have here an example of an application of CBT to inconsistent sets.

of CBT! The first published proof where such a mapping was explicitly constructed was Peano's 1906 (see Chap. 20).[10] Had Russell access to this construction he could indeed have devised a 1–1 correlation from U onto Cls by correlating any term which is not a class to its singleton and this singleton to its singleton, and so on, and all other terms, which are necessarily classes, to themselves. Jourdain suggested this correlation, following his CBT proof, in which the resulting mapping is constructed (Jourdain 1907b p 359 footnote ‡, see Chap. 23). However, Russell was unaware of this construction so instead he patched from the two mappings between U and Cls (POM p 367 #349), a 2–1 mapping from U onto Cls:

> if x be not a class, let us correlate it with {x} [Russell's íx], i.e., the class whose only member is x, but if x be a class, let us correlate it with itself. (This is not a one-one, but a many-one correlation, for x and {x} are both correlated with {x}; but it will serve to illustrate the point in question.)

Russell's patched 2–1 mapping appears in the POM draft of November 1900 (cf. Russell 1993) where neither CBT nor the proof of Cantor's Theorem are mentioned. It is possible that Russell added the reference to CBT in POM both to explain the origin of the patched mapping and to warrant that a 1–1 mapping between U and Cls exists by a rigorous mathematical argument. Russell must have added in POM the proof of Cantor's Theorem to indicate the origin of the diagonal class. Russell continued to obtain his contradictory class:

> Then the class which, according to Cantor's argument, should be omitted from the correlation,[11] is the class ω of those classes which are not members of themselves[12]; yet this, being a class, should be correlated with itself. But ω [. . .] is a self-contradictory class, which both is and is not a member of itself.

Thus Cantor's neutral mathematical language of characteristic functions gave place to the emotionally loaded 'set that is not a member of itself'. No doubt – much of the popularity of Russell's Paradox is due to the existentialist associations its wording provokes. Another nice thing about Russell's Paradox is that the technical details of its derivation disappear once it is established, though U, or some similar class, is still tacitly assumed, for Russell's class collects its members from some realm.

[10] Dedekind's proof (see Sect. 9.2) contained such a construction, but it was published only in 1932. Zermelo's proof, predating Peano's paper, also contained a similar construction but in Poincaré's rendering of it (see Sect. 19.5), published in the same month when Peano's paper was published, the construction was not stated explicitly.

[11] Here Russell applies, tacitly or without notice, the generalized Cantor's theorem.

[12] Actually the class omitted is the class of all terms that do not belong to their image under the patched mapping. Since all the terms that are not classes belong under this mapping to their image, the omitted class is indeed composed of all classes that do not belong to their image, namely, to themselves.

16.3 The Crossly-Bunn Reconstruction

Crossley (1973 p 71) suggested that Russell's set is obtained immediately by Cantor's diagonal argument, when applied to the following setting: Let Cls be the collection of all sets and assume that it is a set; then its power-set P(Cls) exists (by the power-set axiom) and has all sets as its members; therefore Cls and P(Cls) are identical; then the diagonal set construction for the identity mapping between Cls and P(Cls) gives Russell's set.

The third contention of this argument is, however, mistaken because not every set (a member of Cls) is a subset of Cls: There are individuals in the universe that are not sets and sets of individuals are not subsets of Cls.

Bunn (1980 p 239) repeated Crossley's argument while avoiding her mistake by presenting Cantor's Theorem in a slightly generalized way: "For any one-one correspondence f whose domain of arguments is $A' \subseteq A$ and whose range of values is a subset [subclass] of P(A), the class $K = \{x \mid x \in A'. \, x \notin f(x)\}$ belongs to P(A) but not to the range of values of f". The conditions of this version of Cantor's Theorem imply the conditions of the version we provided above: f can be turned into a 2–1 function by arbitrarily giving values to the members of A-A′.[13]

Bunn then sets out from the universal class U and considers its subclass Cls of all classes. As Cls is P(U), f can be the identity from Cls to P(U). Cls is a subset of U, it is the A′ in the quoted passage. K in this case is Russell's set and we have $K = f(K)$, which implies that $K \in K$ iff $K \notin K$, so Russell's Paradox emerges. Bunn thus concluded: "It is by this or some very similar way that Russell discovered his antinomy". We prefer the explanation that rests on the 2–1 mapping mentioned explicitly by Russell, to the assumption that the domain of the mapping from U to Cls be limited to Cls, a gestalt switch that does not appear in Russell's writings.

Unfortunately, Bunn did not clarify the position of his argument in relation to that of Crossley; thus when Grattan-Guinness (1978 p 130) repeated Crossley's argument, referencing both Crossley and Bunn,[14] Coffa (1979 p 35 footnote 11) criticized him for Crossley's mistake.

16.4 Corroborating Lakatos

According to Lakatos' theory (1976), a refutation to a theorem should initiate proof-analysis of the proof of that theorem, which should discover a hidden lemma that is refuted by the said refutation.

Having obtained refutations to Cantor's Theorem, Russell set out, as if directed by Lakatos' theory of proof-analysis, to find a hidden lemma in Cantor's proof. he

[13] Move from the conditions of our version to those of Bunn's version requires AC.

[14] The latter is referenced as 1978.

says: "perhaps we shall find that his proofs[15] only apply to numbers of classes[16] not containing all individuals" and he speaks of "a concealed assumption" (Coffa 1979 p 34; Moore 1995 p 230). For a moment Russell even thought that he identified the hidden assumption to be that "there are classes contained in *u* which are not individuals of *u*" which is not fulfilled for U and Cls (Coffa 1979 p 33). He thus thought that the diagonal class produced by the 2–1 mapping refutes this assumption because it is a member of U. The search for the hidden lemma is mentioned also in POM. Russell says (POM p 362): "it would seem as though Cantor's proof must contain some assumption which is not verified in the case of such classes [as U and Cls]" and he continues (POM p 367 §349) that to find the hidden assumption "it is instructive to examine in detail the application of Cantor's [diagonal] argument to such cases [as the class U] by means of an actual attempted correlation". His paradox emerged only when he realized that the diagonal class cannot be in U and hence that not every propositional function defines a class. The hidden lemma was found in logic.

[15] Plural is used here because Russell did not rule out the possibility that the hidden lemma could be found in one of the lemmas that served to establish Cantor's theorem.

[16] Russell refers here to the formulation of Cantor's theorem stated at the beginning of this chapter.

Chapter 17
Jourdain's 1904 Generalization of *Grundlagen*

In 1904, Jourdain (see Grattan-Guinness 1977 prologue) published two papers. The first, in January (1904a), was titled "On the transfinite cardinal numbers of well-ordered aggregates"[1]; the second, in March (1904b), was titled "On the transfinite cardinal numbers of number-classes in general". The papers are remarkable because they matched Cantor's theory of inconsistent sets and offered a general construction of Cantor's scale of number-classes, both unpublished at the time (see Chap. 4).

Jourdain communicated his results to Cantor in a letter of October 29, 1903. Cantor responded on November 4, 1903.[2] Cantor told Jourdain that he had obtained similar results many years earlier, and mentioned, perhaps to avoid any priority disputes, that he communicated his results to Hilbert and Dedekind.[3]

Cantor encouraged Jourdain to publish his results but refused Jourdain's request to let him publish Cantor's letter (Grattan-Guinness 1971a p 117f). The reason was probably the same reason that prevented Cantor from publishing a third sequel to his 1895/7 *Beiträge* papers: he was not confident in the theory. In his 1904a paper, Jourdain told of his correspondence with Cantor, the latter's priority and encouragement (p 67 footnote †, p 70 footnote †).

[1] Jourdain usually translated Cantor's term *Menge* as 'aggregate'. For example in Cantor 1915. He made an exception in this paper and used 'manifold' for 'consistent aggregates' (see below). In this chapter, we will use both 'aggregate' and 'set', as well as 'class', interchangeably. 'Manifold' we will use when we wish to stress Jourdain's terminology. We might also use 'collection' for inconsistent sets and 'subset' for Jourdain's 'part'. We will use 'similar' interchangeably with Cantor's 'equivalent' and also with Russell's 'ordinally similar' hoping that the context will clarify our intention.

[2] Grattan-Guinness did not bring Jourdain's letter. From Cantor's answer it seems that Jourdain sent him only the content of his 1904a, not that of 1904b.

[3] We agree with Moore (1982 p 51) that the communication to Hilbert is the letter of June 28, 1899, with the attachment to Schoenflies, and the communication to Dedekind is the one of August 3, 1899.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_17, © Springer Basel 2013

An important point regarding Jourdain's papers is that while he was a keen observer his mathematical writings are often confused and his results in the 1904 papers were not rigorous. Evidence of Jourdain's confused style in presenting mathematical content we find in Cantor's remark that he could not understand an example sent to him by Jourdain though he immediately understood an example sent by Hardy (Grattan-Guinness 1971a pp 118, 119). In 1905, Harward published a paper that pointed out some of the confusions of Jourdain. We will review Harward's paper in the next chapter. To Harward's criticism Jourdain retorted in two articles, one, 1907b, with a rigorous, though far from concise, proof of CBT, and the other in 1908a with a rigorous proof of the Union Theorem, using CBT. However, in the meantime, the theorem was proved, using CBT as well, by Hessenberg (1906 Chapter XX; cf. Fraenkel 1966 p 219). We will cover these later papers of Jourdain in Chap. 23. Notwithstanding Harward's and our own criticism, Jourdain's papers are worth reading for through them we can better appreciate the difficulties that Cantor's ideas met when they were dissipated at the time among ordinary citizens of Republica Mathematica, upon which history is only rarely centered.

Jourdain's 1904 papers received little attention (Moore 1978 p 310) besides the attention of Harward. Perhaps because of its style and perhaps because in November 1904, Zermelo published his proof of the Well-Ordering Theorem, which introduced the axiom of choice. Jourdain used infinitely many arbitrary choices to obtain his theory of inconsistent sets, which cannot be gauged by the alephs but contain an image of the collection of all ordinals. But if AC is permitted, then all sets are well-ordered and hence have aleph as power. From this point of view, if for certain collections this conclusion raised a contradiction, it only proved that set theory is ill-based and not that inconsistent sets provide a "somewhat obscurely expressed solution of the Burali-Forti's (ordinal) contradiction" (Jourdain 1908a p 507 footnote **).[4]

Regrettably, also in historical circles Jourdain's reconstruction of Cantor's theory was not appreciated. But why complain about Jourdain's fate when Cantor's own theory of inconsistent sets never received its due attention. Even though it provides the context and explanation to central issues in Cantor's set theory such as comparability (see Chap. 4).

17.1 The Ordinals and the Alephs

Jourdain begins the 1904a paper with several remarks on the series of ordinal numbers and alephs (p 61f where *Grundlagen* endnote 2 is referenced – Ewald 1996 vol 2 p 916f). It appears that Jourdain took it for granted that his readers are

[4] Certain constructs in set theory did become unnecessary in view of AC. Jourdain raised such an argument against Bernstein's concept of 'multiple similarity' (Jourdain 1907b p 363 footnote *).

familiar with the basic notions of Cantorian set theory, so he did not introduce them systematically. Jourdain needed the ordinals and cardinal numbers early on because he used these notions to define the inconsistent aggregates, contrary to the approach of Cantor.

The first point to note is that Jourdain, quite correctly, considers the ordinal numbers of the 1895/7 *Beiträge* and the infinite numbers of *Grundlagen* to be the same. Thus while Jourdain states that the ordinal numbers belong to well-ordered aggregates (p 62), he holds the thesis that Cantor's two generation principles from *Grundlagen* generate all the ordinal numbers (1904b p 300).[5] The class of all ordinal numbers Jourdain denotes by W.

Similarly, Jourdain identifies the earlier 'power' (from Cantor's 1878 *Beitrag*) with the later 'cardinal number' (from Cantor's 1895 *Beiträge*). The alephs, central to the paper, he describes in three different ways: as the powers of "various classes of the transfinite numbers" (pp 61, 62); as "the cardinal numbers of well-ordered aggregates" (p 62); and every aleph as "the cardinal number of some segment of the series" of ordinal numbers (p 64).[6]

The first description suggests that the alephs are the cardinal numbers of the number-classes but Jourdain did not mention the number-classes explicitly in this context. In Cantor's theory only successor alephs are associated with number-classes; a limit aleph $\aleph_\delta$ is associated with the union of the number-classes for $\aleph_\gamma$, with $\gamma < \delta$ successor (see Sect. 2.1). It is possible that Jourdain missed this point and took this union to be the number-class associated with $\aleph_\delta$, for he said (1904b p 295): "According as the corresponding cardinal number has or has not an immediate predecessor, the class [corresponding to that cardinal number] is built up in one of two ways from the lower number-classes".[7] Jourdain's excuse may be that Cantor too did not express himself clearly on this issue, e.g., in endnote 2 of *Grundlagen*, where a correspondence between number-classes, their powers (the alephs) and the infinite numbers (the ordinals), is sketched.

The third description pertains to both the successor and the limit alephs. To have the first and third description coincide for successor cardinals one must prove the Limitation Theorem (see Sect. 2.2): that every number-class is equivalent to the union of it and all previous number-classes. Jourdain attempted to prove this theorem in 1904b, but the proof was a failure, mainly because his proof of the Union Theorem ($\aleph_\gamma{}^2 = \aleph_\gamma$) was worthless (see Sect. 17.6). It is interesting that Jourdain (with many others) regarded, at the time, the Limitation Principle as a comprehension principle for the number-classes (1904b p 300), and did not quite understand its link with the Limitation Theorem. In his 1905a paper (p 466f), however, Jourdain did identify correctly that the collection of all ordinals cannot

[5] Jourdain (1904b p 295 footnote *, 300) rejects Schoenflies view (1900 p 48) that new principles are necessary to generate limit ordinals for longer than ω successions.

[6] 'Segment' in the sense of an initial segment as in Cantor's 1897 *Beiträge* §13, to which Jourdain refers.

[7] Though Jourdain speaks here of 'cardinal numbers' he means the alephs.

have a sequent that fulfills the Limitation Principle and thus that no Burali-Forti contradiction arises, as was Cantor's view. Such sets, the inconsistent sets, he thus held, again with Cantor, to not have a cardinal or ordinal number.

The second description characterizes the alephs analogously to the characterization of the ordinal numbers as the order-types of well-ordered sets (p 62). The second description is compatible with the third: if an aleph is the cardinal number of a well-ordered set, it is the cardinal number of the segment of W defined by the ordinal of that set, as in the third description of the alephs.

Of the magnitude relations between the alephs, Jourdain, in the introduction to the paper, mentions three (p 62f): $\aleph_{\gamma+1}$ is the next greater cardinal number after $\aleph_\gamma$; $\aleph_\delta$, for δ a limit number, is the next greater cardinal number after all the cardinal numbers $\aleph_\gamma$, $\gamma < \delta$; $\aleph_0$ is the smallest cardinal number, namely, for every cardinal number $\mathfrak{a}$, $\mathfrak{a} \geq \aleph_0$, or, every set has a denumerable subset. For the first relation Jourdain referenced his continuation paper 1904b for the argument (see below). The second relation Jourdain took for granted. In fact, it follows from the definition of ω_δ[8] as the first ordinal after all ordinals in the number-classes with initial numbers ω_γ, $\gamma < \delta$, and of $\aleph_\delta$ as the cardinal number of ω_δ. A proof that $\aleph_\delta$ is different from all the $\aleph_\gamma$, for $\gamma < \delta$, is needed (see Sect. 1.3) but Jourdain was probably not aware of that. For the final relation Jourdain only referenced Cantor's 1878 *Beitrag*; in fact, this relation leans on (the denumerable) AC, a point to which Jourdain surely have become aware during his discussions with Russell in late 1905 (Grattan-Guinness 1977 §12).

17.2 The Power of the Continuum

Jourdain proved (1904a §1) that the power of the continuum $\mathfrak{c}$, is $\geq \aleph_1$, by arguing, taking the alephs to be the cardinal numbers of segments of W – not of number-classes, that "it is possible to take elements from the number-continuum corresponding to all the numbers of Cantor's first and second classes of ordinal numbers. For if the process were to stop we would have $\mathfrak{c} = \aleph_0$", contrary to Cantor's Theorem from 1874 *Eigenshaft* that $\mathfrak{c} > \aleph_0$. "Using, then, the theorem of Schröder and Bernstein, we can state that $\mathfrak{c} \geq \aleph_1$" because "$\aleph_1$ is the next cardinal number after $\aleph_0$".[9]

Jourdain then wanted to extend his result to conclude that $\mathfrak{c}$ is either an aleph or greater than all alephs. If $\mathfrak{c} \neq \aleph_1$ then $\mathfrak{c} > \aleph_1$ and so the correspondence he devised earlier can be continued through the third number-class, so $\mathfrak{c} \geq \aleph_2$. Continuing in this way, if $\mathfrak{c} > \aleph_v$ for all finite v, then, $\mathfrak{c} \geq \aleph_\omega$ (p 64). "And so on", Jourdain concluded, for all alephs. Jourdain was informally using for his proofs here

[8] For initial numbers Jourdain introduced the notation ω_γ (1904b p 295 footnote †) which he attributes to Russell (1903 p 322). In fact, Cantor used the same notation in his August 3, 1899, letter to Dedekind (see Chap. 3).

[9] The final quote is first in the original.

transfinite induction, a procedure that was not yet established at the time, though it was used by Cantor.

Jourdain referenced Hardy 1903 for the origin of the argument he used in this proof. He even claimed that his entire paper was influenced by Hardy's 1903 paper, so we have here an overt example of proof-processing. The proof clearly leans on the axiom of choice. Russell, in a letter to Jourdain from April 28, 1905 (Grattan-Guinness 1977 p 48) indicated that he thought Hardy's argument to be fallacious for "it would involve, if written out formally, the assumption of the existence of a multiplicative class". Of course, when Russell wrote his letter AC was already in the open but Jourdain's argument predated Zermelo 1904, so we cannot blame him of using the axiom unintentionally. In his 1907b paper (p 363), Jourdain admitted his guilt.

For his proof that $\mathfrak{c} \geq \aleph_1$, Jourdain invoked CBT. His reason seems to be the following: Cantor defined (1895 *Beiträge* §2) the relation $\mathfrak{a} > \mathfrak{b}$, when $\mathfrak{a}$ is the cardinal number of M and $\mathfrak{b}$ the cardinal number of N, to hold when: (1) No part of N is equivalent to M, (2) there is a part of M equivalent to N. The relation $\mathfrak{a} = \mathfrak{b}$ Cantor defined to hold when M and N are equivalent. Thus, the relation $\mathfrak{a} \geq \mathfrak{b}$, which Cantor never used, would naturally stand for $\mathfrak{a} > \mathfrak{b}$ or $\mathfrak{a} = \mathfrak{b}$. But when Jourdain said that $\mathfrak{c} \geq \aleph_1$ Jourdain meant another definition for $\mathfrak{m} \geq \mathfrak{n}$, namely: 'N is equivalent to some part of M'. Jourdain explicated all this in his 1907b paper (p 355f).[10] Since, with regard to $\mathfrak{c}$, Jourdain only proved that the continuum contains a set of cardinal number $\aleph_1$, there are two possibilities: that a set of cardinal number $\aleph_1$ contains a set equivalent to the continuum, or not. In the second case we have by Cantor's definition of $>$ that $\mathfrak{c} > \aleph_1$; but in the first case, if we follow Cantor's definition, we cannot conclude that $\mathfrak{c} = \aleph_1$ without CBT (cf. p 69). For this reason, in his 1907b paper, Jourdain formulated CBT thus: If a part of M is similar to N, $\mathfrak{m} \geq \mathfrak{n}$.[11] Note that to obtain $\mathfrak{c} > \aleph_1$, the Comparability Theorem for cardinal numbers is not used.

17.3 Inconsistent Aggregates

In analogy to the result for $\mathfrak{c}$, Jourdain concluded (p 64) that "every cardinal number is either contained in the series of alephs or is greater than all alephs". This is the 'Inconsistency Lemma' (see Chap. 4). A cardinal number greater than all alephs

[10] There he attributed this definition to several mathematicians, including Bernstein, Whitehead, Kőnig and Peano. Jourdain gave no references. We suggest the following: Bernstein 1905 p 127ff or Bernstein 1904 p 559, Whitehead 1902 p 380 *3.13 (it is the second *3.12 in the original), Peano 1906 p 360. With regard to Kőnig, probably J. Kőnig, we have not located a reference. Cf. Kuratowski-Mostowski 1968 p 185.

[11] Cf. Sect. 13.1 for Zermelo's Theorem IV, which is a version of CBT using $\geq$. Note that using the Cantorian meaning of $\geq$ and Dedekind's definition of 'infinite', the following formulation of CBT is trivial (Jourdain 1907b p 356): $\mathfrak{a} \geq \mathfrak{b}$ and $\mathfrak{a} \leq \mathfrak{b}$ entails $\mathfrak{a} = \mathfrak{b}$, because under our assumptions, $\mathfrak{a} > \mathfrak{b}$, $\mathfrak{b} > \mathfrak{a}$, $\mathfrak{a} = \mathfrak{b}$ are exclusive of each other (Mańka-Wojciechowska 1984 p 191).

must be greater or equal to the cardinal number of the series of all ordinals, Jourdain concluded (p 64), because the cardinal numbers of the segments of this series are the alephs.

But then the paradox of the greatest ordinal is encountered (p 64f). Jourdain references Burali-Forti for this paradox, which he presents in the following form: If β is the ordinal of W it ought to be the greatest ordinal, but $\beta + 1$ is greater. Jourdain considers then two arguments suggested in the literature to settle the paradox: one, by Burali-Forti, who denied Cantor's trichotomy law for ordinals; the other, by Russell, who denied that the collection of all ordinals is well-ordered. Jourdain rejected both arguments (p 65): The first on the basis that Cantor's proof of the trichotomy law for ordinals "is beyond all possible objections", and the second by the lemma that for a collection to be well-ordered it is enough to show that it is simply ordered and does not have a part of order-type $*\omega$ ($\omega*$ in current notation). He easily demonstrated (p 65f) that both requirements hold for the collection of all ordinals. He did convince Russell with his proof (Grattan-Guinness 1977 p 48). Jourdain noted that the proof that the above lemma entails that the collection of all ordinals is well-ordered, was not mentioned in earlier publications. This proof uses dependent choices (Grattan-Guinness 1977 p 27).

Here a question arises: why is the existence of a cardinal number greater than the cardinal number of W necessary to the emergence of the Burali-Forti paradox? Jourdain did not address the question explicitly but from what follows it appears that Jourdain did not accept unlimited comprehension. He therefore needed an argument that the collection W is a class to be able to talk about its ordinal number and generate the Burali-Forti paradox. Jourdain's rejection of unlimited comprehension comes out also in a letter to Russell of May 24, 1904 (Grattan-Guinness 1977 p 34) where he asks: "Is 'w sim u' (for constant u) always a function which determines a class?" This class is for Russell the cardinal of u. Since it contains a class for each α in W (the class of pairs $< x, \alpha >$ where $x \in u$) it is inconsistent according to Jourdain.[12] Now, if a set exists with cardinal number greater than all alephs, then it has a subset similar to W and the Burali-Forti paradox rolls out.

Faced with the Burali-Forti paradox, Jourdain, determined to substantiate Cantor's *Grundlagen* doctrine (p 64), decided that "certain aggregates have no cardinal number and no ordinal type" (p 66f). He called such aggregates 'inconsistent aggregates'; for the other aggregates, the consistent ones, he reserved the name 'manifold' [*Mannigfeltigkeits*], thus reviving a name which Cantor used for *Menge* during the 1870s, 1880s. Under this doctrine, every aggregate either has a cardinal number which is an aleph (in which case it can be well-ordered) or no cardinal number at all (p 67), in which case it contains an image of W. Jourdain quoted Cantor (*Grundlagen* Ewald 1996 vol 2 p 916 [1], we quote from Ewald) who said "by a 'manifold' or 'set' I understand every multiplicity which can be thought of as one". Jourdain echoed this statement by saying "In conformity with this view . . . an

[12] Russell was aware of the problem (Grattan-Guinness 1977 p 49).

inconsistent aggregate [is] an aggregate of which it is impossible to think as a whole without contradiction". Unknowingly Jourdain expressed here almost literally Cantor's opinion in his letter to Dedekind of August 3, 1899.

Feeling perhaps uneasy in relying for his definition of inconsistent aggregates on Cantor's notion of "can be thought of as one", Jourdain added "for formal purposes, . . . an inconsistent aggregate is an aggregate such that there is a part of it which is equivalent to W". But Jourdain was not happy also with this definition (p 67), because it entails that the notion of consistent set, which is essential to the definition of cardinal and ordinal numbers, relies on the notion W, which makes the definition of W circular.[13] To by-pass this problem Jourdain suggested to use for the formal characterization of an inconsistent set, instead of W, the aggregate "of which every well-ordered aggregate is a segment". In a letter to Russell of March 17, 1904, (Grattan-Guinness 1977 p 27) Jourdain denoted this aggregate by $\mathfrak{B}$[14] and said that "M is a well-ordered series implies that M is ordinally similar to a segment of $\mathfrak{B}$ or to $\mathfrak{B}$ itself". Grattan-Guinness noted (1977 p 34 footnote 2) that in two offprints of 1904a Jourdain similarly corrected the text.

In the paper, Jourdain referenced Schoenflies 1900 (pp 36, 40, 41) for the origin of the idea for $\mathfrak{B}$. Grattan-Guinness repeated this reference (1977 p 27 footnote 3). The following seems, however, to describe more accurately the role of Schoenflies in the emergence of $\mathfrak{B}$: On page 36 Schoenflies noted that all well-ordered sets are of one type, because they are comparable, and thus can be conceived as segments of one well-ordered set. On page 40 Schoenflies repeated this view denoting this set by W. On page 41 he defined W to be the "collection of all ordinal numbers". Because of the circularity noted, Jourdain obviously wanted to keep separated, Schoenflies' two conceptions.

Jourdain may have had another reason to use $\mathfrak{B}$: If one takes W to be the collection of all ordinals defined under the Russell (Frege-like) definition, then, as argued above, the elements of W are inconsistent sets and the segments of W cannot be aggregates. However, there was a weakness also in using $\mathfrak{B}$: no clear definition of $\mathfrak{B}$ is given, say by some propositional function (cf. Grattan-Guinness 1977 p 29). This issue can be settled if we maintain the infinite numbers defined by Cantor's generation principles (under the Limitation Principle) with the ordinal numbers defined by abstraction from well-ordered sets. The infinite numbers form $\mathfrak{B}$ and the segments of $\mathfrak{B}$ represent all well-ordered sets and their ordinals. W is the collection of all ordinals. No circularity arises when this approach is used, which shows the advantage of a representation definition of ordinals over Russell's definition.

Interestingly, Jourdain planned to publish a paper where the representation approach to the definition of cardinal numbers would be propagated (1907b p 355

[13] Harward (1905 p 459) pointed at the circularity of Jourdain's formal definition of W. Jourdain retorted (1907b p 356 footnote * last paragraph) that he did not make this "glaringly vicious circle". He may have meant his covert reference to $\mathfrak{B}$ (see the text).

[14] $\mathfrak{B}$ was apparently mentioned in Jourdain's letter to Cantor for Cantor related to it. On $\mathfrak{B}$ cf. Jourdain 1905b.

footnote †). Under this approach, Jourdain believed, there will be no inconsistent sets. The paper was never published (Grattan-Guinness 1977 pp 66f, 85ff) and we have no clue as to what was Jourdain's idea.

17.4 The Corollaries

With his institution of the inconsistent aggregates, Jourdain obtained the Comparability Theorem for cardinal numbers from the trichotomy of the ordinals. So Jourdain next addressed corollaries B-E of that theorem. With regard to Corollary B, which is CBT, Jourdain could not agree with Cantor to present it as derived from the Comparability Theorem for cardinal numbers because he used it in his proof of the Inconsistency Lemma from which he obtained the theorem. Cantor, in his letter to Jourdain, touched on this point when he said that he did not use CBT in his derivation of comparability. So, Jourdain produced an alternative proof of the Inconsistency Lemma (p 70): Beginning with a set C, he successively (implicitly by transfinite induction) selected elements from it corresponding to the members of W. He made no stops at the various alephs, as he did in his proof regarding the power of the continuum. Thus, he avoided mention of CBT. If C is exhausted, it is well-ordered and its aleph is obtained. Otherwise, C contains an image of W and is inconsistent. In this proof Jourdain repressed his problem with the circularity of W, or used tacitly $\mathfrak{B}$.

Not only was CBT eliminated by the above procedure, Jourdain even added: "With this formulation we obtain a new proof of the theorem proved by Schröder and Bernstein". He must have had in his mind the proof via the enumeration-by method for well-ordered sets, which we mentioned in Sect. 2.2. This proof holds only for consistent sets and indeed, Jourdain stated corollary B only for 'manifold', not 'aggregate' (p 68). Jourdain forgot that CBT was already proven for any set.

Jourdain's new proof of the Inconsistency Lemma is not tenable because how do you know that C is not exhausted by (I) + (II)? You must enter cardinal considerations: you must prove that C is non-denumerable and then that the ordinals can be continued past (II), which requires the full mechanism employed for the construction of the scale of number-classes.

Apparently Jourdain had not noticed that Cantor's corollaries C, E of the Comparability Theorem are equivalent to B, for he said (p 69) that while B was proved directly by Schröder and Bernstein, C, E follow from A (the Comparability Theorem for cardinal numbers). Strangely, he added that thus C, E are "proved independently of the theorem that every cardinal number is an aleph", even though A is proved by him from this theorem.

With regard to Cantor's Corollary D, Jourdain said that it is the only corollary among B-E that can be deduced without circularity from his theorem that every cardinal number is an aleph. It is not clear what Jourdain had in mind when he made these remarks. Anyway, Jourdain did notice the equivalence of D and case (IV) of the scheme of complete disjunction, which he stated in the following form (p 69): "It is impossible that neither a part of M should be equivalent to N nor a part of N should be equivalent to M, provided that both M and N are transfinite". His statement, however, raises unnecessarily the question of what a transfinite set is, which is neither linked to Corollary D nor to case (IV).

Jourdain did not use the name 'scheme of complete disjunction'. In his 1907b paper (p 352) Jourdain brought the scheme (with a slight change in the order of the cases) saying that "[it] leads us to consider all the vital questions in the modern theory of aggregates". Unlike Cantor, Borel and Schoenflies, Jourdain did not demand the subset to be a proper subset (1907b p 353, especially footnote §, 1904a p 69 footnote §). His reason was that under his approach the case that corresponds to CBT holds also for finite sets and the case that corresponds to (IV) does not hold for any set. With this change, Jourdain introduced to the scheme the relation $\leq$, which he himself saw as based on CBT (see Sect. 17.2). Therewith Jourdain destroyed the logical essence of the scheme.

17.5 Jourdain's Rendering of Zermelo's 1901 CBT Proof

After his discussion of the Comparability Theorem for cardinal numbers and its corollaries, Jourdain said (p 70): "The problem as to the relations of magnitudes of any two cardinal numbers is thus completely solved by the consideration of the well-ordered aggregate of cardinal numbers of well-ordered manifolds. This entry of ordinal notions has not appeared satisfactory to Schröder, since the question is one of the elementary properties of cardinal numbers".[15] Jourdain raised here an objection, which he attributes to Schröder, whereby CBT, being a theorem about cardinal numbers, cannot be proved by way of the ordinals. Jourdain direct proof indeed rested on the ordinals (see the previous section). Jourdain partially agreed with this criticism, namely, he agreed that the proof should avoid transfinite ordinals, but thought that the natural numbers are unavoidable in a proof of CBT. In this view Jourdain was a forerunner of Poincaré (see Chap. 19), who held a similar view though on different grounds. Jourdain declared that Schröder's own proof of CBT was marred by the same defect of using ordinals – the natural numbers. As he could not demonstrate his claim on Schröder's proof directly, because Schröder's proof is illegible and erroneous (not that Jourdain made this point), he turned (§8) to the proof that Zermelo provided in his 1901 paper (see Chap. 13), because it also makes use of natural numbers. We will now reproduce in detail his attempt to present that proof.

Denoting the cardinal numbers of M_1, M-M_1, N_1, N-N_1, respectively, by $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{d}$, $\mathfrak{e}$, Jourdain stated CBT as follows (p 70f): If $\mathfrak{a} = \mathfrak{d} + \mathfrak{e}$ and $\mathfrak{d} = \mathfrak{a} + \mathfrak{b}$, then $\mathfrak{a} + \mathfrak{b} = \mathfrak{d} + \mathfrak{e}$. By substitution $\mathfrak{a}=\mathfrak{d} + \mathfrak{e} = (\mathfrak{a} + \mathfrak{b}) + \mathfrak{e} = \mathfrak{a} + (\mathfrak{b} + \mathfrak{e})$. Hence, by re-substituting $\mathfrak{a}$ on the right side ν times, ν finite, we get $\mathfrak{a} = \mathfrak{a} + \nu(\mathfrak{b} + \mathfrak{e})$. If we have the lemma that this entails $\mathfrak{a} = \mathfrak{a} + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e})$ then we can say that

[15] Jourdain references Schröder 1898 without giving a page number and we could not locate where Schröder's view is expressed.

$\mathfrak{a} = \mathfrak{a} + \aleph_0 \cdot \mathfrak{b} + \aleph_0 \cdot \mathfrak{e} = \mathfrak{a} + \aleph_0 \cdot \mathfrak{b} + (\aleph_0 + 1) \cdot \mathfrak{e} = \mathfrak{a} + \aleph_0 \cdot \mathfrak{b} + \aleph_0 \cdot \mathfrak{e} + \mathfrak{e} = \mathfrak{a} + \mathfrak{e}$, and similarly $\mathfrak{a} = \mathfrak{a} + \aleph_0 \cdot \mathfrak{b} + \aleph_0 \cdot \mathfrak{e} = \mathfrak{a} + (\aleph_0 + 1) \cdot \mathfrak{b} + \aleph_0 \cdot \mathfrak{e} = \mathfrak{a} + \aleph_0 \cdot \mathfrak{b} + \aleph_0 \cdot \mathfrak{e} + \mathfrak{b} = \mathfrak{a} + \mathfrak{b}$, so that $\mathfrak{a} = \mathfrak{d}$ and hence $\mathfrak{a} + \mathfrak{b} = \mathfrak{d} + \mathfrak{e}$ as required.

Jourdain's poor style of presenting mathematics is evident here: The thesis is simply $\mathfrak{a} = \mathfrak{d}$ and the derivation of $\mathfrak{a} = \mathfrak{a} + \mathfrak{e}$ is redundant.

The first part of the proof resembles Zermelo's proof in its reliance on properties of $\aleph_0$ and on cardinal arithmetic. Jourdain was correct to note that the first part is purely cardinal and so Schröder's criticism does not apply to it. The ordinal part comes in the proof of $\mathfrak{a} = \mathfrak{a} + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e})$. The passage from the equations $\mathfrak{a} = \mathfrak{a} + \nu \cdot (\mathfrak{b} + \mathfrak{e})$ to the equation $\mathfrak{a} = \mathfrak{a} + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e})$, Jourdain calls the conclusion "from $\{\nu\}$ to $\aleph_0$" and describes it as the "extended principle of cardinal induction" (p 71 footnote).[16] That this principle is not always valid Jourdain demonstrated by noting the counterexample that $\aleph_0^{\nu} = \aleph_0$ while $\aleph_0^{\aleph_0} = 2^{\aleph_0} > \aleph_0$.[17] He then claimed that he had found that this principle applies to the case at hand in October 1902, independently of Zermelo.[18]

For the proof of the principle in the present case, Jourdain argues as follows: Because $\mathfrak{a} = \mathfrak{a} + (\mathfrak{b} + \mathfrak{e})$, from M_1 a manifold P_1 can be removed with cardinal number $\mathfrak{b} + \mathfrak{e}$, while the remainder M_2 still has cardinal number $\mathfrak{a}$. The process can be repeated without end to generate the sequence P_ν of pairwise disjoint manifolds and the manifold M_ω which he describes by saying that it is the "first (in the above process) of all manifolds which is not contained in all the manifolds P_ν". For the definition of M_ω Jourdain references Schoenflies 1900 p 14 where M_ω is defined naturally as the intersection of the M_ν (see Sect. 12.1). As each P_ν is of cardinal number $\mathfrak{b} + \mathfrak{e}$, then, denoting by $\mathfrak{g}$ the cardinal number of M_ω, we have: $\mathfrak{a} = \mathfrak{g} + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e}) = \mathfrak{g} + 2 \cdot \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e}) = \mathfrak{g} + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e}) + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e}) = \mathfrak{a} + \aleph_0 \cdot (\mathfrak{b} + \mathfrak{e})$ which is the required lemma. Jourdain employed here the reemergence argument (see Chaps. 13 and 14).

Clearly Jourdain relied on no new principle of induction in this proof and he proved no such principle. What Jourdain did here was a proof of the lemma required in the CBT proof given above without appeal to ordinals other than the natural numbers. Jourdain stresses that "whatever may be the cardinal numbers of M and N, the proof requires only an enumerable manifold (of type $\omega + 1$) of steps".[19]

---

[16] In the same footnote Jourdain compares this procedure to transfinite induction, "ordinal induction" he calls it or "the conclusion from $\{\nu\}$ to ω". He references Schoenflies 1900 (pp 45, 52, 60, 67) for examples where this procedure is used. Jourdain's differentiation between cardinal and ordinal induction seems superficial.

[17] The counterexample against the general validity of the procedure applied is entirely superfluous, as it comes out of the next paragraphs.

[18] Incidentally, this was the month in which Whitehead's 1902 paper was published, with an erroneous proof of CBT by Russell (see Chap. 15). We can speculate that Jourdain found the argument he mentioned when he tried to proof-process Russell's CBT proof. Jourdain clearly studied Whitehead 1902 carefully because he attempted to prove two open problems stated in that paper (see the next section).

[19] In fact, the proof requires the natural numbers and properties of $\aleph_0$ (not the type $\omega + 1$).

Jourdain's remark serves to soften Schröder's criticism if not to annul it entirely; one can hardly say that the status of the natural numbers and of the transfinite ordinals is the same.

The proof is similar to part of the proof of Zermelo's Denumerable Addition Theorem. Harward (1905 p 455ff Note A) criticized Jourdain's proof on grounds that one could not infer from a finite process that goes on without end that the entire process is complete and all the P_ν are given so that M_ω can be defined. Harward's criticism is, however, incorrect, though his suggestion to take a 1–1 mapping φ between M_1 and M_2 is advisable because then all the P_ν are given at once and AC is avoided. This is indeed how Zermelo approached the scene. Jourdain pointed out this fact when he accepted, in half voice, Harward's suggestion in his 1907b paper (p 356 footnote *). Because of its resemblance to Zermelo's proof, we do not compare Jourdain's 1904a CBT proof to earlier proofs. We nevertheless do attach to it a metaphor: 'conclusion from ν to $\aleph_0$'.

It is not known if Jourdain's worry of Schröder's criticism had any influence on Zermelo,[20] but in 1905 Zermelo devised a new proof, which avoids the notion of natural numbers and complete induction. Zermelo communicated the proof (June 1905) to Hilbert (Ebbinghaus 2007 p 89). As a result, when Poincaré, in his 1905/6 papers debating logicism, challenged Couturat to produce a proof of CBT making no use of natural numbers and complete induction, Zermelo could respond immediately. We will come back to Zermelo's second proof in later chapters.

17.6 The Sum and Union Theorems

Jourdain (p 72) references Whitehead 1902 (p 368) as the place where two open problems are stated, which he sets out to prove. These two problems are: If $\mathfrak{a}$, $\mathfrak{b}$ are two cardinals, $\mathfrak{a}$ infinite and $\mathfrak{a} \geq \mathfrak{b}$, then $\mathfrak{a} = \mathfrak{a} + \mathfrak{b}$ and $\mathfrak{a} = \mathfrak{a}\mathfrak{b}$.

It is easy to see that if $\mathfrak{a} = \mathfrak{a} + \mathfrak{a}$[21] then for every $\mathfrak{b} \leq \mathfrak{a}$, $\mathfrak{a} = \mathfrak{a} + \mathfrak{b}$, because $\mathfrak{a} \leq \mathfrak{a} + \mathfrak{b} \leq \mathfrak{a} + \mathfrak{a} = \mathfrak{a}$ and by CBT the result follows.[22] Similarly, if $\mathfrak{a}^2 = \mathfrak{a}$[23] and $\mathfrak{b} \geq 1$, we have $\mathfrak{a} \leq \mathfrak{a}\mathfrak{b} \leq \mathfrak{a}^2 = \mathfrak{a}$ so that by CBT $\mathfrak{a} = \mathfrak{a}\mathfrak{b}$. It is easy to see that $\mathfrak{a} = \mathfrak{a} + \mathfrak{a}$ for all $\mathfrak{a}$ such that $\mathfrak{a} = \aleph_0\mathfrak{b}$ for some $\mathfrak{b}$[24] (Whitehead 1902 p 381f *4.3, *4.31, *4.38). Similarly $\mathfrak{a}^2 = \mathfrak{a}$ for all $\mathfrak{a}$ such that $\mathfrak{a} = \mathfrak{b}^{\mathfrak{c}}$ because for $\mathfrak{c}$ (the power of the continuum) we have that $\mathfrak{c} = \mathfrak{c} + \mathfrak{c}$ (Whitehead 1902 p 393f *32.2).[25] Zermelo showed that the set of all $\mathfrak{b}$ such that $\mathfrak{a} = \mathfrak{a} + \mathfrak{b}$ has certain closure properties. Jourdain attempted to extend Zermelo's closure conditions to include the case $\mathfrak{b} = \mathfrak{a}$ (1904b p 301).

[20] Prior to his 1908a paper Jourdain did communicate with Zermelo (see below).

[21] This result, for $\mathfrak{a}$ an aleph, follows from the Sum Lemma (see Sect. 2.2).

[22] Russell used the mentioned inequality in his proof of CBT (see Sect. 15.1).

[23] This result, for a an aleph, follows from the Union Lemma (see Sect. 2.2).

[24] Without loss of generality we can take $\mathfrak{b} = \mathfrak{a}$ (1904a p 73).

[25] Jourdain explicitly referenced the pages mentioned here from Whitehead 1902.

To prove that $\mathfrak{a} = \mathfrak{a} + \mathfrak{a}$, Jourdain (p 73) took a manifold M of cardinal number $\mathfrak{a}$, it is composed of denumerable partitions with perhaps one finite partition. This assertion, which is correct when $\mathfrak{a}$ is an aleph and which amounts to the theorem that every ordinal is divisible by ω with a finite residue, Jourdain stated without any justification; see Fraenkel 1966 p 210f for its proof. He then suggested that each element of the set be replaced by two new elements (each different from any other). We get then on the one hand the possibility of splitting the set into two sets each equivalent to the original set, and on the other hand that the set's power does not change because the denumerable partitions remain denumerable partitions and the finite partition can be embedded in one of the denumerable partitions. But, as Jourdain's inconsistent aggregates theory is built by the axiom of choice, which also entails *4.3 from Whitehead's 1902 (see Chap. 15), Jourdain offered nothing to the results already obtained there.

Early on in the paper (p 62) Jourdain mentioned Cantor's *Grundlagen* (§12, 13) proof that "$\aleph_1$ is the next greater cardinal number to $\aleph_0$". There Jourdain stated a generalization of this theorem, namely, that for every ordinal γ, $\aleph_{\gamma+1}$ is the next cardinal number after $\aleph_\gamma$. In 1904b he attempted to prove this generalized theorem by generalizing Cantor's proof of *Grundlagen* (compare Chap. 1). It seems that Jourdain regarded this result as the main purpose of his paper (1904a pp 62, 68 footnote §, 74f) because with it the series of alephs is established with the scale of number-classes. It is to Jourdain's credit that he apprehended, and was the first to apprehend, that the Next-Aleph Theorem is essential to the construction of the scale of number-classes, which is the backbone of Cantor's theory of transfinite numbers. However, Jourdain's attempt was a failure on several counts, as we will demonstrate now.

Jourdain began (1904b p 295f) with a the proof for the case of $\aleph_2$. He said that the proof consists of two steps: that every subset of the third number class is either finite or denumerable or of the power $\aleph_1$ or of the whole third number-class, and that $\aleph_2$ is different from $\aleph_1$. Cantor had the two steps in reverse order. For his first step, which Cantor called the Fundamental Theorem (*Grundlagen* §13, see Sect. 1.2), Jourdain gave the following argument:

> Denoting by ω_1 the first number of the third number-class, the whole class is formed as follows: – First after ω_1 comes the series represented by $\{\omega_1 + \alpha\}$, where α takes, in order of magnitude, the values of all the numbers of the first and second number-classes. Next after this series comes the number $\omega_1 + \omega_1 = \omega_1 \cdot 2$, which is followed by the series $\{\omega_1 \cdot 2 + \alpha\}$, and so on. It is evident that the numbers of the first and second number-classes, arranged in order of magnitude, are similar to the series $\{\omega_1 + \alpha\}$.
>
> We may then conclude, in a precisely similar way to that which Cantor has proved that any part of the second number-class has a cardinal number which is either finite, or $\aleph_0$, or that of the second number-class, that any part of the third class has a cardinal number which is either finite, or $\aleph_0$, or $\aleph_1$, or that of the whole third class[26]

[26] Similarly pointless observations, but intuitively necessary and naively satisfying, regarding the numbers in general number-classes, Jourdain brought in §3 of 1904b.

The first paragraph quoted is clearly irrelevant to the discussion. The final sentence of this paragraph is a triviality that has no import on the proof proposed. The preceding part of the paragraph is a careless attempt to copy Cantor's way of populating the second number-class. The paragraph as whole suggests no advance to a proof of the required theorem. The second paragraph quoted is a joke: it tells us that we should copy Cantor's proof of the Fundamental Theorem to the present context, which is actually what we expected Jourdain to do.

For the second step of the proof, which is a proof of the Sequent Lemma (see Sects. 1.2 and 2.2), Jourdain (1904b p 296f) followed in the footsteps of Cantor (*Grundlagen* §12) until he broke loose: If $\aleph_2$ is equal to $\aleph_1$ then (III) can be arranged in a sequence of type ω_1. From this sequence Jourdain extracted a monotonic increasing subsequence, which he then complemented into a succession (see Sect. 1.1) by adding to it all the numbers preceding or in between the members of the subsequence. At this point Jourdain leaps to the conclusion that if it is proved that ${\aleph_1}^2 = \aleph_1$ (the Union Theorem for $\aleph_1$) the proof that $\aleph_2$ is different from $\aleph_1$ is complete.

Jourdain failed to explicate several issues here: that the succession contains all the members of the original sequence of the members of (III); that the succession must have a sequent, either because it is well-ordered or by Cantor's second generation principle; that the subsequence partitions the ω_1 enumeration of (III) into at most ω_1 partitions each with at most ω_1 members; that by the Union Theorem, and CBT for $\aleph_1$,[27] the sequent is of power $\aleph_1$; that hence the sequent must belong to the succession – a contradiction, which entails that the assumption that (III) can be enumerated by $\aleph_1$ must be rejected. Note that for the application of the Union Theorem here, the axiom of choice is invoked, because a mapping has to be chosen from each partition to ω_1.

With regard to the proof of the Union Theorem ${\aleph_1}^2 = \aleph_1$, Jourdain (1904b 298f) envisaged a table with ω_1 rows and ω_1 columns, the elements of which he designated by $u_{\alpha,\beta}$.[28] This table is composed of sub-tables each of ω rows and ω columns. For these tables we have Cantor's Theorem that the table can be enumerated by ω. Jourdain merely hand-waived how to each sub-table a different section of ω_1, equivalent to ω, is allocated and concluded that the $\omega_1 \times \omega_1$ table can also be enumerated by ω_1. The proof cannot even be rendered rigorous. But Jourdain felt that from this proof it can be concluded that ${\aleph_v}^2 = \aleph_v$ for every finite v (p 300). The passage to ${\aleph_\omega}^2 = \aleph_\omega$ he then described as "obvious" and hence he concluded "we are certain that the equation $\aleph_\gamma \aleph_\gamma = \aleph_\gamma$, where γ is any ordinal number, holds". Under the inconsistent sets doctrine Jourdain declared that thus for every cardinal number $\mathfrak{p}$, $\mathfrak{p}^2 = \mathfrak{p}$. Jourdain did not bother to go back and state that with the proof of the Union Theorem, the proof that $\aleph_{\gamma+1}$ is the next cardinal

[27] Cantor too had not stressed that his proof of the Sequent Lemma, required CBT.

[28] u is here a dummy that fixes the order of α, β. Apparently, ordered-pairs were not yet invented in 1904. Cantor used indexed dummy with the same purpose in letters to Dedekind of the late 1873 (see Sect. 7.3). In Jourdain's 1908a, ordered-pairs are already used.

number after $\aleph_\gamma$ is also complete. For Harward's criticism of Jourdain proof and Harward's own proofs see Sect. 18.5.

17.7 Comparison with Cantor's Theory

Jourdain developed independently of Cantor, the theory of inconsistent sets, and even came up with the same term. When asked by Cantor where did he obtain the term 'inconsistent aggregate', he referenced Schröder, who used it for sets that have incompatible elements (Grattan-Guinness 1977 p 28f). In addition, Jourdain realized the centrality of the Union Theorem for the construction of Cantor's scale of alephs. Therewith Jourdain identified that Cantor intended his Comparability Theorem for cardinal numbers to be reduced to the comparability of the alephs. On several issues, however, there are differences between Jourdain and Cantor.

Cantor did not use choice, as did Jourdain, in proving the Inconsistency Lemma, but Corollary D (see Chap. 4). Cantor thus knew that a new postulate was necessary to establish the theory of inconsistent sets, while Jourdain was not aware of this point. Cantor was aware of the need to lay down segregation rules to segregate the consistent sets from the inconsistent ones. Jourdain used these rules without notice (p 67). Harward filled the gap for Jourdain on this point (see below). Cantor avoided the circularity, which worried Jourdain, regarding W and the definition of an inconsistent set. Cantor defined an inconsistent set as a set that the assumption of its oneness leads to a contradiction. Though this definition seems non-mathematical and Jourdain, with good sense, deliberated on it, it was indeed adopted even in formal set theory: see Levy 1979 (I.3) proof that Russell's class is not a set.[29]

Cantor proved that W is inconsistent and, using Corollary D, that a set which is not equivalent to any aleph contains W and is therefore also inconsistent. Case (IV) was likewise settled for Cantor by postulating Corollary D, so that it is obtained directly, not by way of the Comparability Theorem for cardinal numbers, without mention of inconsistent sets. Jourdain had perceived the relation between Corollary D and case (IV), an observation which (it too) was missed by most scholars of Cantor, but not to its full extent: He missed noticing that D and B entail A*.

[29] In terms borrowed from Lakatos' theory of research programs (Lakatos 1978b) it may seem that the inconsistent sets emerged as a protective belt to Cantor's naive set theory while the consistent sets belong to the core of the theory. However, Cantor embraced the inconsistent sets as an essential part of his theory, not as a protective maneuver.

Chapter 18
Harward 1905 on Jourdain 1904

Not much is known about A. E. Harward (Moore 1976). According to the information in the paper (1905 p 439, subtitle and footnote) he was working with the Indian Civil Service. The signature of the article states: United Service Club Calcutta. His 1905 paper was published in October, by the same journal, Philosophical Magazine, which published Jourdain's 1904 papers. The same journal also published the only other known publication of Harward, from 1922, about Einstein's relativity theory.

According to information supplied by Harward in the mentioned footnote, the paper was completed in December 1904, communicated to Jourdain who made some comments and then submitted to the journal. Final submission occurred probably in May. A sequel paper was submitted (footnote on p 459) to the same journal during March 1905, before the first paper was returned from the publisher for revisions; apparently, it was never published.

Harward explains (p 439) that he was motivated to write his paper following his reading of Jourdain's 1904 papers because

> there appears to be considerable logical confusion about the subject [transfinite numbers] at present, and as Mr. Jourdain's proof of Schröder and Bernstein's theorem[1] and the theorems that $\mathfrak{a} = \aleph_0 \cdot \mathfrak{a}$ [Harward's Partitioning Theorem] and that $\mathfrak{a}^2 = \mathfrak{a}$ [Union Theorem][2] appear to me to be open to serious criticism, I have attempted in this article to give a brief restatement of the whole subject with rigorous proofs of the above theorems.

Harward's paper is indeed an improvement over Jourdain's papers in structure, clarity and rigorousity. Still it has some weak points upon which we will touch below.

Harward tells (p 439 footnote *) that on set theory he only read Russell 1903, Hardy 1903 and Jourdain 1904, not Cantor. This is quite amazing in view of the results he had obtained. Harward claimed that the distinction between aggregates

[1] This is how Harward called CBT.

[2] It seems that Harward was the first to publish a proof of the union theorem in its general form and his proof did not employ the axiom of choice.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_18, © Springer Basel 2013

and unlimited classes, Harward's terminology for consistent and inconsistent aggregates, he had found independently of Jourdain. His theory included the segregation axioms between these two types of sets, the same that were postulated by Cantor in a letter to Dedekind. To Jourdain he said that he owed the notion of well-ordered sets, which enabled him to "see how the distinction could be practically brought into evidence in mathematics" (p 439). Harward also touched on Cantor's epsilon numbers and the generalized continuum hypothesis (p 455). Unlike Jourdain, however, Harward did not relate to the comparability of sets, Corollary D and case (IV) (see Sect. 17.4). In addition, his proof of CBT is less interesting because it assumes the well-ordering principle, which narrows the theorem to its original scope in Cantor's theory.

18.1 Proof of CBT

Harward presented CBT, in its two-set formulation, as follows (p 445): It is not possible for two dissimilar aggregates to be each similar to a part of the other. For proof Harward argued thus (p 445f): Let A be an aggregate of cardinal number $\mathfrak{a}$ and B an aggregate of cardinal number $\mathfrak{b}$ and let A be similar to a part of B and B to a part of A. Let α be an ordinal that well-orders A. Let B_1 be a part of A that is similar to B. Let β_1 be an ordinal that well-orders B_1 by the well-ordering of A. If $\beta_1 = \alpha$, the proof is complete. Otherwise, B_1 contains a part A_1 similar to A and let α_1 be its ordinal under the well-ordering of A. If $\alpha_1 = \beta_1$, the proof is complete. Otherwise we continue this procedure. As there can be no infinite descending sequence of ordinals, there is an n such that $\alpha_n = \beta_{n+1}$ or $\alpha_{n+1} = \beta_n$. Then $A \sim A_n \sim B_{n+1} \sim B$, or the alternative, so that $\mathfrak{a} = \mathfrak{b}$.

Harward notes (p 445 footnote *, †) that his proof is not exposed to his criticism against Jourdain (see Sect. 17.5) because there is no repeated selection here: all the A_n, B_n are given at once. Indeed, if one chooses the mappings that provide the similarities mentioned in the theorem, the same mappings provide all the nesting subsets addressed in the proof. Moreover, the well-ordering of all the subsets is chosen only once, for the well-ordering of A.

Harward's CBT proof was an original in the history of CBT. Like Cantor's proof Harward's proof rests on the assumption that the sets concerned are well-ordered. Thus his CBT was also limited to aggregates and it does not apply to his unlimited classes. However, instead of Cantor's enumeration-by procedure that used transfinite induction, Harward generated Schröder's *Scheere* gestalt and focused on its exhibition in one of the sets, but not because he shifted to the single-set formulation. We name Harward's gestalt 'half Scheere'. Unlike Borel, Harward did not switch to the frames gestalt; instead he took the nesting sets gestalt to its ordinal counterpart, the descending sequence of ordinals that is necessarily finite (the metaphor). Note that the assertion that the sequence of ordinals is necessarily finite is equivalent to an application of transfinite induction. From this perspective too, then, Harward's proof is no real improvement on the proof of Cantor but it definitely rectifies Schröder's proof.

From the supposition of the well-ordering principle, Harward obtained the Comparability Theorem for cardinal numbers (p 445f), in the following way: Take two different cardinal numbers and take sets that are of these cardinal numbers. Assume the sets are well-ordered, then one of them is (ordinally) similar to a segment of the other. Therewith the relation of size of the cardinal numbers is implied, for the other set cannot be similar to a subset of the first as otherwise by CBT the cardinal numbers would have to be equal. Harward noted (p 446 footnote *) that his way of deriving comparability is simpler than Cantor's way, because of his use of CBT. By Cantor's way of deriving comparability Harward probably meant Jourdain's way. Jourdain, after he had established that all cardinal numbers are alephs (see Sects. 17.3 and 17.4) concluded that the comparability of cardinal numbers follows from the comparability of the alephs, which he took as obvious (1904a p 68). The trichotomy of the alephs follows from the lemma that all alephs are different which follows from the theorem that a well-ordered set is not ordinally similar to any of its segments, established by Cantor in his 1897 *Beiträge*. Since Harward's proof of CBT relied on the same theorem, through its use of the lemma that a descending sequence of ordinals is finite, his proof by way of CBT does not seem to us to be in any way simpler than the proof of Jourdain-Cantor.

18.2 Harward's Unlimited Classes and Other Basic Notions

Harward distinguished between aggregates (limited classes) and unlimited classes (p 439f). According to Harward, unlimited classes are those classes the individuals of which cannot be thought of collectively as a whole. That there are unlimited classes Harward argues as follows: "If all classes could be thought of collectively as a whole, the whole would be identical with one among many individuals comprised in itself, which is a contradiction in terms." Harward is clearly concerned here with the problem of foundation, as Dedekind was (see Sect. 8.4). Harward noted that the class of ordinals as well as the scale of number-classes, and thus the alephs, are unlimited classes.

Harward rejected (p 459) Jourdain's formal definition of an inconsistent aggregate as such that contains a part which is equivalent to the collection of all ordinals W (1904a p 67). His reason was that the notion of ordinals cannot be defined before making the distinction between aggregates and unlimited classes.[3] Harward, like Cantor, felt comfortable with the characterization of inconsistent sets as such that the assumption that they are sets leads to contradiction.

Harward suggested several rules to segregate between aggregates and unlimited classes. The rules help us decide if a class is an aggregate or is unlimited. With his idea of segregation rules, Harward hit upon Cantor's own ideas and expanded on

[3] Jourdain himself was unhappy with his formal definition for the same reason (see Sect. 17.3).

Jourdain's understanding of the distinction between the two sorts of aggregates. The rules correspond to three operations involved in class formation: subsets, union, replacement (p 440). As is the case with regard to Cantor's rules, Harward did not offer his rules in any existence or closure capacity but only for segregation purpose.

To the rules suggested also by Cantor Harward added (p 441) the following: The multiplicative class[4] of an aggregate of aggregates ("the class each of whose terms is an aggregate formed by taking one and only one element from each of those aggregates") is also an aggregate. This axiom, unlike the previous ones, is an existence, not a segregation, axiom. It provides that the existence of the power-class of a class A, apprehended as the multiplicative class of the aggregates {0, 1} and A.[5] Harward (p 441 footnote † and Note B p 459f) regarded his 'multiplicative axiom', as well as the segregation rules, as provisional, until the introduction of aggregates and unlimited classes is made formally, apparently by an axiom system. He declared (p 459) his intention to expand on the rules, their informal intuitive grounding and on the formal definition of unlimited classes, in the sequel paper that never came out.

The well-ordering of the aggregates Harward obtained as follows: Harward defined cardinal numbers, as Russell, as the equivalence classes of the relation of similarity (p 440f), Cantor's equivalence. It seems that he held that since only aggregates can be members of classes, not unlimited classes, only aggregates can have cardinal numbers. Harward then defined ordinal numbers, again like Russell, as the equivalence classes of the relation of ordinal similarity (p 443f). Thus, again, only aggregates have ordinal numbers. However, contrary to his own rules, Harward conceived of the class of all ordinals, Jourdain's W,[6] though the ordinals are not aggregates. It is possible that Harward held silently to the idea that ordinals are not equivalence classes but are represented by certain sets because he said (p 444) that for transfinite ordinal β,[7] the series of ordinals less than β form a series of ordinal β (and for finite β, the series of preceding ordinals is β less one because, like Cantor, Harward did not consider 0 as one of the ordinals). It takes only a single step to define from this observation every transfinite ordinal as the aggregate of its preceding ordinals. This can explain how he avoided Jourdain's twists between W and $\mathfrak{B}$ (see Sect. 17.3). Using the class of ordinals, let us denote it still by W, Harward thought like Jourdain that any aggregate can be well-ordered by choosing sequentially a member from it to correspond to each member of W (p 444). This is how Harward justified the supposition that all aggregates are well-ordered.

It is noteworthy that Harward was very much aware of transfinite induction and was perhaps the first to introduce it systematically (p 449), delineating the two

---

[4] Harward took the term from POM p 119.

[5] Namely, the Cartesian product of {0, 1} and A.

[6] A notation that Harward did not use.

[7] A cardinal (ordinal) of an infinite aggregate is called transfinite (p 441). Infinite aggregates Harward characterized after Russell 1903, who followed Dedekind on this matter (see Sect. 8.1), as aggregates with a reflection.

induction steps for successor and limit ordinals.[8] He even applied this method to the definition of the exponentiation of ordinals (p 449), as did Cantor in his 1897 *Beiträge* (Cantor 1932 (1915) §18; cf. Zermelo's remark in Cantor 1932 p 355 [24], Purkert-Ilgauds 1987 p 141). Since ω exponents will be used extensively below, let us give its Harward definition in detail (p 449): ω^2 means $\omega\omega$, with the multiplication of ordinals defined (p 442) as usual. If $\beta = \gamma + 1$ then $\omega^\beta = \omega^\gamma\omega$. Otherwise, ω^β is the first ordinal after all the ordinals $\omega^\gamma, \gamma < \beta$. Here Harward uses his replacement rule (p 440) to warrant that the collection of all ω^γ is an aggregate. Since this aggregate is well-ordered it has an ordinal. Notice that from the definition it follows that if $\beta < \beta'$ then $\omega^\beta < \omega^{\beta'}$, a lemma that Harward will be tacitly using.

18.3 Harward's Partitioning Theorem

For the Sum Lemma Harward (p 446) argued that every aggregate can be well-ordered by a limit ordinal (because it is an aggregate its cardinal is an aleph and AC is not invoked) and by replacing each element by two elements we can on the one hand split the set into two sets each equivalent to the original set, while on the other hand the replacement does not change in fact the ordinal of the set so neither is its cardinal number changed (see Sect. 17.6). He further suggested (p 446) a similar proof for $\mathfrak{a} = \aleph_0\mathfrak{a}$, which is direct, i.e., by-passing Zermelo's Denumerable Addition Theorem (see Sect. 13.1, Theorem II): For every limit ordinal α there is an ordinal β such that $\alpha = \omega\beta$.[9] Instead of replacing each member of the set of cardinal number $\mathfrak{a}$ with two elements, as in the proof that $\mathfrak{a} = \mathfrak{a} + \mathfrak{a}$, every element is replaced by ω elements. The result is of cardinal number $\aleph_0\mathfrak{a}$ and of type $\omega^2\beta$. But as every section of type ω^2 can be replaced by a section of type ω, this set is also of cardinal number $\mathfrak{a}$ and hence the required result. We will call this result Harward's Partitioning Theorem; it is to be distinguished from Zermelo's Denumerable Addition Theorem and from Banach's Partitioning Theorem to be discussed later.

18.4 Constructing the Number-Classes

Next (p 447f) Harward denoted by $C(\beta)$ the cardinal numbers of a series of ordinal number β and stated that if $\beta < \gamma$ then $C(\beta) \leq C(\gamma)$. Harward defined the order relation between cardinal numbers (p 441) like Cantor in his 1895 *Beiträge*, but he did not define $\leq$ explicitly. Apparently he had in mind that $C(\beta) \leq C(\gamma)$ because β can be 1–1 mapped into γ. Thus the remark that Jourdain had noted (see Sect. 17.2)

[8] Jourdain only referred to the 'generalized theorem of complete induction' in his 1908a paper (p 509 footnote †) and even there he did not differentiate between the two cases of the induction step.

[9] Harward skipped the proof, which runs as follows: $\omega\alpha \geq \alpha$ so there is β minimal such that $\omega\beta \geq \alpha$. If there is $\alpha < \gamma < \beta$ then $\alpha < \gamma < \omega\gamma < \omega\beta$, a contradiction to the minimality of β.

applies and CBT is necessary to conclude from $C(\beta) \leq C(\gamma)$ and $C(\beta) \geq C(\gamma)$ that $C(\beta) = C(\gamma)$.

Harward then defined that β, γ are said to belong to the same number-class when $C(\beta) = C(\gamma)$. Harward noted that if $\beta < \delta < \gamma$ and if $C(\beta) = C(\gamma)$ then $C(\delta) = C(\beta)$ which implies that the number-classes are segments of the class of ordinals and that the number-classes form a scale. However, Harward did not point out that the lemma he mentions here requires CBT.

Harward argued (p 447) that there can be no greatest number-class (cardinal) because by his multiplicative axiom and Cantor's Theorem, the aggregate of sub-aggregates of any number-class is an aggregate of greater cardinal number. Like Cantor, Russell and Jourdain, Harward denotes the initial number of the γ number-class by ω_γ. The scale of cardinal numbers associated with the scale of number-classes is comprised of the $C(\omega_\gamma)$ which is $\aleph_\gamma$ (p 448). This is a smoother definition then Jourdain's oblique observation of two types of number-classes (see Sect. 17.1).

Because each initial number is the first ordinal with cardinal number different (and hence necessarily greater) than all previous initial numbers, it seemingly follows that $\aleph_{\gamma+1}$ is the next cardinal number after $\aleph_\gamma$ (and similarly $\aleph_\delta$, δ limit, is the next cardinal number after all $\aleph_\gamma$, $\gamma < \delta$). Thus the Next-Aleph Theorem is seemingly very simply proved. But Harward did not explain why a number-class is not an unlimited class. To this end Cantor introduced the Limitation Principle and applied (implicitly) the Limitation Theorem.

Even though Harward was unaware of the Limitation Theorem, he still attempted to prove it (p 448f), namely, that $\aleph_{\gamma+1}$ is the cardinal number of the γth number-class (that starts with ω_γ). Assuming that x is the cardinal number of that number-class, by the definition of $\aleph_{\gamma+1}$ we have the equation $\aleph_{\gamma+1} = \aleph_\gamma + x$; if $x < \aleph_{\gamma+1}$ then $x \leq \aleph_\gamma$, because the ordinal of x must be either in the γth number-class or in one with smaller index. Thus, seemingly, by the Sum Lemma and CBT, we would have $\aleph_{\gamma+1} = \aleph_\gamma$, a contradiction to the definition of $\aleph_{\gamma+1}$. Harward's argument follows the alternative path we mentioned in Sect. 2.2 following the Sum Lemma and it suffers from the same lacuna mentioned there.

18.5 The Union Theorem

Harward (p 457) criticized Jourdain's proof of the Union Theorem for the matrix $\omega_1 \times \omega_1$ (see Sect. 17.6). He claimed that Jourdain showed how to correlate disjoint sections of the matrix containing the pairs (α, β) and bordered by the condition that $\alpha + \beta$ is between two ordinals (infinitely apart) from (II), to the section defined by these ordinals. This is possible because both sets correlated are denumerable. Harward then contended that though this process, which evidently requires transfinite induction, can be continued indefinitely, it cannot be concluded that it terminates and that thus the matrix and (II) correlate.

Interestingly, Harward did not present Jourdain's proof (see Sect. 17.6) but an improvement thereof. Jourdain did not use the idea of dissecting the matrix to slices by the values of $\alpha + \beta$. Anyway, Harward's criticism of the improved proof was for the lack of a demonstration that the entire matrix is covered by the process.[10] When Jourdain retorted, in his 1907b paper (p 356 footnote * second paragraph), he did not answer Harward's criticism and only mentioned that the multiplicative axiom can be proved in the case of $\omega_1 \times \omega_1$. This was Jourdain's odd way of saying that $\omega_1 \times \omega_1$ can be well-ordered. Clearly, given this well-ordering it can be guaranteed that every pair will be correlated to a number in (II). In his 1908a paper (see Chap. 23) Jourdain indeed defined a well-ordering on $\omega_1 \times \omega_1$ but apparently he must have found that well-ordering at a late stage because in his 1907a paper,[11] he noted that the multiplicative axiom must be assumed for the proof. So Harward was justified in his criticism of Jourdain 1904.

Harward thought that by Jourdain's method it can be argued that, since every number of (II) is denumerable, also ω_1 is denumerable. Take a denumerable sequence and correlate every alternate member to a member in a segment of ω_1 and then every alternate member of the residue to a member of a following segment and so on. Obviously this argument is false since ω_1 is not denumerable and Harward maintained that the same applies to Jourdain's argument in his proof of the Union Theorem. But obviously this is not correct. In the example, only a denumerable sequence of numbers in (II) can be generated by its argument and this sequence does not have ω_1 as its sequent. In Jourdain's proof there was no question that there are enough elements in the matrix to cover ω_1.

As we already explained, Harward did not need the Union Theorem for his construction of the scale of number-classes. Still, because of Jourdain's engagement with this theorem Harward was drawn to it too and he gave it two proofs.

We will present first the second proof that Harward gave to the Union Theorem (p 458) because it is based on the gestalt of Cantor's proof that the rationals are denumerable, though apparently Harward had not seen Cantor's proof. Harward brought the proof just after the mentioned criticism of Jourdain's proof, in an appendix to the paper (Note A), saying that he had recently found it. It seems that Harward found the proof after he sent the paper to Jourdain and perhaps after it was first submitted to the journal.[12]

Harward took the pairs (α, β) in the square of pairs with $\alpha, \beta < \omega_1$, on the lower side of the diagonal of pairs (α, α) without the diagonal. This triangular matrix has the power of $\omega_1 \times \omega_1$ when the rows are placed one after the other. However, when the columns are placed one after the other, every segment of the sequence is denumerable (because $\aleph_0\aleph_0 = \aleph_0$) but the ordinal of the sequence is greater than

[10] Proving that the correlation is on (II) is not necessary because CBT can be used. In both the theories of Jourdain and Harward CBT is proved earlier than the proof of the union theorem. Cantor, in his 1883 *Grundlagen*, had to prove both in tandem.

[11] The paper was presented to the London Mathematical Society in April 1906 but the proceedings for 1906 appeared in 1907.

[12] In a footnote to the appendix (p 455), Harward tells that in the copy of the article that he had sent Jourdain the attached note was defective. He adds that Jourdain commented on it but that his comments do not "meet the real objection".

every denumerable ordinal and so the ordinal of the entire sequence is ω_1. This proof can be generalized by transfinite induction to every ω_γ, hence $\aleph_\gamma{}^2 = \aleph_\gamma$. When in his 1908a paper Jourdain (p 511) rejected Harward's criticism of his 1904 proof of the Union Theorem, he noted that in Harward's proof the following mapping is implicitly defined: to every pair of ordinals α, $\beta < \omega_\gamma$ the ordinal $\Sigma_{\zeta<\beta}\zeta+\alpha$ is related. This mapping is a 1–1 mapping from the matrix $\omega_\gamma \times \omega_\gamma$ onto ω_γ. Thus the proof generalized Cantor's proof that $\omega \times \omega$ is denumerable (1895 *Beiträge* §6) by the mapping $\lambda = \alpha + ½(\alpha + \beta\text{-}1)(\alpha + \beta\text{-}2)$. In this formula the second summand is indeed the sum of all numbers preceding the sum $\alpha + \beta$-1, which enumerates the diagonals of the matrix $\omega \times \omega$. In the case of this matrix we had to look on the diagonals to switch from the gestalt of the square to the gestalt of the triangle. The triangular gestalt also appears in Cantor's proof that the algebraic numbers are denumerable, where in each column the numbers with the same height appear. The triangular gestalt may have had its origin in Greek pebble mathematics (triangular numbers).

The other proof given by Harward (pp 451–454) rolls out like this: Let (A) denote the two statements[13]: For every transfinite ordinal β, $C(\omega^\beta) = C(\beta)$ and $C(\beta)^2 = C(\beta)$. The proof of (A) is by transfinite induction.

For $\beta = \omega$ both statements follow from $\aleph_0\aleph_0 = \aleph_0$ (p 450f). Assume that (A) holds for all $\gamma < \beta$. If $\beta = \gamma + 1$, clearly $C(\beta) = C(\gamma)$ and $\omega^\beta = \omega^\gamma\omega$ so $C(\omega^\beta) = C(\omega^\gamma\omega) = C(\omega^\gamma)\aleph_0 = C(\gamma)\aleph_0 = C(\gamma) = C(\beta)$. The equality $C\ (\omega^\gamma\omega) = C(\omega^\gamma)\aleph_0$ is by the definition of the multiplication of ordinals and of cardinals which Harward took from Russell. $C(\omega^\gamma)\aleph_0 = C(\gamma)\aleph_0$ follows from the induction hypothesis $C(\omega^\gamma) = C(\gamma)$. $C(\gamma)\aleph_0 = C(\gamma)$ follows from Harward's Partitioning Theorem.

If β is a limit ordinal then the set of all ordinals smaller than ω^β is partitioned into the β partitions of the ordinals between ω^γ and $\omega^{\gamma+1}$ for every $\gamma < \beta$. Each such partition has a cardinal number $\leq C(\omega^{\gamma+1})$, which is by the induction hypothesis $C(\gamma)$, which is $\leq C(\beta)$ by definition. Therefore, we have $C(\omega^\beta) \leq C(\beta)^2$.[14] Since trivially $C(\omega^\beta) \geq C(\beta)$, we have to prove that $C(\beta)^2 \leq C(\beta)$ to obtain the desired result by CBT. If now β is not an initial number there is an initial number $\gamma < \beta$ such that $C(\beta) = C(\gamma)$. Then, by the induction hypothesis, $C(\beta)^2 = C(\gamma)^2 = C(\gamma) = C(\beta)$ as required.

Otherwise, β is an initial number, say ω_κ. Harward presents all the ordinals less than ω_κ in the following matrix for every $\delta < \beta$: In the first row are all the ordinals of the form ν and $\omega\delta + \nu$ where ν is finite not 0. In row γ, $\gamma < \beta$, are all the ordinals of the form $\omega^\gamma\nu$ and $\omega^{\gamma+1}\delta + \omega^\gamma\nu$ where ν is finite not 0. There are $C(\beta)$ matrices, and each matrix has $C(\beta)$ rows and $\omega 2$ columns. By Harward's Partitioning Theorem each matrix is of cardinal number $C(\beta)$ and all the matrices together of cardinal number $C(\beta)^2$, provided all the ordinals in the matrices are different.

---

[13] The need to prove by transfinite induction several theorems in tandem, characterizes also Cantor's CBT proof. Perhaps more study is required to explain this phenomenon.

[14] We need to choose a mapping between $\omega^{\gamma+1}$ and γ so AC is invoked.

Harward next proves that the cardinal number of each ordinal in the matrices is smaller than $C(\beta)$. Hence he would be able to conclude that the number of ordinals in the matrices, namely, $C(\beta)^2$, is $\leq C(\beta)$, as required. Note that Harward does not claim that every ordinal $< \beta$ is in the matrices. By the induction hypothesis, we have: $C(\delta)^2 = C(\delta)$; $C(\omega^{\gamma+1}) = C(\gamma + 1) = C(\gamma) = C(\omega^{\gamma})$, and $C(\delta), C(\gamma) < C(\beta)$ because β is assumed to be initial number. So:
$C(\omega^{\gamma+1}\delta + \omega^{\gamma}\nu) = C(\omega^{\gamma+1}\delta) + C(\omega^{\gamma}\nu) = C(\omega^{\gamma+1})C(\delta) + C(\omega^{\gamma})C(\nu) = C(\gamma)C(\delta) + C(\gamma)\nu$. The second term equals $C(\gamma)$ by the Sum Lemma,[15] and the first, by the induction hypothesis, is either $C(\gamma)$ or $C(\delta)$, whichever is greater. So finally, $C(\omega^{\gamma+1}\delta + \omega^{\gamma}\nu)$ is also equal to the greater of $C(\gamma)$ or $C(\delta)$, and so it is $< C(\beta)$.

Finally Harward proves that all the ordinals in the matrices are different. For otherwise we would have $\omega^{\gamma_1+1}\delta_1 + \omega^{\gamma_1}v_1 = \omega^{\gamma_2+1}\delta_2 + \omega^{\gamma_2}v_2 = \kappa$. So on the one hand κ ends with a segment of ordinal ω^{γ_1} while on the other hand it ends with a segment of ordinal ω^{γ_2}. This is possible only when "$\gamma_1 = \gamma_2$ and it obviously follows that $\nu_1 = \nu_2$ and $\delta_1 = \delta_2$". Clearly, $\omega^{\gamma} < \omega^{\gamma+1}$ and thus if $\gamma_1 < \gamma_2$ then $\omega^{\gamma_1} < \omega^{\gamma_2}$. It is easy to prove by transfinite induction that the end segment of ω^{γ} has the ordinal ω^{γ}. Thus if $\alpha + \omega^{\gamma_1} = \beta + \omega^{\gamma_2}, \gamma_1 < \gamma_2$, necessarily an end segment of ω^{γ_2} is similar to a subset of an end segment of ω^{γ_1}, which is a contradiction because an ordinal cannot be similar to a subset of a smaller ordinal. A similar argument proves that $v_1 = v_2$. If δ_1 and δ_2 are not equal we would have a section of ordinal $\omega^{\gamma+1}$ similar to a subset of $\omega^{\gamma}v$, which is a smaller ordinal, and the same contradiction arises.

[15] Not mentioned in this context by Harward.

Chapter 19
Poincaré and CBT

In the November 1905 and January 1906 issues of the *Revue de mètaphysique et de morale ('Revue')*, Poincaré published in two parts an article titled *Les Mathematiques et la Logique* (Poincaré1905; Poincaré 1906a). The article was aimed to refute the two reductionist research programs, Russell's logicism and Hilbert's formalism. What triggered Poincaré to his attack was mainly a series of papers by Couturat in 1904–1905 who expounded the logicist point of view.[1]

Poincaré's refutation was by a methodological argument. He claimed that both programs tacitly employ complete induction before introducing it into their system. Formalism, which will occupy us no more here, requires complete induction for a proof that its constructions are consistent (1905 p 833f); logicism requires it for a proof of CBT. This theorem Poincaré describes as "the Fundamental Theorem of the theory of infinite cardinal numbers" (1906a p 27). Thus Poincaré regarded CBT as one of the first mathematical theorems that must be reduced to logic, if logicism is to be upheld. However, Poincaré (first) maintained that CBT proofs necessarily use complete induction, which, contrary to Couturat, Poincaré regarded as irreducible to logic, being grounded in intuition (McLarty 1997). Poincaré saw in complete induction the paradigm of Kant's synthetic a priori judgment. Couturat's anti-Kantian stance irritated Poincaré and as a result, the article is tainted with sarcastic tone.

Poincaré even maintained that CBT must be obtained, under the logicist program (1906a p 27), before the finite number theory is constructed. Such view is justified because under the logicist program numbers are equivalence classes for Russell's relation of similarity ('equivalence' in Cantor's terminology); CBT is required to warrant the "smoothness" of these equivalence classes, that when $A \subset B \subset C$ are classes such that A and C are similar, B is similar to both.[2] In addition, CBT enables the definition of the inequality of numbers, as an antisymmetrical relation.

[1] Russell's 1905 paper, which criticized a 1905 paper of Boutroux (Poincaré's nephew), was published in the same issue of the *Revue* so it did not trigger Poincaré's article.

[2] This justification is not explicitly made by Poincaré.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_19, © Springer Basel 2013

Interestingly, both in Whitehead's 1902 paper (Sect. 15.3, see Chap. 15) and in his "Principles of mathematics" (POM), Russell disregarded this point and defined the numbers and their inequality well before he mentioned CBT, though in POM the relevance of CBT to the definition of the inequality of numbers is noted (p 306). Only in Principia Mathematica (see Chap. 26), the order of presentation follows the correct methodology.

To illustrate his point Poincaré brought in the second part of his article, a proof of CBT, adapted from Borel 1898. Poincaré detailed the use of complete induction in the proof, which we will present and discuss below. Following the proof, Poincaré challenged Couturat to produce a proof that avoids complete induction (1906a p 29). Poincaré's challenge was a turning point in the history of CBT. Several mathematicians responded to the challenge and provided such proofs: Zermelo (already in January 1906 by a letter to Poincaré; the proof was published by Poincaré in his 1906b and later appeared in Zermelo 1908b), Peano (May and August 1906), J. Kőnig (June 1906) and Bernstein (December 1906). The first two are similar as are the last two.[3] Other reactions to the article appeared as well, by Couturat and Pieri. In addition, a paper by Russell (Russell 1906a) was published in March and though it was written before Poincaré's article came out, it was relevant to Poincaré's criticism and was among the reasons that provoked Poincaré to publish a sequel (1906b) to his 1905–1906 article.

Before Poincaré published his 1906b, he became acquainted with Richard's paradox and accepted Richard's explication of it. The result was his doctrine of barring "non-predicative [impredicative][4]" definitions from mathematics (cf. Fraenkel et al. 1973 p 193). Poincaré (1906b) was published in the May 1906 issue of the *Revue*. There Poincaré rejected, for reasons of impredicativity, a proof of CBT that he constructed out of Schoenflies 1900, and to which he amalgamated an argument used by Russell in Whitehead 1902. Below we bring the proof, rounded with a regular argument, and the Russell-like argument constructed by Poincaré with its rejection.

Poincaré's 1906b sequel, further brought and rejected on grounds of impredicativity the proof of Zermelo, an argument that Cantor used in his proof of CBT for sets of the power of II (see Chap. 1) and an argument that Poincaré derived from Zermelo's proof of the Well-Ordering Theorem. We cover all this below and also Zermelo's response to Poincaré's impredicativity criticism from his 1908a paper.

In retrospect, Poincaré's criticism may seem cranky (Lakatos 1976 p 42), for the programs that he attacked took center stage in the following years. Yet at the time, Poincaré's criticism was important for it contributed to the crystallization of ideas driving both logicism and formalism: the mathematics of metamathematics, the vicious circle principle, impredicative definitions and the predominance of CBT

[3] We will cover these proofs in the following chapters.

[4] We use the current 'impredicative' interchangeably with Poincaré's 'non-predicative'.

over the notion of cardinal number. In his 1906b paper (which appeared in the September 1906 issue of the *Revue*) Russell responded to the non-predicativity and some other points in the criticism of Poincaré. Poincaré retorted in his 1906c (titled *A propos de la logistique,* appearing in the discussion part of the November 1906 issue of the *revue*). Below we address the issues raised in these two papers, which ended the polemic.

Poincaré calls CBT "theorem of Bernstein" (1906a p 27) ignoring Schröder, unlike the German set-theorist of the time who named the theorem the "Equivalence Theorem" or the British, Whitehead-Russell and Jourdain, who used both the names of Bernstein and Schröder in calling the theorem. It is possible that in his naming decision Poincaré expressed his criticism of Schröder's proof.

19.1 The First Proof by Complete Induction

Poincaré stated CBT in its two-set formulation (1906a p 27) as follows:

> The theorem of Bernstein teaches us that if: $A_0 \equiv B_1$ and $A_1 \equiv B_0$, A_1 being a subset[5] of A_0, and B_1 a subset of B_0, we have likewise that: $A_0 \equiv B_0$.

Preceding the theorem Poincaré defines $A \equiv B$ for two sets A, B by:

> When the elements of the two sets can be made to correspond in such way that for every element of one of them corresponds one and only one element of the other, the sets are said to have the same cardinal number and we write $A \equiv B$.

Like Borel, Poincaré does not define the notion of cardinal number but only of when two sets, have the same cardinal number ('power' in Borel). The proof of the above theorem Poincaré provided as follows:

> From $A_0 \equiv B_1$ it follows that to every element of A_0 corresponds an element of B_1 and as A_1 is a subset of A_0,[6] to the elements of A_1 correspond elements of B_1, the set of which, B_2, is thus a subset of B_1 and we have $A_1 \equiv B_2$ and A_0-$A_1 \equiv B_1$-B_2.
>
> Similarly the set A_2 is defined, which is a subset of A_1 and such that $A_2 \equiv B_1$ and B_0-$B_1 \equiv A_1$-A_2.
>
> Now since $A_1 \equiv B_2$ and A_2 is a subset of A_1, there exists a set B_3 that is a subset of B_2 satisfying the following conditions:
>
> $A_2 \equiv B_3$ and A_1-$A_2 \equiv B_2$-B_3.
>
> We similarly define A_3 *and so on*[7] until we have a sequence of sets $A_0, A_1, \ldots, A_n \ldots$, $B_0, B_1, \ldots, B_n \ldots$,[8] such that each A_{n+1} is a subset of A_n and B_{n+1} a subset of B_n and so we have:
>
> $A_n \equiv B_{n+1}$, A_{n-1}-$A_n \equiv B_n$-B_{n+1}

[5] For Poincaré a subset is a proper subset, like Cantor and unlike Dedekind. Poincaré uses '*être partie*' for 'being subset' and '*faire partie*' or '*appartenir*' for 'being member'.

[6] There is a typo in the original and B_0 is printed instead of A_0.

[7] Italics here and below are in the original for later reference.

[8] There is a typo in the original and B_n is printed instead of A_n.

$B_n \equiv A_{n+1}$, B_{n-1}-$B_n \equiv A_n$-A_{n+1}.

Let now C be the set of all elements common to the different sets $A_0, A_1, \ldots, A_n \ldots$, and D the set of all elements common to the different sets $B_0, B_1, \ldots, B_n \ldots$; we have:

$A_0 = \Sigma(A_n\text{-}A_{n+1}) + C$ and $B_0 = \Sigma(B_n\text{-}B_{n+1}) + D$.

Because *when, in an indefinite series of sets, each is a subset of its predecessor, the first is formed of all the elements that belong to all these sets and all those that belong to one of them without belonging to its sequent.*

This principle that I have just underlined [italicized] is clearly evident; but it seems that it supposes an appeal to intuition; I will not insist on this point. We demonstrate now that: $C \equiv D$.

In fact, to a member of C, being a member of A_0, corresponds an element of B_1,[9] because of the correspondence defined by the relation: $A_0 \equiv B_1$.[10]

As this element belongs to A_n and the same correspondence defined by $A_0 \equiv B_1$ is the same as that which is defined by $A_n \equiv B_{n+1}$, the corresponding element is a member of every B_{n+1}; it is therefore a member of all the B's and as a result of D. Inversely, for each element β of D corresponds an element α[11] of A_0 by the same correspondence[12]; and as this element β belongs to B_{n+1}, the element α belongs to A_n and to *all* A_n and as a result to C; we have therefore: $C \equiv D$.[13]

And by bringing together all our equations: $A_0 \equiv B_0$.

Here are some comments on the historic and heuristic background of the theorem and proof:

- The use of zero indices is justified by the construction used in the proof, but not the provisions of the theorem itself. Such use seems not heuristic and for this reason, perhaps, it was avoided by all other proofs of CBT. It may have been a later addition and the cause of the typos.
- The definition of the A and B sequences appears in Schröder (1898) but Schröder missed the gestalt switch that shifts the focus from the nesting sets to their differences which is central for Poincare. It is in Borel's proof (1898) that the differences (frames) between the nesting sets first appear and clearly, Poincaré's proof is based on Borel's gestalt, though Borel is not mentioned. The gestalt that we identify in Poincarè's first proof we describe as the *scheere* of frames.
- However, Borel, after the definition of A_2, constructed only the sequence of A's, and thus shifted to prove the single-set formulation of CBT. Poincaré stayed with the two-set formulation, as did Schoenflies (1900).
- In addition, Borel chose a new mapping at each step of the construction of the sequences of nesting sets, which entangled him, unknowingly, with the axiom of choice. Poincaré is using the same mappings throughout, those given by the equivalences $A_0 \equiv B_1$ and $A_1 \equiv B_0$. However, he does not name them explicitly as did Schröder.

[9] There is typo in the original and B is printed instead of B_1.

[10] Again B is printed instead of B_1.

[11] In this sentence α is printed for β and β for α but we interchanged between them according to the printing in the next sentence.

[12] If β is in D it is in B_1 so it has an origin in A_0 under the mapping between A_0 and B_1.

[13] The second step of the proof provides that the mapping from C into D is on.

- Borel emphasized that the equivalence of the frames follows from the fact that the inner set and the frame partition the outer set, and not from the mere equivalence of the inner and outer sets. Poincaré made no similar point explicitly but he carefully derived the equivalence of the frames, e.g., A_0-$A_1 \equiv B_1$-B_2, in tandem with the equivalence of the inner sets (A_1, B_2).
- In Borel's proof the notion of 'equal power' is used repeatedly, at the point where a new mapping is chosen for the next step of the definition, though essentially the proof is in the language of sets and mappings. Poincaré makes no similar digressions and maintains all along his discourse in that language only. In this Poincaré's proof differs from the proofs of Zermelo 1901 and Russell 1902.
- The definition of A_2 and consequently of all other nested sets is by what we called, following Schröder, 'Dedekind's Lemma' (see Sect. 10.1).
- Poincaré's proof has an algorithmic character: once A_1 and B_1 are defined, it turns out that they fulfill the same conditions given for A_0 and B_0. Awareness to this point provides the repetitive metaphor of the proof.
- In the construction, even frames in the sequence of A's (between even and odd members) correspond to odd frames in the sequence of B's, and vice-versa. This creates a gestalt of crisscross correspondence between the frames. In view of this and the previous point, we depict Poincaré's proof with the image 'shoe lacing'.
- The proof that the residues are equivalent was perhaps clear to Schröder too but he thought the residue is also equivalent to each of the nesting sets, which is nonsense (the residue could be empty, see Chap. 25).

Regarding the role of complete induction, Poincaré says (1906a p 29) that it is applied, at the italicized words *and so on*: "our sets A_n and B_n are defined by induction and induction is used to reason over them." The inductive nature of the definitions of the sets is clear. Note that the proof of $A_{n+1} \subset A_n$, and of the equivalence of the frames, do not require a new act of inductive reasoning as they are obtained at the definition stage. Proof by complete induction is used to prove that the intersection of the nested sets is equal to the residue obtained after removal of the frames. Thus the proof that A_0 (respectively B_0) is equal to the sum of the frames of A's and the intersection C does not require a vague special act of intuition, as Poincaré suggests in the italicized passage, but simply a proof by complete induction (and extensionality). Defining C, D as the residues will not avoid use of induction because proof that the residue is equal to the intersection is still necessary. Note that Cantor used diminution by a sequence of frames in his 1883 paper presentation of derived sets (Cantor 1932 p 160).

19.2 The Second Proof by Complete Induction

Poincaré returned to CBT in the third part of his article (1906b p 313) to demonstrate that if Russell's definition of complete induction is used in the proof, and thus intuitive complete induction is seemingly avoided, the result is defective because it uses impredicative definition. The proof that Poincaré considered he probably devised from the proof of *2.6 of Whitehead 1902, which states that every infinite class has a subclass of cardinal number $\aleph_0$. In this section we will present the proof up to the point where Poincaré brings in Russell's inductive notions. From that point on we will complete the proof by regular complete induction, as was done in Fraenkel 1966. In the next section we return to Poincaré's Russell-like argument. Here is the theorem and the first part of its proof (italicized text in the original):

> If A_0 can be decomposed into three subsets H_0, Q_0 and A_1, so that $A_0 = H_0 + Q_0 + A_1$ and if A_1 is equivalent to A_0 so that $A_1 \sim A_0$,[14] then $A_1 + Q_0$ will also be equivalent to A_0.
>
> In fact if A_1 is equivalent to A_0, to every element of A_0 corresponds an element of A_1 that we may call its image. If B is a set contained in A_0, the images of the elements of B form a set that we may call the image of B and that we denote by $\varphi(B)$.[15] We have therefore $\varphi(A_0) = A_1$. So put:
>
> $\varphi(A_1) = A_2$, $\varphi(H_0) = H_1$, $\varphi(Q_0) = Q_1$,
> $\varphi(A_2) = A_3$, $\varphi(H_1) = H_2$, $\varphi(Q_1) = Q_2$,
> ....
>
> *and so on*. We have $A_n = H_n + Q_n + A_{n+1}$ when the index n is any inductive number.[16]
>
> Let C be the set of *all* elements that belong to *all* A_n for which n is an inductive number. It is easily demonstrated that $\varphi(C) = C$.
>
> I say now that (1) $A_0 = H_0 + Q_0 + H_1 + Q_1 + \ldots + C$, that is to say that all elements of A_0 that do not belong to any H_n or to any Q_n for which the index n is an inductive number, must belong to *all* A_n of inductive index and therefore to C.

Denoting by C' the set that remains of A_0 after the removal of all H_n and Q_n, the final sentence quoted says that C' is equal to C, which justifies (1). A member of C is clearly not a member of some A_k-A_{k+1} for otherwise it would not be a member of A_{k+1}. So C is a subset of C'. To prove that C' is a subset of C, and thus that C = C' by extensionality, it is necessary to prove that a member of C' is a member of every A_n. This is easy by complete induction.

The proof can now be completed by defining the mapping ψ between A_0 and $A_1 + Q_0$ as the identity on the Q's and C and φ on the H's. The proof obviously combines elements from both Borel's proof (Chap. 11) and the proof given by Schoenflies (Chap. 12). From Borel it is taken that the proof is for CBT in its single-set formulation, from Schoenflies was taken the naming of the frames, though Poincaré uses different letters. Poincaré made no reference to these earlier sources. Unlike Schoenflies, but like Zermelo (Chap. 13), Poincaré names the correspondence φ between A_0 and A_1. As in his first proof, and Schröder's proof, Dedekind's Lemma is used to generate the nesting sets.

Fraenkel (1966 p 73ff) brought a similar proof [17] and he illustrated it with the following drawing (with different notation); C is the vertical rectangle on the left crossing the horizontal rectangles of the A's (Fig. 19.1):

[14] Poincaré switched to reverse tilde instead of ≡ for equivalence. We use regular tilde.

[15] Here φ is, rather strangely, introduced.

[16] The definition of this notion is given in the next section. Here n should be interpreted as a natural number and the definition of the sequences of sets as contemplated by complete induction.

[17] With regard to this proof Fraenkel remarked (1966 p 77) that the mapping established in it uses the theorem that the union of a finite number of finite or denumerable sets with at least one set denumerable, is denumerable, or the relation $\aleph_0 = \aleph_0 + 1$. We failed to understand the reason for this remark.

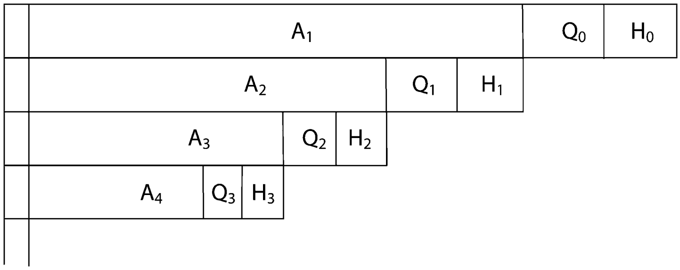


Fig. 19.1 Drawing for Poincaré's second proof

We give this proof the gestalt 'hanging staircase' and the metaphor 'pushdown part of the stairs'. Cleary we have here two chains: one of Q_0, the other of H_0. But Poincaré is not using the chain concept. In this gestalt the nesting subsets of the A_n vanish and it is realized that the original set is composed only of the two chains and the residue. Poincaré's gestalt is similar to the gestalt in Harward's proof, which we named half *Scheere* but it is half of a *Scheere* of frames, similar in its 'half' part to the gestalt of Schoenflies' proof. In its pushdown metaphor the proof of Poincaré is similar to the pushdown the chain metaphor used in Dedekind's proof.

19.3 The Russell-Like Argument

Poincaré's proof continued as follows:

> And in fact the class K formed by the A_n to which belongs an element that is not a member of any A_k-A_{k+1} whose index k is an *inductive* number is a recurrent class [for this notion see below].
>
> It is clear that the equality (1) entails the stated theorem; but because in the definition of the class K appears the notion of inductive number we find the same vice of the outlined demonstration of §XI. What does this mean? the demonstration of Bernstein's theorem remains legitimate, but on the condition that the principle of induction is regarded there as synthetic judgment and not as definition, because this definition is non-predicative.

The logicist terms used in the proof were introduced by Poincaré[18] thus (p 308): "Call *recurrent class* every class of numbers that contains 0 and contains $n + 1$ if it contains n. Call *inductive number* all numbers that belong to all the recurrent

[18] Poincaré references Russell's “"recent memoir” for the origin of his definitions but no publication of Russell prior to Poincaré's paper contains exactly these definitions. The notion of inductive number appears in Russell 1906a p 49. The notion of recurrent class appears within Russell's definition of the class of finite numbers, it is s in Whitehead 1902 III.*1.6:
$Nc\_fin = \{n \mid n \in Nc \wedge \forall s(s \in Cls \wedge 0 \in s \wedge \forall m(m \in s \cap Nc \rightarrow m + 1 \in s) \rightarrow n \in s)\}$. Cf. Russell 1903 p 115, p 123, 1906a p 36 footnote. It appears that Poincarè did a close study of the logicist notions.

classes.” Obviously, the notion of ‘(cardinal) number’ must be defined to put any sense into these definitions. Poincaré does not define this notion and so we assume he had in mind the definition of Russell (POM p 115): the number of a class is the class of all classes that are equivalent to that class. Next, we assume that Poincaré also had in mind Russell’s definition of $n + 1$ (given in generalized terms in POM p 117f): it is the number of the union of a set of number n and a disjoint set of number 1.

Clearly, if n is inductive so is $n + 1$. The inverse is also true: otherwise, and n is not 0, the set of all inductive numbers except n and $n + 1$ would also be recurrent and so $n + 1$ would not be an inductive number. For every number for which A_n, H_n, Q_n, are defined, the definition of A_{n+1}, H_{n+1}, Q_{n+1}, can be contemplated as indicated in the proof using φ. Therefore the class of all numbers for which the definition of A_n, H_n, Q_n, is obtained, is recurrent so that the definition is obtained for all inductive numbers. Proof that $\varphi(C) = C$ can be the following: For x in C, we have to prove that its φ-image is in C and that it is the φ-image of a y in C. If x is in C, x is in every A_n with an inductive index and thus its image in every A_{n+1} with an inductive index. As the intersection of all A_n with an inductive index is the same as the intersection of all A_{n+1} with an inductive index, the image of x is in C. As x is in A_1, there is a $y \in A_0$ such that x is the image of y. If $y \notin A_n$ for some inductive n, then x is not in A_{n+1}, contrary to the assumption that x is in every A_n with an inductive index. So $y \in A_n$ for all inductive n, hence $y \in C$ and the proof is complete.

To justify (1) note first that for every inductive n, A_n, H_n, Q_n, $\subseteq A_0$ and thus also $C \subseteq A_0$. Hence, assuming extensionality, we have only to prove that every element of A_0, which is not in any H_n, Q_n, where n is an inductive number, is in C. To this end, Poincaré introduces the class K. However, K is not correctly introduced. Poincaré defined a recurrent class as a class of numbers but with regard to K, he says that it is a class of A’s, though the A’s are sets not numbers. A corrected definition of K should be: K is the class of those n for which A_n obeys a certain requirement. The requirement that Poincaré puts forward is that to A_n “belongs an element that is not a member of any A_k-A_{k+1} whose index k is inductive”. This requirement also seems mistaken: we need to replace “an” with “every” (or use a different K_x for every x). Thus the requirement on K should be: to K belong all n such that A_n contains every element of A_0 that is not in any A_k-A_{k+1} with inductive k. This K is indeed a recurrent class and must contain all inductive numbers. So the members of A_0 that do not belong to any of the frames belong to every A_n and is therefore in C. Thus it is proved that the intersection of the A_n’s is equal to the residue of A_0 after removal of all frames.

Now, Poincaré maintains that (1) entails the theorem. He no doubt intends that we can obtain a 1–1 mapping between A_0 and A_1 if we combine φ on the H_n’s and C with the identity on the Q_n’s. However, this proof is not valid, contends Poincaré, because in the definition of K there appears the notion “inductive number”, while in that notion the notion of recurrent class, such as K, appears; thus K is used in its own definition and so the definition of K is circular and K cannot be contrived. A similar failure would occur, concluded Poincaré (p 314), in any attempted definition of the principle of induction.

Instead of stating the final argument of rejection, the vice, Poincaré directs us to §XI (1906b p 309), where a similar construction of a recurrent K is rejected. The lemma that he brings there appears in Whitehead 1902[19] as line (2) of theorem *2.6. Here is the lemma with its Poincaré proof:

> If n is not inductive and m is, n-m is non-inductive.
>
> And in fact, the class of numbers m such that, n being an arbitrary not inductive number, n-m is not inductive, this class, I say, is recurrent; therefore if m is inductive it must belong to that class.
>
> It is easy to verify that this class that we call K is recurrent; for zero in fact belongs to it because n-0 [namely, n] is not inductive, if n is not inductive; in addition m + 1 belongs to it if m does; because if n-m is not inductive, it is the same of n-m-1.

But here lies a defect, says Poincaré: a recurrent class, in the definition of which appears the notion of inductive number, such as K, cannot be one of the classes that are used in the definition of inductive number. Otherwise, the definition contains a vicious circle and defines nothing. Therefore, we cannot conclude that all inductive numbers belong to K.

19.4 On Impredicativity and Poincaré's Influence on Russell

Poincaré took the term "non-predicative" (1906b p 307) from Russell (1906a p 34; cf. Poincaré 1909 p 200). Russell used the term for propositional functions ('norms' in the terminology of his 1906a, after Hobson 1905) "which do not define classes". For such norms, the assumption that they do define a class leads to a contradiction. Poincaré, however, attributed the term to definitions "that contain a vicious circle" (1906b 307), namely, which reference the defined. In his answer to Poincaré (Russell 1906b), Russell adopted Poincaré's descriptor 'vicious circle' to definitions which Poincaré called non-predicative. The term for such definitions that finally prevailed (we do not know when this happened exactly) is 'impredicative definition'.

Russell (1906b p 646) seemingly accepted Poincaré's criticism regarding the circularity in the definition of K but he claimed that it can be avoided, referring to the introduction of what later would be called the 'axiom of reducibility'. Poincaré, however, who identified that Russell is introducing a new axiom, saw no solution in this move (1906c p 868) for the new axiom is not more evident than the principle of induction.

Interestingly, it seems that by singling out the notion of recurrent class and shaping around it the definition of inductive number, Poincaré made an important contribution to logicism. For in Whitehead-Russell 1910–13 ('PM', vol I p 543, 544; cf. Russell 1919 p 21f, Grattan-Guinness 1977 p 65; also Dedekind's notion '*abhängig*' in Dugac 1976 p 296, Ewald 1996 vol 2 p 788 footnote c), in a more general setting, the notion "hereditary class", which no doubt leveraged on Frege's 'hereditary property' (van Heijenoort 1967 p 65; Grattan-Guinness 2000 p 327;

[19] Poincaré references Whitehead's paper but not the theorem.

Russell 1906a p 36), is defined like Poincaré's recurrent class. Likewise the notions of ancestor-descendent follow on Poincaré's definition of inductive number. In PM vol II *120 it is with these notions that the inductive numbers are introduced.

A distinction should perhaps to be made here between two problems: The first concerns the place of non-predicative propositional functions and/or the inconsistent classes defined by them, this problem was first identified by Cantor and was later connected to the antinomies. The second is that of circular definitions, extracted by Poincaré from the semantical paradoxes. Cantor accepted the inconsistent sets into mathematics; Russell accepted the non-predicative propositional functions into logic, without any ontological commitment (1906b p 647ff) regarding their classes. Later Russell attempted to solve both problems by the same solution, that of type-theory and the axiom of reducibility. Poincaré had seemingly no hesitation in rejecting all non-predicative definitions/propositional-functions/classes. However, Poincaré's solution was nothing but a mirage, as Russell pointed out (1906b p 633, 644): the notion of impredicativity is itself impredicative (see Sect. 1.7).

Indeed, unlike Poincaré's original attack on the proofs of CBT, which was based on his Kantian philosophy of mathematics implying that the principle of induction has intuitive grounding, his impredicativity argument appears ad-hoc, a reaction drafted to face Zermelo's proof. Poincaré adapted Richard's own explication of his paradox (1905), which he presented under the title "the true solution" (1906b p 309), grafted on it Russell's term 'non-predicative' that was carved in the context of the inconsistent sets, and, enjoying the negative winds against the antinomies, suggested a doctrine to bar impredicativity from mathematics, even when no harm was evidenced by its use. Thus he said, with regard to the lemma of §XI proved above, from Whitehead 1902:

> The reasoning of Whitehead is circular; it is the same that has conducted to the antinomies; it was illegitimate when it gave false results; it rests illegitimate when it conducts, by chance, to a true result.

Poincaré's move can be described in Lakatosian terms (1976 I.4.c) as "strategic withdrawal" to safe harbor, which decreases mathematical content.

Incidentally, Poincaré regarded the origin of non-predicative definitions in the tacit assumption of an actual infinity (1906b p 316; cf. 1906a p 18, 25), an infinite set which serves as the domain of discourse. Signs of this infiltration of the actual infinity Poincaré saw in the use of the word *all* which he italicized in both his proofs of CBT described above. Thus in the second proof, even if we drop the word 'inductive' from the proof, it will still not be acceptable by Poincaré because actual infinite collections are assumed in it, such as the collection of all A_n.

19.5 Criticism of Zermelo's Proof

After finishing off with the second CBT proof and Russell's definition of induction, Poincaré turned to criticize Zermelo's proof. Poincaré presented Zermelo's proof, maintaining his own notation introduced above, as follows:

> Consider all the sets B [subsets of A_0] that contain Q_0 and that contain their own image φ(B) [A_0 is a B]. Let R be the set formed by all the elements common to all the sets B; we see that the image of R, that is to say φ(R), is formed by all the elements common to all the φ(B); these elements are members of all the B, and therefore of R, whence it follows that R, which contains Q_0, contains likewise φ(R).
>
> We show next that R = Q_0 + φ(R); for otherwise Q_0 + φ(R) would be a B set, that does not contain R, because on the contrary it will be its subset. This point established, we see that
>
> $Q_0 + A_1 =$ [20]$Q_0 + (A_1\text{-}\varphi(R)) + \varphi(R) = (A_1\text{-}\varphi(R)) + R$ is equivalent to
>
> $(A_1\text{-}\varphi(R)) + \varphi(R) = A_1$ because φ(R) is equivalent to R[21]; and finally to A_0 because A_1 is equivalent to A_0.
>
> The fault is again the same; R is the common subset of *all* the sets B; to avoid a vicious circle this should mean: to all sets B in the definition of which the notion of R does not enter. This excludes the Q_0 + φ(R) that depends on R. Thus the definition of the set R is non-predicative.[22]

Zermelo's proof is based on Dedekind's theory of chains (1963) and is similar to the proof of CBT which Dedekind devised already in 1887 (see Chap. 24). R is, in Dedekind's terminology, the chain of Q_0. Note that R is the union of the Q_n of Poincaré's second proof, but while in that proof the chain generated by H_0 is "pushed down" by φ to obtain that A_0 is equivalent to $A_1 + Q_0$ here the chain generated by Q_0 is pushed down by φ to obtain that $A_1 + Q_0 \sim A_1$. The difference in the two proofs is that in the first the chain generated by H_0 is constructed by the union of the H_n, all images of each other, while in the second proof the chain is generated by the intersection of a family of sets. These two equivalent definitions of a generated chain originate from Dedekind (1963 #37, #131 p 92).

From the notion of a chain generated by a single element and the assumption that there exists an infinite set, Dedekind devised a model for the theory of natural numbers. In figurative language we can say that the axiom of infinity with its reflexive mapping of the infinite set (a spiral down) entails the axiom of complete induction (a spiral up).[23] It turns out that to prove CBT one needs either to assume the theory of the natural numbers or to assume something equivalently strong. Denying non-predicative definitions, which denies Zermelo's axiom of subsets, is a denial of the alternative gestalt to induction, not a move towards greater rigorousity.

---

[20] Here Poincaré wrote 'where' but the equality sign is implied.

[21] And $(A_1\text{-}\varphi(R))$ corresponds to itself by the identity mapping. Zermelo's proof was thus the first published proof to provide the mapping between the sets found to be equivalent in the thesis of CBT.

[22] We will expand on Zermelo's proof in Chapter 24.

[23] Continuing in this fashion we can say that the antinomies arise when an attempt is made to condense the spiral down into one single loop, as in Escher drawing of a hand that draws a hand that draws it (see Sect. 9.4.).

19.6 Criticism of Cantor's Proof

In his 1883 *Grundlagen*, Cantor proved CBT for sets of the power of the second number-class (II) (see Chap. 1). A crucial step in the proof is the lemma that the power of (II), the set of all denumerable numbers, is different from the power of the set of finite numbers (I). Poincaré (1906b p 308) noted that the argument by which Cantor justified the lemma appears to him "totally like" the argument driving the Burali-Forti paradox. He therefore said that he is not sure that $\aleph_1$ exists. Poincaré thus indirectly criticized the validity of Cantor's CBT proof for sets of power (II). In defense of Cantor, we wish to show that this criticism is not valid.

With regard to the Burali-Forti paradox, Poincaré says (p 307):

> In it a set E is introduced of all ordinal numbers; this should mean all ordinals that can be defined without introducing the notion of the set E itself; thus the ordinal number that corresponds to the order-type defined by this set E is excluded.

The argument, taken from the explication of Richard's paradox, is dubious even with regard to the Burali-Forti paradox and is irrelevant with regard to lemma about the power of (II), namely, $\aleph_1$. Here is why: The proof of the lemma and the argument of the Burali-Forti paradox resemble in that in both a sequent (by Cantor's second generation principle from *Grundlagen*) to a succession of numbers is generated which ought to belong to the succession, hence a contradiction arises (because a sequent is always outside its defining collection). However, there is a difference: in the proof of the lemma, Cantor assumes that (II) is denumerable, so that it satisfies the Limitation Principle. As a result, Cantor derives the (temporary) conclusion that (II) has a denumerable sequent, which ought to belong to (II). The contradiction proves that the assumption cannot be maintained and hence the proof of the lemma that (II) is not denumerable. In the Burali-Forti paradox, however, no similar assumption is made and compliance with the Limitation Principle is disregarded. If an assumption would be made, say that the collection of all ordinals has some aleph as its power, the argument used for the lemma would apply to the set of all ordinals, to the effect that the assumption would be rejected (cf. Sect. 17.3.). From Cantor's perspective, no contradiction is obtained in Burali-Forti paradox because the Limitation Principle is not obeyed. The Limitation Principle requires of a succession to have a sequent that its power is either that of a number-class or that of the union of a singular number of number-classes.

Against Cantor's proof that (II) is not denumerable, Poincaré could have raised the objection that the sequent obtained under the assumption that (II) is denumerable, is obtained by a non-predicative definition (because it would have to belong to (II)). Under this criticism, which could well be what Poincaré had in mind, (II) is accepted but there is nothing beyond (II). This appears to be a conventionalist stratagem (Lakatos 1976 p 99) to block comprehension, which seems to go against Poincaré's own convention that in mathematics consistency is the criterion for existence (Poincaré 1905 p 819). Poincaré may have exhibited here a form of double standard, to protect his Kantian ideology.

19.7 CBT from the Well-Ordering Theorem

Poincaré, determined to bar any possibility to prove CBT except by assuming complete induction, added the following (Poincaré 1906b p 315):

> It is possible to deduce Bernstein's theorem from the celebrated theorem of Zermelo, but we bump into the same obstacle.[24]

Poincaré is referring here to Zermelo's Well-Ordering Theorem (the first proof of 1904). To tie it to CBT Poincaré goes to Cantor's assertion of CBT in his 1895 *Beiträge* as a corollary to the Comparability Theorem for cardinal numbers, which itself is a consequence of the Well-Ordering Theorem. Poincaré's attack is not on the derivation of CBT from the Comparability Theorem for cardinal numbers but on the Well-Ordering Theorem itself.

In Zermelo's proof, Γ is the union of all well-ordered subsets of a set E, the order of the subsets being compatible assuming the axiom of choice. If Γ is not equal to E then it can be extended to the subset Γ' $=$ Γ∪{A}, where A is the distinguished element of E-Γ provided by AC. Γ' being a well-ordered subset of E must be a subset of Γ contrary to the choice of A as a member of E not in Γ. Poincaré claims that the definition of Γ is non-predicative because among the subsets unified to obtain Γ there is the subset Γ' which is defined by reference to Γ.[25] Poincaré obviously rejects the Well-Ordering Theorem because he sees in its proof common elements to Richard's paradox (1905). However, the difference between the two is obvious too: the definition of E in Richard's paradox changes the meaning of "definability" while no similar mutation occurs in the terms of the Well-Ordering Theorem (cf. Goldfarb 1988 p 77).

Poincaré concludes his criticism of the Well-Ordering Theorem with an interesting remark:

> Though I am inclined to accept the axiom of Zermelo, I reject his demonstration [of the Well-Ordering Theorem], which for a moment made me think that aleph-one could exist.

Why does the Well-Ordering Theorem imply the existence of $\aleph_1$? Presumably because if all cardinals are comparable and there are non-denumerable cardinals (the continuum), there is in particular the cardinal next following $\aleph_0$.

In the second part of his 1908a paper, Zermelo answered the various objections raised against his first proof of the Well-Ordering Theorem. There he also answered Poincaré's objection to the impredicative definition in his first proof of the Well-Ordering Theorem, which, Zermelo notes, could similarly apply also to his second

[24] Medvedev (1966 p 239) notes that Zhegalkin (1907) made such a deduction. We have not seen Zhegalkin 1907.

[25] Again we note the duality symmetry between the impredicativity identified in the Well-Ordering Theorem that employs union and the impredicativity in Zermelo's proof of CBT that employs intersection. Note that a definition can be impredicative when it explicitly uses neither intersection nor union. Thus the definition of Nc_fin given in Sect. 15.2. is impredicative because this class is among the classes s mentioned in it.

proof (the objection to the CBT proof is not mentioned). First Zermelo points out that impredicative definitions have been used for a long time in mathematics without any objection, as an example he cites Cauchy's proof of the fundamental theorem of algebra.

Zermelo's second argument is, similar to an argument raised by Russell (1906b p 633), that the notion of impredicativity is itself impredicative, "for, unless we already have the notion, we cannot know at all what objects might at some time be determined by it and would therefore have to be excluded". Thus, this very notion cannot be used to discredit notions that can be qualified by it.

Zermelo thirdly agrees that the question whether an object is to be subsumed under a definition needs to be decidable by an objective criterion, namely, such that does not refer to the notion defined; but once such a criterion is given, "nothing can prevent some of the objects subsumed under the definition from having in addition a special relation to the same notion and thus being determined by, or distinguished from, the remaining ones – say, as common component or minimum. After all, an object is not created through such 'determination'". Thus R, from Zermelo's proof, is among the B sets by an objective criterion, and Γ, of the first proof of the Well-Ordering Theorem, is among the well-ordered subsets, and besides this R is minimal and Γ the maximal among the sets fulfilling the criterion.

Finally, Zermelo points out that an object may have different determinations and these can be equivalent, namely, have the same extension: "Indeed, in every definition the definiens and definiendum are equivalent notions, and the strict observance of Poincaré's demand would make every definition, hence all of science, impossible." This same point is made by Russell when he explains the intricacies of the liar paradox and the way the axiom of reducibility can bypass the problem of self-reference by a bounded variable (Russell 1906b p 648).

The debate between Poincaré and Zermelo continued in the context of Zermelo's definition of the finite numbers and the principle of induction. Our impression is that Poincaré had eventually backed-away from his dogmatic stance against non-predicative definitions. In his 1910 paper (p 1074 [10]) CBT, in its single-set formulation, is mentioned briefly and Poincaré says that if the given correlation between M and M'' is predicative, so is the constructed correlation between M and M'. Thus, Poincaré no longer thought that the proof itself inserted non-predicative noise.

Chapter 20
Peano's Proof of CBT

Shortly after Poincaré provoked Couturat to produce a proof of CBT that does not use complete induction, Peano[1] proposed such a proof in an article titled "*Super* [on] *theorema* de Cantor-Bernstein" (Peano 1906, dated March 1906 and published May 17, in the *Rendiconti del circolo matematico di Palermo*).

Peano's article is written in two languages invented by Peano. The first is the sign language (pasigraphy) in which Peano writes mathematical propositions. The second language is his language of discourse "*Latino sine flexione*", one of the first artificial languages. The more complex propositions in the paper Peano translates from the sign language to the language of discourse[2]; perhaps specifically for Poincaré who expressed his ignorance (and in fact – dislike, 1905 p 822ff) of Peano's sign language. We will partly replace Peano's pasigraphy with the now current sign language used in Levy 1979.

Despite his quick reaction to Poincaré's challenge Peano was not quick enough: Zermelo, already in January 1906 (Zermelo 1908a p 191 footnote 8), had communicated a similar proof to Poincaré, who published it in his 1906b paper (see Chap. 19). The 1906b paper also appeared in May; it criticized Zermelo's proof for its use of impredicative definition and it contained no reference to Peano's proof. This may have provoked Peano to republish his paper (Peano 1906a), this time in *Revista de Mathematica* – a journal that he himself had founded, with an added appendix, titled *Additione* and dated August 23, 1906.

There are two main subjects in the *Additione* and both concern Zermelo. First Peano discusses Zermelo's axiom, as the axiom of choice was called in those days. He indicates his priority (1890) in noticing the problem with infinite number of arbitrary choices and the reason for his rejection of such possibility: making an arbitrary choice in an assertion calls for a specific clause in that assertion in any formal rendering of the argument; making infinitely many arbitrary choices means

---

[1] On Peano see Kennedy 1980 and Kennedy 2002.

[2] The translated text Peano puts between double quotation marks (") which we render, when we quote from Peano, by an apostrophe (').

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_20,

that infinitely many clauses must be added; but our languages are all finite (cf. Moore 1978 p 317). Peano ignores Zermelo's point that with the axiom of choice only one arbitrary choice is made: that of the choice function (cf. Korselt 1906 p 218; Sierpiński 1918 p 210) from the collection of the relevant choice function that is not empty by the axiom of choice. As the axiom of choice is not relevant to Peano's paper it seems that Peano brought the subject up only to discredit Zermelo in a possible priority dispute regarding the solution to Poincaré's challenge. (Peano did not know when Zermelo communicated his proof to Poincaré.)

The second subject is Peano's objection to Poincaré's criticism of impredicative definitions. Peano brings several examples to such definitions in traditional mathematics since antiquity, to which no objection ever rose. Peano does not explain why he brings this subject up but it seems doubtless that he wanted to defend his proof against the objection raised by Poincaré against Zermelo's proof, which could likewise be directed at his own. At this part of his article Peano did not mention Zermelo.

Zermelo was quite aware of Peano's attitude towards him and in his 1908a paper, after he brought his own criticism of Poincaré's stance against impredicative definitions (p 190f), he said (p 191 footnote 8): "Why does Peano avoid mentioning my name here, where I agree with him, and then direct his opposition to the principle of choice… so expressly at me?" Hinting at the possible reason for Peano's behavior he continued: "It would seem obvious, after all, that not mathematical principles, which are common property, but only the proofs based upon them can be the possession of an individual mathematician".[3] Zermelo also noted the date of his letter to Poincaré, perhaps in order to subdue any conspiracy theory by Peano.

Interestingly, there was a third proof at the time suggested in response to Poincaré's challenge, by J. Kőnig (1906, see the next chapter). Peano mentions J. Kőnig's paper in his *additione*, but does not say a word about its relevance to Poincaré's challenge. Surely, Peano was not concerned regarding a priority dispute with J. Kőnig, because the first printing of his proof was earlier and because his and J. Kőnig's proofs were different.

20.1 Peano's Inductive Proof

Peano begins his paper with a statement of CBT in its cardinal form that he says comes from Cantor 1895 *Beiträge* p 484:

> If x, y are two cardinal numbers such as $x \geq y$ and $x \leq y$ then $x = y$.

[3] Zermelo is using here language akin to the legalistic perception of patents, which must be the application of principles and not principles themselves. Zermelo was indeed to register several patents in 1911 and in the 1930s (Ebbinghaus 2007 p 118).

In the form cited by Peano, CBT does not appear where referenced or elsewhere in Cantor's writings. The Comparability Theorem for cardinal numbers given by Cantor in the referenced place is indeed in the language of cardinal numbers but the three formulations of CBT presented there as its corollaries (B, C, E) are in the language of sets and mappings (see Chap. 4). Moreover, Cantor never defined or used the relation $\leq$ between cardinals.

Another oddity is that in the cited statement of CBT Peano refers to x, y as cardinal numbers, namely, as entities; however, in what follows, he, like Borel, defines cardinal numbers only in their equality and inequality ($\leq$) relations and not as entities. Thus, after introducing these notions, Peano translated the theorem to his pasigraphy and shifted it to the single-set formulation in the language of sets and mappings, saying:

> we can eliminate the idea of cardinal number and after substituting the defined terms by the defining symbols Cantor's Theorem receives the following form:
>
> (1) $(a, b, c \in \text{Cls} \wedge c \subseteq b \subseteq a \wedge g \in (\text{cFa})\text{rcp}) \rightarrow \exists(\text{bFa})\text{rcp}$.
>
> 'if a, b, c are three classes, and class a contains class b which contains c, and if g is a reciprocal function, or correspondence, between a and c, then there exists a reciprocal correspondence between a and b.' [4]

It seems that Peano presented CBT in the language of cardinal numbers in order to exhibit how cardinal numbers in their order relations are introduced in his formalism and how 'nominal definitions' can be eliminated (contrary to 'definitions by postulates' which cannot be eliminated – see below).

With regard to the theorem (in its changed version) Peano said:

> Cantor never proved this theorem. Bernstein published a demonstration in Borel.

Peano is twice mistaken here as we have seen in Chaps. 1 and 11: Cantor did prove CBT and Borel's proof was adapted from the proof of Bernstein and not its exact rendering.

"After partial introduction of symbols and various modifications", says Peano, Borel's "demonstration assumes the following form":

> If u is a subclass of a, ... gu denotes the image under g of u. Thus we can consider the sequence of classes u, gu, g^2u, g^3u, ..., g^nu,.... We define by Zu their logical sum [union]
>
> (2) $u \in \text{Cls'a} \rightarrow \text{Zu} = \cup\{g^n u \mid n \in N_0\}$ Df.[5]
>
> Now a can be partitioned into three parts: $a' = Z(a\text{-}b)$, $a'' = Z(b\text{-}c)$, and $a''' = \cap\{g^n a | n \in N_0\}$.... b [of (1)] is also divided into three parts: $b' = g(Z(a\text{-}b))$, $b'' = a''$, $b''' = a'''$. g is a one-one correspondence between a' and b'; the identity transforms a'' onto b'' and a''' onto b'''. Therefore Bernstein's proof assumes the form:

[4] Peano uses 'class' where Cantor uses 'set; this seems to have no relation to the issue of the antinomies. Peano uses 'reciprocal' and 'one-to-one' interchangeably as well as 'reciprocal function' and 'correspondence'. Peano notes that the use of the signs $\in$, $\subseteq$, $\exists$, etc., in (1) places it in the science of 'mathematical-logic'.

[5] Peano is using Cls' for Russell's Cls'. Peano writes n'N_0 for $n \in N_0$. N_0 appears to be Peano's sign for the class of natural numbers including 0.

(3) The hypothesis of (1), and (2), imply that g|Z(a-b)∪1|(a-Z(a-b)) is a 1–1 mapping between a and b [so the thesis of (1) is obtained].[6]

What Peano denotes by Zu Dedekind denoted in *Zahlen* by u_0 and called the chain of u. Peano's gestalt is that a, b are composed of two chains of equivalent frames and a residue, all carved by the same mapping g. This gestalt is similar to the one in Poincaré's second proof by induction (see Sect. 19.2). Only here, without the drawing we took from Fraenkel, the gestalt can be perhaps better described as: two staircases leading to a common landing, like in the old opera house in Paris. Though Peano discerned two chains in a, b, in the proof he used the gestalt and metaphor used by Dedekind (see Chap. 9): he partitioned the sets into a chain and a complement (Z(a-b), a-Z(a-b)) and pushed down the chain, combining g on one partition with the identity on the other. Peano had obtained this, not yet published at the time, feature in CBT proofs, independently of Zermelo who had obtained it in the proof he submitted to Poincaré (see Sect. 19.5).

Of course, Peano did not know of Dedekind's proof, which was published only in 1932. Still, it is reasonable that Peano had obtained the chain gestalt from Dedekind's *Zahlen* (1963), though there is no direct evidence that he did. On the other hand, the gestalt and metaphor of Peano are clearly different from those of Borel, who rearranged the pairwise equivalent frames, each by a different mapping. Yet Peano says that his proof is a representation of Borel's proof. It is possible that Peano had reached his proof while processing Borel's proof against elements he had obtained previously from Dedekind's *Zahlen*, though, Dedekind is not mentioned in Peano's paper.

Following the above proof Peano says: "Identical demonstration occurs in (Schröder 1898)" and he added, "Schröder, in defining Z, introduces a sequence of numbers N_0 and 'limit'". Peano demonstrates here a very relaxed way of using the word "identical". Schröder, who attempted to prove the two-set formulation of CBT without gliding to the single-set formulation as Borel and Peano, did use complete induction but he never mentions the Z operator or the chain gestalt. Schröder was not aware of the sequences of frames (generated by a-b and b-c) and though he did see the emergence of the equivalent residues, his limit idea is a joke: he claimed that $a \sim a''' \sim b$. It must be that when Peano said 'identical' he did not refer to the text but to the image he construed in his mind when he browsed the text, the effects of his proof-processing of the text against the background of Borel's proof.

Similarly, when Peano continued and said that Poincaré "reproduced the preceding demonstration", referring to the first of Poincaré's proofs, we must interpret "reproduced" with leniency. After all, Poincaré's algorithmic shoe-lacing proof is in the two-set formulation, it does not contain the chain gestalt or the pushdown metaphor and so it is different from the proof of Peano.

[6] We have replaced in (3) some of Peano's pasigraphy with his words. 1 is the identity mapping denoted 'idem' by Peano. | is the reduction of domain sign.

All these examples teach us that mathematicians do not read a mathematical text as it is but through a medium of associations that they have acquired in their past proof-processing experience.

20.2 Addressing Poincaré's Challenge

Peano presents Poincaré's challenge, in his own pasigraphy terms, as follows: In (1) only logical sign occur[7] while in (2) the arithmetical sign N_0 occurs; so the challenge is to provide a proof from which the sign N_0 is eliminated. À priori, says Peano, it is doubtful that N_0 can be eliminated because it is not introduced by a nominal definition but by five postulates.[8] Nevertheless, Peano asserts, in the particular case proposed by Poincaré, the elimination is possible.

In this form the challenge falls under Occam's razor and the methodological directive of Craig's Interpolation Theorem, rather than the original context of intuition vs. logicism from which Poincaré's criticism emerged. However, not only is the context shifted, the entire backdrop is changed (history is rewritten while still in the making): Poincaré did not pose any problem; a problem is posed from a different set of mind and with different expectations regarding the possible result. Poincaré did not expect an answer when he sarcastically challenged Couturat; he expected that no one would be able to meet the challenge. Moreover, Poincaré was critical of the logicist reducibility position and he would not pose a problem formulated in the terminology of the point of view that he opposed. Poincaré maintained that intuition is the only way to establish the number concept and complete induction and that these two are necessary for the proof of CBT. Peano was surely aware that when he answered Poincaré's challenge, not only did he solve a mathematical problem but he also refuted Poincaré's anti-logicist stance. In Lakatosian terms (1976) this case can perhaps teach us that when intuition is called in a proof, proof-analysis in search of a hidden lemma is necessary, to eliminate that call.[9] This is the road to axiomatization.

To eliminate N_0 from the definition of Z Peano gave Z a new definition:

"(4) When $u \subseteq a$, Zu $= \cap\{v|\ gv \subseteq v$ and $u \subseteq v\}$". [As the domain of g is a, $v \subseteq a$ is implied.]

---

[7] At the time, the notions of set, mapping, membership, etc., where considered logical notions. With the axiomatization of set theory by Zermelo these notions became part of set theory just as N_0 belongs to arithmetic.

[8] Peano refers to his system of axioms as presented in his *Formulario* and not as in his 1889 paper where the number of axioms is 19.

[9] J. Kőnig, though he was a formalist and thus a reductionist too, was more attentive to Poincaré's ideological arguments. It would be interesting to see if Lakatos' "Proofs and Refutations" could be broadened to cover also cases of "Ideologies and Refutations".

It is quite amazing that Peano does not mention Dedekind here, for Dedekind (1963 #44) defined by (4) the chain of a subset *u* of *a*. Dedekind's definition preceded his construction of the natural numbers which led him to show (#131) that definitions (4) and (2) give the same result. Anyway, Peano's definition is the same given by Zermelo and published by Poincaré in his 1906b (see Sect. 19.5).

To prove (1) Peano simply states:

(5) The hypothesis of (1), and (4), imply that
$g|Z(a\text{-}b) \cup 1|(a\text{-}Z(a\text{-}b))$ is a 1–1 mapping between a and b.

As Peano notes, (5) and (3) are identical, and he, interestingly, adds: "Verification of the thesis is intuitive" referring no doubt to the thesis of (5). It is not that Peano invoked here Poincaré's intuition to verify his thesis in (5); rather he seems to refer to the argument that since we have a partitioning of a, a $=$ Z(a-b) + (a-Z(a-b)) and a partitioning of b, b $=$ g(Z(a-b)) + (a-Z(a-b)), then the mapping identical with g on Z(a-b) and with the identity on (a-Z(a-b)), gives a mapping from a onto b.

Peano then added: "we can decompose the affirmation of the thesis into elementary affirmations, and determine many rules of logic that are applied in an implicit way in the preceding demonstration." In 24 steps Peano then provided a formal proof of assertion (5), which we will omit.[10]

20.3 A Model for Arithmetic

Peano now opens a new section in the (otherwise section less) paper, starting with the words: "from the presented formula I deduce new consequences". Maintaining the discussion in logic, so that the notions 0, N_0 and + have no assigned meaning, Peano adds a new assumption, that a-b is not empty, and he denotes by 0 an arbitrary element of a-b. Peano then defines $N_0 = Z(\{0\})$ and for every $x \in N_0$ he defines $x+ = gx$. Then Peano proves that N_0 satisfies his five axioms of arithmetic. Peano concludes:

> This proves (if proof is necessary) that the postulates of Arithmetic do not involve contradiction.

The remark in parenthesis, which goes against expressed views of Hilbert (1904) and Poincaré (1905, Poincaré 1906a), is not a slip of tongue; Peano continues with some very interesting remarks regarding the need for consistency proofs for the basic branches of mathematics:

> But proofs that systems of postulates of Arithmetic or Geometry do not involve contradiction, are not, I believe, necessary. For we do not create postulates in arbitrary, we chose as postulates simple proposition, written in explicit or implicit mode, in all works of Arithmetic or Geometry. Our analysis of the principles of these sciences reduces common

[10] This is an example to the elimination of intuition by axiomatization mentioned above.

> affirmations to a minimal number, necessary and sufficient. Systems of postulates of Arithmetic and Geometry are satisfied by the idea that number and point have in all writings of Arithmetic and Geometry. We think of number therefore number exists. Proof of the coexistence of system of postulates may have utility, if the postulates are hypothetical, and do not correspond to the real facts.

Peano expresses here the mentalist variant of Platonic ontology: instead of existing ideas imprinting our thoughts, our thoughts imprint existence. Thoughts are real and therefore can be used in a mathematical model. His view is reminiscent of Dedekind's view with regard to his infinite set of all possible thoughts and Cantor's definition by abstraction. Mentalism, however, seems to differ from the realism of Russell (1906a p 41; cf. Grattan-Guinness 1977 p 124) and Gödel (1944).

Strangely, Peano is not aware that by this demonstration of a model of arithmetic actually undermines his claim that his proof of CBT is free of any reference to the number concept or complete induction. His proof shows that under the assumptions of CBT and the non-emptiness of a-b, the entire structure of axiomatic arithmetic can be constructed, which entails that the assumptions of CBT are as strong as arithmetic and therefore the attempt to bypass arithmetic in the proof of CBT is futile. This is in fact the correct answer to Poincaré's challenge (see the end of Sect. 19.5).

Another amazing point is that Peano seems to be totally unaware that his "new consequences" are nothing but the construction of simple chains as models of arithmetic in Dedekind's 1888, *Zahlen* (1963). And it is not that Peano never read Dedekind; Peano references Dedekind's *Zahlen* in his monograph on the Principles of Arithmetic (van Heijenoort p 86; cf. Ferreirós 1999 p 251 on Dedekind's influence on Peano). Could it be that Peano reproduced the content of the *Zahlen* without becoming aware of its origin? This phenomenon is known to occur with composers so it is possible that it occurred to Peano. We call this "Peano *Obliviose*", which in *Interlingua* (an offspring of *Latino sine flexione*) means "forgetful Peano". It may indicate that the conscious workings of proof-processing, affects the subconscious without necessarily leaving a conscious association trail.[11] Poincaré (Newman 1956 p 2041) described several incidences where his mathematical thinking was subconscious. This subject requires more study (see Hadamard 1954).

[11] Peano's generation of Dedekind's CBT proof from Borel's, mentioned earlier, is another example.

Chapter 21
J. Kőnig's Strings Gestalt

In response to Poincaré's challenge, two more proofs of CBT were produced in 1906, by J. Kőnig and by Felix Bernstein. Both proofs were presented by Poincaré in sessions of the French Academy of Science. J. Kőnig's paper was published in the proceedings of the academy for July 1906, and that of Bernstein in December 1906. In both papers CBT is called the Equivalence Theorem; J. Kőnig adds "of Mr. Cantor" and Bernstein "of set theory". It is likely that the authors did not actually present their papers in person and that Poincaré merely presented the papers for publication in the proceedings. In this case the delivery of the papers to Poincaré was surely accompanied by some letter exchange between the authors and Poincaré but we know of none.

As Poincaré presented the papers to the academy, one is inclined to assume that Poincaré agreed with their contention that they indeed answer his challenge. However, since it is quite obvious that these proofs do make use of natural numbers and complete induction, this supposition cannot be maintained. So it appears that Poincaré sponsored the presentation of the papers because he valued their mathematics, or at least their authors.

J. Kőnig planned "to make [CBT] evident without employing the concept of number", nor "the principle of complete induction, which, as Mr. Poincaré correctly remarked (1906), all the proofs published until now did employ". Describing the circumstances that led him to present his proof J. Kőnig noted: "I did not want to give it except in the exposition of synthetic Logic, that I hope to publish soon and that I already given in my course of this year. But the interest taken today in these things made me publish this note". Kőnig does not say when he had found the proof but what he says about his course, which was probably given in Budapest not in Paris, can equivocally indicate that he presented the proof in that course. In this case it is possible that Kőnig hinted at the early origin of his proof so that he could be considered to have priority, over Zermelo and Peano, in meeting Poincaré's challenge, which may have raised at the time some interest in mathematical circles.

J. Kőnig's proof certainly merited Poincaré's attention. It brought a new gestalt to CBT proofs which had "remarkable generalizations" (Fraenkel 1966 p 77 footnote 1) in new contexts that could not have been foreseen by either J. Kőnig or

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_21,

Poincaré. It was thus a good thing that J. Kőnig did not wait to include his CBT proof in his planned book, for the book appeared only posthumously in 1914 (cf. Franchella 2000), titled *Neue Grundlagen der Logik, Arithmetik und Mengenlehre*. The editor of the book was J. Kőnig's son, D. Kőnig, who in a series of five papers (1908, 1914(1923), 1916, 1926, Kőnig-Valko 1926, see Chap. 22), leveraged on his father's 1906 gestalt, to produce results in set theory, graph theory, and other branches of mathematics (cf. Franchella 1997). In the 1829s, J. Kőnig's gestalt was further employed in the hands of members of the Polish school of logic (see Chaps. 28, 32, and 34).

Bernstein's paper, on the other hand, turned out to be of little significance. It is in fact a variant of Peano's first proof, though Bernstein correlated his proof to J. Kőnig not to Peano. It is articulated so as to conceal its use of natural number and complete induction. Bernstein somewhat twisted Poincaré's position when he said that to prove the principle of complete induction it is indispensible to prove first CBT without using that principle. Actually Poincaré thought that the principle of complete induction cannot be proved and that CBT can be proved only by way of complete induction. To answer Poincaré's challenge as he understood it, Bernstein demonstrated how his original proof, as dressed by Borel, can be modified to avoid complete induction. Bernstein mentioned Zermelo's 1906 proof and the similar proof of Peano 1906, and the "irrefutable criticism" of Poincaré against the use of chains in these proofs.[1] He pointed to J. Kőnig's proof as the first to answer Poincaré's challenge. From that proof he borrowed an argument (see below) to justify his contention that his proof does not use the natural numbers or induction. For this reason we discuss Bernstein's 1906 proof in this chapter.

21.1 J. Kőnig's Ideology

J. Kőnig's paper is only two pages long. It has clear mathematics[2] but is vague regarding J. Kőnig's position on the Poincaré debate against the logicist. Thus with regard to CBT Kőnig says: "the Equivalence theorem is a theorem of intuition", which is Poincaré's claim, while he also declares that his intention is to demonstrate it without appeal to the notion of number or complete induction, the basic intuitive notions of Poincaré. Further J. Kőnig states that: "with regard to the concept of number, it is quite true that we need to construct it ourselves", which is the stance of the logicist and formalist schools.

J. Kőnig was apparently not a logicist for he did not accept the notion of set as a fundamental, logical notion; he says: "To demonstrate it [CBT], I will employ the terminology of Mr. Cantor; while underlining at the same time that a broader and

[1] Yet another sign of the rivalry between Bernstein and Zermelo (Peckhaus 1990 p 48, Ebbinghaus 2007 §2.8.4).

[2] Fraenkel (1966 p 77 footnote 1) complements J. Kőnig's CBT proof for its lucidity.

more precise exposition cannot further use the words set, etc." By 'etc' Kőnig could have meant the notions of 'mapping' and 'element of a set' that he uses in the proof. His position is somewhat clarified when he says that[3] "the concepts *follow* and *sequence* need be well accepted as definitive logical concepts", and in the closing lines of the proof, which also end the paper: "It goes without saying that this exposition has still many inconveniences; because we have not discussed in depth the logical concepts here found. Such are also the expressions *to the right* or *to the left"*. These words echo Hilbert 1904, which established the formalist approach (van Heijenoort p 129).

A more obscure statement that can nevertheless be construed to further explain J. Kőnig's views is this: "The spiritual and profound criticism of Mr. Poincaré is irrefutable, so I believe, in its negative parts. That which we have come to possess until now was perhaps necessary for the development of the new science of logic; but certainly this does not give what we seek: the bases of this new science." By the "negative part" of Poincaré's criticism Kőnig could have meant Poincaré's criticism that previous attempts to define the number concept failed and his point that the available proofs of CBT use complete induction. With this criticism J. Kőnig agreed. By the positive part of Poincaré's criticism Kőnig may have meant the view that the number concept and complete induction are grounded upon intuition and cannot be reduced to simpler notions. With this view Kőnig seems not to agree. He seems to say that even if natural numbers and complete induction were used for developing the new science of logic (what "we have come to possess"), they cannot serve as its base, as would, in his view, the aforementioned formalist notions of relative position.

Still, J. Kőnig did not expel intuition from reasoning for he added: "There are quite a number of things in the immediate intuition, a *deed lived* or an *experience*; but this residue is of total necessity". Here Kőnig seems to shift intuition from the role given to it by Poincaré in basing mathematics to a role in basing epistemology. Perhaps he saw the basic notions of logic as emerging from sensory feelings (cf. Franchella 2000).

J. Kőnig thought highly of his proof,[4] for he says: "The new demonstration of the Equivalence theorem of Mr. Cantor that I will give in these lines has, so I believe, quite a great importance, in view of the present discussion on the foundation of logic, mathematics and the theory of sets". Kőnig may have had in mind the view that since his proof grounds CBT in formalism it thereby allows large parts of Cantor's set theory into formalism as well. As it happened, the scene soon changed with Zermelo's axiomatic set theory. With the growing interest in symbolic logic (Whitehead-Russell PM), interest in formalism had faded for a while; then came Gődel.

---

[3] Italics in the original.

[4] Kőnig's self-praise and covert priority concern provide us with some clues on his personality. He emerges as a passionate, romantic person that sees much drama in the development of ideas. An expression of his character we find in his already quoted description of Poincaré's 1905-1906 article as "spiritual and profound". It seems that Kőnig had the feeling that great strides are made in his time towards important mathematical truths and he believed that his contributions had an important role in the events. He certainly took center stage with his 1904 Heidelberg address.

21.2 J. Kőnig's CBT Proof

We quote the proof in full. In the footnotes we add our comments, with more comments in the following section.

> Let X and Y be two determined sets,[5] X_1 and Y_1 subsets of X and Y respectively. We have to prove that, given $X \sim Y_1$ and $Y \sim X_1$, we always have $X \sim Y$.[6]
>
> The proposition $X \sim Y_1$ signifies the assumption of the following law (I)[7]: an arbitrary element x of X determines one and only one element y of Y[8]; thus this y determines also the corresponding x.[9] However, there are one or many elements of Y that don't figure up in this law.[10]
>
> Similarly, the equivalence $Y \sim X_1$ signifies the assumption of a law (II), that it will be superfluous to detail too.
>
> Take then an arbitrary element x_1 of X; by (I), a determined element y_1 of Y is given to us; this element y_1 gives us, by the law (II), a determined element x_2 of X_1, etc. In so doing, we are not *counting*; there is nothing but the employment of the signs 1, 2, ... for distinguishing the elements of X. But the concepts *follow* and *sequence* need be well accepted as definitive logical concepts.[11]
>
> Thus the sequence[12] $x_1y_1x_2y_2\ldots$ can always be continued to the right, but not always to the left. If x_1 is an element of X_1, the law (II) gives an element y_0 that immediately precedes x_1 in the sequence; but if x_1 is an element of X, which is not an element of X_1, the sequence cannot be further continued to the left.
>
> It is seen therefore that the possible cases are three: The sequence starts with an element of X. The sequence starts with an element of Y. The sequence can always be continued to the left. The elements x_1' and x_1'' of X thus give us two corresponding suites:
>
> (1) $x_1'y_1'x_2'y_2'\ldots$
>
> (2) $x_1''y_1''x_2''y_2''\ldots.$
>
> If there is a common element in the sequences (1) and (2), the element that follows it is determined by the law (I),[13] as a result it will be the same in the sequences (1) and (2), likewise with the preceding element if such exists.[14]

---

[5] The term 'determined' seems to stand for Cantor's term 'well-defined'. It is not a condition of the theorem but on the domain of discourse.

[6] This is the two-set formulation of CBT. The notion of 'equivalence' (~) is not explicitly introduced.

[7] Like Cantor, and most everyone else at the time, Kőnig insists that a 1–1 mapping must be by a law.

[8] 'determines' here stands for 'corresponds to'; it has some vague relation with the previous 'determined'.

[9] By the inverse mapping which exists since equivalence is 1–1.

[10] Namely, it is assumed that the subsets are proper subsets.

[11] This statement hints at J. Kőnig's formalism. See the previous section.

[12] We prefer the term 'string'.

[13] Only if this element belongs to X; otherwise it is law (II) that will determine the following element. To remove ambiguity X and Y must be assumed disjoint, a point ignored in the proof.

[14] The uniqueness of the following and preceding elements results from the 1–1 nature of the given equivalences.

That is to say: An arbitrary element of X determines always the corresponding sequence.[15] It is not necessary to detail the special case of a periodic sequence.[16] It is evident, that a periodic sequence can always be continued to the left.

The law of equivalence, the expression of which is X ~ Y, is found determined by these considerations. Let x be an arbitrary element of X; we have the instruction for forming of the corresponding sequence. If this sequence starts with an element from X, or it can be continued to the left, we choose as the element in Y corresponding to x to be the element that follows x in the sequence. If, however, the sequence starts with an element from Y, we set as corresponding in Y that which immediately precedes x in the sequence.

Thereby the equivalence X ~ Y is fixed. Pure intuition directs us to recognize its existence.[17]

It goes without saying that this exposition has still many shortcomings; because we have not discussed in depth the logical concepts here found. Such are also the expressions *to the right* or *to the left*.[18]

21.3 More Comments on the Proof

An explicit statement that every *y* will be in some string is missing from the proof, no doubt because it is trivial: If *y* belongs to Y_1 it appears in the sequence of the *x* that corresponds to it under the first law; if *y* does not belong to Y_1 it belongs to the sequence of the *x* it corresponds to under the second law.

Obviously, the roles of X and Y in the proof can be interchanged to provide a mapping from Y to X. Note that while the correspondence between X and Y is seemingly defined by Kőnig on a "per element" basis, actually it is defined on a "per string" basis: all the elements of a string are decided at the same time as to the way they correspond: either to the left or to the right. Note further that there is no redundancy in the condition of the theorem: though only one of the given mappings is used in the definition of the mapping for a certain string, the other mapping is still necessary to establish the entire string and for the definition of the law within some of the other strings.[19]

The 1–1 nature of the mapping between X and Y follows from the 1–1 nature of the given mappings and the disjointed nature of the strings. Note that this law is not instructive in the sense that given an element there is no way of telling how its corresponding element is defined. One needs to know to what type of string it

[15] This is true of any element of X + Y: it can belong to only one string.

[16] A periodic string consists of a repeating finite string.

[17] This statement is perplexing for it fits more Peano's mentalistic view regarding mathematical ontology than Hilbert's axiomatic approach. Perhaps J. Kőnig used here 'intuition' loosely, to signify that the discourse was not formal, but "naïve".

[18] Kőnig actually uses the terms interchangeably with "follow" and "precede". This statement too hints at Kőnig's formalism.

[19] For this reason Mańka-Wojciechowska (1984 p 194) classify this proof as a back-and-forth argument. Cf. Silver 1994 and the back-and-forth argument in Sect. 32.2.

belongs. The situation in J. Kőnig's proof in this regard, however, is not worse than with any of the other proofs of CBT.

Kőnig binds the periodic strings with the left-extendible ones so that they do not require special attention when defining the mapping between X and Y. Alternatively, periodic strings, which can be regarded as finite closed loops, could be coupled with the not-left-extendible strings in the definition of the mapping between X and Y because they are always extendible to the right. The same applies to the left-extendible strings, which can be regarded as periodic strings with infinite period (as a straight line is regarded a circle in projective geometry). Also, we can, in either of the cases or both, correspond to x the y preceding it. The arbitrariness of assignment cannot be maintained if we add the demand that of the two possibilities of assignment one is preferred, say because of some affinity ("love") between the elements.

Note that the strings are equivalence classes for the equivalence relation 'pertaining to the same sequence'. Equivalence relations and their related classes were not a new subject in 1906 but it seems that the rhythm of their introduction (through their reflexivity, symmetry and transitivity properties),[20] taking stage with the development of mathematical structuralism, was not yet common at that time. So Kőnig did not care to point out this essential structure within his proof.

We do not think that J. Kőnig's proof withstands Poincaré's criticism of previous CBT proofs. Indeed, Fraenkel (1966 p 77 footnote 1) classified Kőnig's proof among the proofs that rest upon properties of the positive integers. This was also Zermelo's view (1908b p 209 footnote 11). One cannot agree with J. Kőnig when he says that he is not counting when he attaches signs to the elements of the strings. For what is counting if not the attachment of signs to the elements counted? And how do you generate infinitely many signs or define the general term of a string, without induction? It is hardly conceivable that any formalistic grounding of his theory would change our judgment. J. Kőnig's view was repeated by Couturat (1906 p 246 published in November). Interestingly, Couturat added that complete induction has in a proof of CBT the same status as an auxiliary construction has in a geometric proof, in the sense that although it does not figure in the theorem it figures in its proof. The analogy, charming as it is, seems irrelevant for there is no demand for a geometry without auxiliary constructions while there was a demand for CBT without complete induction (cf. Medvedev 1966 p 237). Furthermore, complete induction is directly employed in J. Kőnig's proof: the observation that every element of a string determines the string requires complete induction to prove that the nth following or preceding element is uniquely determined. The stated uniqueness of the preceding or following elements of any given element is used at the induction step.[21] So we conclude that J. Kőnig's claim that his proof meets

[20] Originating with Peano (Grattan-Guinness 2000 p 240).

[21] Use of complete induction could be avoided using impredicative definition for the strings as in Banach 1924 (Chap. 29).

Poincaré's challenge, cannot be maintained. Franchella's (1997 p 5 footnote 9) acceptance of J. Kőnig's claim seems to be unwarranted.

21.4 Bernstein's 1906 Proof

The proof can be summarized as follows: If A corresponds by a 1–1 mapping φ onto $T \subset A$ and U contains T and is contained in A, then setting $R_1 = A\text{-}U$[22] and $R_{n+1} = \varphi(R_n)$, and taking ψ to be φ on the union R of the R_n and the identity on A-R, a 1–1 mapping is defined between A and U.

Bernstein sees "the capital point where there is difference between the previous demonstration [Borel's] and the new demonstration" in his definition of R. Because instead of creating in an inductive process the pairwise denumerably many equivalent frames, and proving by complete induction that the frames are all disjoint (and, we may add, that their union with the residues gives the sets A and U), we have in the new proof the equivalence of R and $\varphi(R)$ given directly. Bernstein adds that the definition of R is not impredicative because R is not among the R_n. Further, Bernstein explicitly relies on J. Kőnig's position (see the previous section) that the use of numbers as indices of the R_n is only as signs; no logical argument (namely, induction) is carried over the indices and their logical quality does not furnish anything to the proof. Thus Bernstein believed that the proof stands to Poincaré's challenge.

How Bernstein could disregard the inductive definition of the R_n is hard to comprehend. And he had in front of him Peano's identical proof where Peano acknowledged its reliance on complete induction! It is perhaps a point of historical justice that just as reliance on a theorem of Bernstein led J. Kőnig to his mistake in refuting the continuum hypothesis at the Heidelberg congress (Dauben 1979 p 247), so reliance on J. Kőnig led Bernstein to his mistaken view regarding his CBT proof of 1906.

In passing let us note that Bernstein makes an interesting remark with regard to R_1; he says that both the new and the original proofs share in that they lead to the conclusion that R_1 can be neglected. This remark seems to support our view that Bernstein's original proof was by abstraction (see Sect. 11.2).

21.5 Comparison with Earlier Proofs

J. Kőnig's proof[23] is for the two-set formulation of CBT and establishes the desired mapping from the strings of members of the two sets induced by the given mappings. Perceiving the zigzagging strings in the CBT setting is the gestalt of

[22] There is a typo in the original: when R_1 is introduced it is written as 'R,'.

[23] Bernstein's 1906 proof, being similar to Peano's first proof is here ignored.

J. Kőnig's proof and the pairing of adjacent elements of the strings to form the desired mapping lies its metaphor. None of the proofs given to CBT before J. Kőnig[24] used the string gestalt or the pairing metaphor from which the desired mapping emerges.

In his 1914 book (pp 62, 215) J. Kőnig said of the equivalence between sets that it is always a relationship between the members of the sets. No doubt this view was the origin of his gestalt for his 1906 proof. This position would change with the Whittaker-Tarski-Knaster move (see Chap. 31) of the context of CBT to the power-sets of the given sets.

Strangely, in the mentioned book (p 219 footnote 1), J. Kőnig described his 1906 proof to be similar to the proof which Hessenberg published in the same year (1906 §21). The view was repeated uncritically by Medvedev (1966 p 239). Actually Hessenberg's proof is similar to the first proof of Peano (see Sect. 20.1) and it lacks the fruitful gestalt of the strings as equivalence classes. However, J. Kőnig was so convinced in his view that he dropped his own proof and presented in the book the proof of Hessenberg.

In Schrőder's proof of the two-set formulation the zigzag *appears in the Scheere*, but it was between nesting sets, not in strings or even between frames. In Bernstein's original proof of the single-set formulation, no strings or zigzagging were spotted; it was instead realized that the first frame could be neglected after abstraction. Borel who identified the frames, correlated them pairwise, without zigzag, and ignored their elements. Schoenflies' (1913 p 37) statement that J. Kőnig's proof is identical with that of Bernstein (referring to Borel) is thus groundless.[25] Schoenflies and Poincaré in his first proof, took the same approach as Borel, working for the two-set formulation and cascading the generation of the frames in tandem. All the proofs that spotted the chain as composed of denumerably many frames (Zermelo 1901 (under its cardinal numbers mantle), Jourdain 1904/1907, Peano's first proof, Bernstein 1906) worked in the single-set formulation; they focused on partitioning the set, pushing down the frames of the chain and generating the required mapping from the given one and the identity, each on its own partition. These proofs ignored the elements as well. Those proofs that saw the chain as generated by the impredicative definition (Dedekind, Zermelo 1906/1908, Peano's second proof) collapsed the chain into itself, ignoring the fate of the elements in the process. Though Dedekind did use the string gestalt in his simple chains, not-left-extendible strings, which he devises as a model for the natural numbers. Harward associated ordinals to the nesting sets and so was off the context of the elements of the sets. Cantor did correlate the elements of the sets but he assumed that they are well-ordered and thus he had one string instead of many.

[24] J. Kőnig mentions only the proofs of Bernstein, Schrőder and Zermelo.

[25] In his 1913 work (p 36f), Schoenflies brought J. Kőnig's proof in detail. He noted that the left-extendible strings belong to the residue, ignoring the possibility that some of them can be finite.

The two sequences of frames that appear, in the proofs based on Dedekind's theory of chains, in each of the sets given in the conditions of CBT, can be regarded as "lateral" partitions of the sets, with the residue giving the "bottom" partition. Then J. Kőnig's strings can be regarded as "vertical" partitions. Under the lateral gestalt corresponding frames are equivalent and so are the residues; combining these equivalences gives the equivalence of the two sets. Under the vertical gestalt, each string defines a mapping between its consecutive members and the desired equivalence is obtained by unifying all these "local" assignments. The not-left-extendible strings have one and only one member in common with each of the frames and in general their zigzag pattern follows the *Scheere* correspondence between the frames. The left-extendible infinite and finite strings have their members coming only from the residues – the intersection of all the nesting sets. Interestingly, J. Kőnig ignores the lateral partitioning of the previous proofs of CBT as a contrasting metaphor to his own vertical construction.

Obviously, the vertical and lateral partitioning exist in any reflexive set, namely, an infinite set, which, according to Dedekind definition, has a 1–1 mapping into itself. Note that when an infinite set is mapped *onto* itself, only the vertical partitioning gestalt remains and the strings are all finite or infinite left-extendible.[26]

[26] This paragraph is an example of proof-processing. We will have occasion to relate to the observations obtained within.

Chapter 22
From Kings to Graphs

We review in this chapter the research project launched by Dènes Kőnig, son of J. Kőnig, to implement his father's strings gestalt to produce a proof of BDT. His research led D. Kőnig to the theorem and lemma in graph theory that bear his name. It seems that D. Kőnig held to the methodological principle that if a gestalt can be applied to the proof of a certain theorem it can also be applied to other, similar proofs. This principle resembles Lakatos' heuristic principle (1976) that when a hidden lemma is spotted in the context of one theorem, related theorems should be searched to see if they too require the hidden lemma.[1] The strong connection between the proof of BDT and that of CBT[2] made the first a natural candidate for an application of the mentioned principle. (cf. D. Kőnig 1916 p 461 footnote ***). We have here an example of the theory that proofs drive mathematical development perhaps even more than do theorems and that proofs are often more important than theorems. After all, both CBT and BDT were already proved when J. Kőnig and D. Kőnig took an interest in them.

22.1 D. Kőnig's Proof that $\mathfrak{m} = \mathfrak{m} + \mathfrak{m}$

The background of D. Kőnig's first paper in his research project is this: In his 1901 paper (see Chap. 13), Zermelo studied certain closure properties of the set of all powers $\mathfrak{p}$ such that $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}$. For instance, that this set, which we denoted by $C_{\mathfrak{m}}$, is closed under addition, under denumerable addition and under diminution. However, Zermelo did not raise the question whether m is among these powers. We have noted that Zermelo probably realized that this result requires a new postulate. We have also noted that Russell, in his 1902 proof of CBT (see Chap. 15), having

[1] Lakatos mentions the principle when discussing the origin of uniform convergence.

[2] Which D. Kőnig calls the Equivalence Theorem and says that it was first proved by Bernstein, with no mention of Schröder.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_22, © Springer Basel 2013

postulated that every set can be partitioned to denumerable partitions, obtained that $\mathfrak{m} = \mathfrak{m} + \mathfrak{m}$. Cantor must have obtained $\mathfrak{m} = \mathfrak{m} + \mathfrak{m}$ already in 1882, when he proved the Union Theorem, that $\mathfrak{m} = \mathfrak{m}^2$ (see Sects. 1.3 and 2.3), but Cantor's result was for m a power in his scale of number-classes.

D. Kőnig's 1908 paper continued Zermelo's research,[3] along Russell's line of thought (see Schoenflies 1913 p 43f.). By leveraging on J. Kőnig's strings gestalt[4] D. Kőnig noticed that any 1–1 mapping from an infinite set to itself partitions the set into countable (finite or denumerable) partitions, which are strings (see the remark at the end of Sect. 21.5). By assuming a certain principle (axiom) D. Kőnig was able, by repeated partitioning, to derive a partitioning of any set into denumerable partitions so that the equality $\mathfrak{m} = \aleph_0\mathfrak{n}$ would emerge from which the desired result $\mathfrak{m} = \mathfrak{m} + \mathfrak{m}$ can easily be obtained.

D. Kőnig was aware that the result he was looking to establish is trivial when assuming the axiom of choice but he believed that the principle he had laid down was weaker. He had not noticed, however, that he did use the axiom of choice directly at other points in his proof.

D. Kőnig's principle, assumed in his 1908 paper, was the following[5]:

> If a set contains more than one element[6] then this set can be mapped on itself in such a way that no element is mapped to itself.

The mapping provided by D. Kőnig's principle institutes a J. Kőnig vertical partitioning in any set M. Denoting the partitioning by T_0, then the same principle can be applied to the set of partitions and provide a partitioning of that set, denoted by T_1. By taking the union of the partitions of T_1 one obtains a new partitioning of M, in which the size of the finite partitions must increase, because the partitions of T_1 must contain more than one whole partition of T_0 by virtue of the property of the mapping that it does not map an element to itself. Applying the principle to the set of the new partitions gives the partitioning T_2 from which again a new partitioning of M is obtained.[7] Having defined the partitioning T_ν for finite ν, a partitioning of M can be defined by collecting into a partition all the elements of M that appear in a

---

[3] The paper is not about two theorems of Bernstein as Franchella (1997 p 6) claims.

[4] D. Kőnig references Zermelo 1901 for the result $\mathfrak{m} = 2\mathfrak{m}$ is equivalent to $\mathfrak{m} = \aleph_0\mathfrak{m}$ but he did not mention Russell or J. Kőnig. Zermelo obtained this result from his Denumerable Addition Theorem $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}$ entails $\mathfrak{m} = \mathfrak{m} + \aleph_0\mathfrak{p}$. The proof of the latter theorem used gestalt and metaphor similar to the ones used in Peano's first proof of CBT (see Sect. 20.1). Thus the question whether $\mathfrak{m} = \mathfrak{m} + \mathfrak{m}$ was placed in D. Kőnig's research project.

[5] We do not know if the principle which D. Kőnig stated is indeed weaker than the axiom of choice.

[6] The interesting case is when the set is infinite, non-denumerable, as indeed D. Kőnig assumes in the proof.

[7] It seems that the proof works also if we take at each stage the set of all finite partitions only.

partition obtained following stage ν (and thus henceforth).[8] Clearly all the final partitions are denumerable because each is the union of denumerably many finite or denumerable sets (the Denumerable Union Theorem).[9]

However, the Denumerable Union Theorem requires the denumerable axiom of choice for its proof. In addition, the axiom of choice is used to enable a choice of a mapping satisfying the principle at every step ν of the construction (denumerably many choices from non-denumerable sets). Therefore, D. Kőnig did not meet his own expectation of proving the theorem without using AC, and, at the time, he was not aware of his failure. Still his proof is interesting for the iterated process of constructing compatible, expanding equivalence classes. "Repartition the partitions and unify" are the gestalt and metaphor we attach to the proof. It would be interesting to find the proof from which this method was proof-processed and proofs that proof-processed it.

22.2 D. Kőnig's 1914 Proof that $\nu\mathfrak{m} = \nu\mathfrak{n} \to \mathfrak{m} = \mathfrak{n}$

During the April 1914 congress on the philosophy of mathematics held in Paris, D. Kőnig presented a proof of BDT: $\nu\mathfrak{m} = \nu\mathfrak{n} \to \mathfrak{m} = \mathfrak{n}$ (see Chap. 14). The lecture notes were published only in 1923.[10] In the meanwhile D. Kőnig published his 1916 paper that contained the same results with even greater detail.

In the introduction to the 1923 paper D. Kőnig noted that the Well-Ordering Theorem trivializes theorems like BDT. Thus it appears that he believed that his proof makes no use of that theorem. Nevertheless, as in his 1908 paper, D. Kőnig did use the axiom. Still, he was correct to note (p 444; cf. 1916 p 461 footnote **) that his proof is simpler than the proof given by Bernstein.[11] Besides, Bernstein detailed the proof only for $\nu = 2$ while D. Kőnig proved the theorem for any finite ν. Yet Bernstein's proof had the advantage of making no appeal to the axiom of choice. Another point made by D. Kőnig, in the introduction to 1923, was that direct proofs, which do not employ the notion of well-ordering "permit better view to the bottom of things", remains valid despite his unnoticed use of AC, for proofs have

[8] D. Kőnig is not using the term 'equivalence relation' or 'equivalence class' but he does say that given a relation among members of M which fulfills the conditions of reflexivity, symmetry and transitivity, terms which D. Kőnig does not use, M is partitioned into disjoint partitions each composed of the members of M which are in that relation to each other. A style change over his father.

[9] Under this construction we can modify the mapping φ provided by the principle at T_0 so that for no finite subset S of M, $\varphi(S) = S$. If φ has this property to begin with the construction is not necessary and all T_0 partitions are denumerable. Cf. *Zahlen* p 42, Tarski 1924 p 92.

[10] World War I alone does not explain the delay in publication.

[11] The complexity of the proof, D. Kőnig notes, is evidenced by its exclusion from the very detailed Schoenflies 1913. One of the reasons that Bernstein's proof appears complex is that his proof is full of typos and lacks in systematic notation, as is clarified by our discussion of it (see Sect. 14.2).

their value even if they fail to prove what they were intended to prove (Lakatos 1976, 1978a p 61). Along this line of thought comes Wagon's view (1985 p 113f) that though CBT and BDT are "quite easy" assuming the axiom of choice, their direct proofs are valuable. We certainly agree with Wagon for we believe that it is from proof-processing that a mathematician's arsenal is acquired and proofs are a prerequisite to proof-processing.

With regard to his proof of BDT D. Kőnig stated (Kőnig 1923 p 444; cf. Kőnig 1916 p 461f footnote ***; Kőnig 1926 p 126f) that it implements an idea from his father's 1906 proof of CBT. Interestingly, D. Kőnig began to mention his indebtedness to his father only after the latter had passed away; in the 1908 paper, which likewise leans on J. Kőnig's 1906, no link is acknowledged explicitly. Nevertheless, in his 1926 paper (p 127 footnote 3) D. Kőnig noted that already in his 1908 paper he had the train of thoughts for his research program to implement his father's proof of CBT to a proof of BDT. He must have meant the footnote (1908 p 340) where it is observed that from the theorem $\mathfrak{m} = \mathfrak{m} + \mathfrak{m}$ the more general theorem follows that an infinite set can be put in pairs $\mathfrak{m} = 2\mathfrak{n}$ (see the following sections where the notion of pairing is developed). Note that the result of 1908 trivializes BDT so the 1908 paper is indeed within the research program of D. Kőnig for a proof of BDT.

Anyway, the idea implemented is the idea of partitioning a set into finite or denumerable partitions. D. Kőnig assumes as given two collections of v[12] equivalent sets M_i, N_i, $i = 1, \ldots, v$, all disjoint.[13] For an element a_1 (x_1) of M_1 (N_1), its corresponding element in M_i (N_i), $i > 1$, is denoted by a_i (x_i). The set of v-tuples ($a_1, a_2, \ldots, a_v$) is denoted by M (respectively, N is the set of v-tuples ($x_1, x_2, \ldots, x_v$)). Clearly M is equivalent to every M_i and N to every N_i. The condition of the theorem provides that the unions of the M_i and N_i, are equivalent and it requires that M and N be equivalent too. The given equivalence of the unions provides that every v-tuple of M corresponds to v, not necessarily distinct, v-tuples of N, and vice versa. Thus conditions similar to the conditions of CBT are established.[14] Now, if M + N can be partitioned in such a way that for each partition the number of M v-tuples is the same as the number of N v-tuples, then the required result would be obtained by unifying all the equivalences of the partitions.[15] Following J. Kőnig's lead, M + N is partitioned[16] to networks of corresponding v-tuples, the networks being the counterparts of J. Kőnig's strings. D. Kőnig calls 'chains', a sequence of points such that every two consecutive points in the sequence correspond. To distinguish

[12] In the paper the proof is for $v = 3$, perhaps because it was given in lecture. We give the general case.

[13] This point D. Kőnig neglects to make. He did make it in his 1926 (p 115) presentation of the same theorem.

[14] In D. Kőnig's courage to work with the esoteric 1-many mappings, lies the winning move of his strategy in this proof.

[15] But because the equivalence in each partition is not specific the derived equivalence cannot be obtained "effectively", as Sierpiński would later note (see Chap. 28).

[16] D. Kőnig refers to the partitions as 'subsets'.

from Dedekind's chains we will use 'strings' for D. Kőnig's 'chains'. A string in a network has the look and feel of J. Kőnig's string. Strings provide the connectivity between tuples of the same network and can measure the distance between tuples. The above sequence of definitions and observations erects the gestalt of the proof.

It is easy to see that the number of tuples in each network is finite or denumerable. Taking an arbitrary tuple S, the upper limit of the number of tuples with distance k from S is v^k and thus a partition has either a finite number of tuples or is the union of denumerably many finite sets of tuples and hence has denumerable number of tuples. Obviously, if the number of tuples is denumerable, the number of strings can be equal to the continuum! The method of partitioning into networks allowed D. Kőnig to reduce the question on the comparison of general powers to the comparison of finite or denumerable sets.

Now, let M* (N*) be the set of v-tuples from M (N) in a network with powers $\mathfrak{m}^*$, $\mathfrak{n}^*$. If both M* and N* are denumerable then there is nothing left to prove.[17] Obviously, both M* and N* must be finite if one is, so the alternative case is that M*, N* are finite; then the number of corresponding v-tuples can be counted either from the point of view of the members of M* or of N*; thus the following equation emerges: $v\mathfrak{m}^* = v\mathfrak{n}^*$, which trivially entails, $\mathfrak{m}^*$, $\mathfrak{n}^*$ being finite, $\mathfrak{m}^* = \mathfrak{n}^*$. Thus BDT is proved. Note that if $v\mathfrak{m} = \mu\mathfrak{n}$ and all the partitions are denumerable we obtain that $\mathfrak{m} = \mathfrak{n}$, but this does not hold in the general case. Note also that the proof proves that if a set M has 1–v correspondence to N and vice versa, the two sets are equivalent. But this result can be obtained directly, with AC, from CBT.

Finally, note that D. Kőnig's proof employs the axiom of choice twice. First, in the passage from the equivalence of the M*, N* of every partition (network) to the equivalence of M, N: when there are infinitely many partitions, AC must be used to choose a representative from the equivalences of M*, N* (cf. Zermelo 1908a p 188 (2)). Then, in case the partition is not finite, to prove that it is denumerable, because it is the union of denumerably many finite sets, the denumerable axiom of choice is necessary.

22.3 Into the Land of Graphs

In the second part of the 1923 paper D. Kőnig translated the discussion into what he calls "geometric language", which is the language of the theory of graphs. If D. Kőnig appears hesitant in his introduction of the notion of graph to his discussion, it was no doubt because at the time the theory of graphs was a rather esoteric body of mathematical knowledge. It is perhaps pertinent to our story to quote here a passage from Paul Erdos' reminiscences of his friend Tibor Galai (1982): "We both

[17] D. Kőnig did not separate the denumerable case and argued that it follows from the equation $v\mathfrak{m}^* = v\mathfrak{n}^*$ as the finite case discussed below.

went to the lectures of Dénes Kőnig on graph theory. In those "dark and uncivilized" times, graph theory and combinatorics were held in amused contempt by most mathematicians. A friend of my parents once said: "Dénes Kőnig is great in his art but his art is so small", and the great English topologist J. H. C. Whitehead[18] referred to graph theory as the "slums of topology". Now I think most mathematicians would remember J . Kőnig – an eminent set theorist – first of all as the father of Dénes Kőnig, which was very different in 1930."[19]

After the translation to the language of graphs, BDT suggested a Factoring Theorem for graphs. Factoring a graph was a notion that D. Kőnig took from Petersen 1891 (Biggs et al. 1986 p 190). A factor of a graph is a graph with the same vertices but only some of the edges. If it is a regular graph of degree n,[20] then it is an n-factor. Thus a 1-factor has only one edge incident to every point, so that it is composed of pairs of vertices linked by an edge.

Here is D. Kőnig's translation of the BDT discussion into the language of graphs as provided in the second section of 1923: The members of M + N he calls vertices and of two vertices he says that they are connected by an edge when they correspond by the correspondences given above. So two vertices may be connected by more than one edge and each vertex is incident on ν edges. The resulting graph is thus a regular graph of degree v. D. Kőnig notes that in this graph every circuit has an even number of vertices and edges.[21] The reason is that of two vertices linked, one belongs to M and the other to N. He then states the converse assertion that in every graph of even circuits the vertices can be divided into two sets such that only vertices that belong to different sets are linked. D. Kőnig proved this assertion in his 1916 paper (p 454): One begins with an arbitrary point assigned arbitrarily to one of the sets and each linked point to the other set and so on. A problem may arise only in a circuit with an odd number of points, a possibility that is ruled out by the assumption. In his 1916 paper D. Kőnig changed the term 'graph of even circuit' to 'coupled (*paare*) graph'. We will use for such graphs the term common today: 'bipartite graphs'. The partitioning of M + N described in the previous section can now be applied to any regular bipartite graph and so the theorem that the powers of M and N are equal becomes a general theorem for such graphs.

Compared to J. Kőnig's CBT proof, however, it appears that something is missing in this result. Though the strings of J. Kőnig's proof are not regular graphs, vertices beginning a not-left-extendible string are incident to only one edge, still the proof gives more than equality of power: it gives a factoring of the graph into 1-factors.[22] No similar factoring was offered by D. Kőnig's proof of BDT. It seems

[18] Not to be confused with his uncle A. N. Whitehead, Russell's collaborator.

[19] This oblique father-son appreciation is similar to the Pierce father-son story.

[20] A graph is regular of degree n when all its vertices are incident on n of edges (Petersen 1891; Biggs et al. 1986 p 190).

[21] From the context a circuit is a closed string, such that its first and last points coincide.

[22] Note that a 1-factor of a bipartite graph represents a 1–1 mapping between the two sets of vertices of the graph.

probable that D. Kőnig noticed the graph in his father's proof and its factoring[23] before his involvement with BDT began. Thus he stipulated in the second section of 1923 the following theorem: A. Every regular bipartite graph has a 1-factor.[24] After omitting from the given graph the edges of its 1-factor another regular bipartite graph is obtained, with degree decreased by 1. Hence D. Kőnig generalized the following: B. Every regular bipartite graph is the product[25] of 1-factors; the number of factors is equal to the degree of the graph. Obviously the two theorems are equivalent. The proof of either A or B, D. Kőnig promised to give at a later occasion and in a footnote dated 1923 he refers to his 1916 paper where the proofs are given for finite graphs.

Still, for the case $v = 2$, D. Kőnig did indicate (in the 1923 paper) how factoring can be attained: by omitting every second edge, which can be done in two ways (depending on the first edge omitted). These factorings are, in fact, J. Kőnig's possible definitions of the mapping between X and Y. However, while in J. Kőnig's proof the choice of which edge is omitted (or maintained) could be made uniformly for all strings of a type, and there are there only a finite number of such types, in the general case for a graph of degree 2 no such choice can be made and hence the axiom of choice is invoked. Interestingly, while D. Kőnig missed noticing his application of the axiom of choice in the first part of the 1923 paper, at this point of his paper (p 448)[26] he became aware of it.[27]

D. Kőnig closed the 1923 paper with an application of the Factoring Theorem for graphs to solve a particular case of the four colors problem. This is interesting because it demonstrates how proof-processing associates context to context, surely by association of gestalt and metaphors arising in the proof-processing of each context. In the short introduction to the 1916 paper,[28] D. Kőnig makes a similar observation when he explains that the paper deals with problems in topology, determinants theory and set theory, tied by the notion of graph: "the method of graphs… shows the equivalence of seemingly remote investigations". Our story echoes Lakatos' (1976) methodological observation of "translation into a new dominant theory".

[23] Porting J. Kőnig's proof to the language of graphs appears at first sight to be poorly justified, for the new dominant theory (graph theory) seems to add nothing to the understanding of the proof. This was also D. Kőnig's opinion (1916 p 462 footnote), but, as it turned out, this porting had consequences, for CBT could be generalized to a theorem on paths in graphs. See Diestel-Thomassen 2006 and the bibliography there.

[24] This became known as D. Kőnig's theorem.

[25] The notion of product of graphs (of the same set of vertices) comes from Petersen; it simply means that the graph is obtained by superimposing the multipliers, namely, all the edges are added while the points remain the points of the factors.

[26] There is a typo in the original: instead of "M* et N*" it should be 'M et N'.

[27] When a string is denumerable the situation reminds us of Russell's example with a denumerable pile of socks.

[28] In an endnote D. Kőnig remarks that the results of the paper were presented to the Hungarian academy of science in November 1915.

22.4 Factoring Finite Graphs

In the 1916 paper D. Kőnig proved his factoring Theorems A and B, for finite graphs, by the following Theorem C: If every point of a [finite] bipartite graph is incident to at most v edges then each edge of the graph can be assigned one of v indices [say, 1, … ,v] in such a way that all the edges incident to a point have different indices.[29] Thus in a regular graph all indices occur at every point and by taking the edges of the same index a 1-factor of the graph is obtained and so A, B follow from C. Conversely, Theorem B entails Theorem C. The proof (p 464) completes the graph in Theorem C to a regular graph by duplicating its vertices, connecting vertices and their duplicates with as many edges as necessary to make the total graph regular. The idea on how to duplicate a graph may have originated in Schröder 1898 p 325 where it is suggested to place a mirror over a set to get new elements. Note that Theorem C explains why in the graph generated by J. Kőnig's 1906 proof of CBT a 1-factor emerges: because when $v = 2$ all points with two edges have edges numbered 1 and 2 and thus it does not matter which number the one point with a single edge in every not-left-extendible string is assigned: still all edges with the same number form a 1-factor of the string. It is possible that D. Kőnig obtained the clue for Theorem C by further proof-processing his father's 1906 construction in light of a procedure used in the 1880s by Tait (Biggs et al. 1986 p 103; Tait is explicitly referenced in D. Kőnig 1916 p 457; Kőnig 1926 p 128).

The proof of C is carried by complete induction on the number of edges of the graph. It is trivial when this number is $\leq v$ so assume it holds when the number is $<$N and prove it for a graph G with N edges. Remove an edge e from the graph, incident to the points α, β. The graph remains a bipartite graph with at most v edges incident at a point (but not regular even if G was regular),[30] so under the induction hypothesis we can assume its edges numbered as stipulated by Theorem C. If there is an index which does not appear as an index of an edge incident at α or β, then giving this index to e provides the indexing of G. Otherwise, the edges incident at β do not have all the indices $<v$ because as e was removed there are less than v edges incident at β. Let the missing index be v_1 and by the assumption there must be an edge incident at α indexed by v_1. Let v_2 be an index that does not appear as an index of an edge incident at α. Let α_1 be the second point incident on the edge indexed by v_1 emanating from α. If there is no edge incident at α_1 indexed by v_2 then we can change the index of the edge indexed v_1 to v_2 and free the index v_1 for e. Otherwise, we continue to form a string α, α_1, α_2, … such that between consecutive points the

[29] D. Kőnig here treats the general case not the case $v = 3$ of 1923, and he uses k for the 1923 v, which we maintain.

[30] It is probably because the induction step destroys the regularity of the graphs that the theorem speaks of non-regular graphs.

edges are indexed either by v_1 or v_2.[31] A point in the string other than α cannot be retraced because then it will have three edges indexed by v_1 or v_2 contrary to the indexing provided by the induction hypothesis. Neither can α be retraced because it does not have an edge indexed v_2 and cannot have two edges indexed v_1. Also β cannot appear in the string because it cannot be reached with an edge indexed v_1 because of how v_1 was chosen, and if it is reached with an edge indexed v_2 the string would have a pair number of edges and we would obtain a closed circuit in G, before the omission of e, with an odd number of edges, contrary to the condition that G is a bipartite graph. So the string will eventually terminate (the graph is finite). Interchanging the indexes of the edges between consecutive points of the string we free v_1 to be used to index e. This proves Theorem C and hence also Theorems A and B. The courage to walk on a string to establish the required permutation of the indices which is the main metaphor in the proof of Theorem C, may have come to D. Kőnig from his familiarity in threading along his father's strings, but mainly, it seems, from a similar procedure used by Kempe in 1879 (Biggs et al. 1986 p 94ff) as D. Kőnig noted explicitly (1916 p 457; Kőnig 1926 p 128).[32]

After mentioning several applications of his factoring results to the four colors problem (mentioned already in the 1923 paper) and determinants theory, D. Kőnig came back, in the third section of the 1916 paper, to set theory. This time he partly replaced the geometric language he used in 1923 for the graphs with set theoretic language: a graph is composed of a set of points (not vertices) and pairs of points are the edges.[33] The advantage of the translation lies for D. Kőnig in the possibility to discuss infinite graphs with infinitely many points and edges,[34] without relying on geometrical intuition.

Also to BDT D. Kőnig gives a new setting: Let M be a set of power m and N be a set of power n. D. Kőnig says that there is a bi-v mapping[35] between M and N when to every member of M correspond v, not necessarily different, members of N and vice versa, in such way that if to a corresponds b with multiplicity s then to b corresponds a with the same multiplicity. BDT now becomes: if there is a bi-v mapping between M and N then there is a 1–1 mapping between them. For this

[31] The alternating nature of the edges of the string is reminiscent of the alternating nature of J. Kőnig's strings.

[32] The need to change a pre-given arrangement appears also in Bernstein's original proof of BDT (see Sect. 14.2). Note that the results of Tait and Kempe mentioned above were obtained in the context of the four-color problem.

[33] Multiplicities can conveniently be denoted by adding a number to each pair. D. Kőnig still uses a geometrically inclined language (points, edges) but the meaning of these terms is no longer geometric.

[34] The number of edges incident at every point is required to be finite; hence connected graphs must be finite or denumerable. D. Kőnig notes that it is meaningless to say that two points are connected through infinitely many edges.

[35] D. Kőnig's term is "reversible (1, v) relation". We follow Sierpiński (see Sect. 28.2) as did D. Kőnig later (see Sect. 30.3).

Theorem D. Kőnig repeats the proof from his 1923 paper (first part, see Sect. 22.2).[36]

Then D. Kőnig generalizes the theorem by requiring that the 1–1 mapping correspond only such members of M and N that correspond under the bi-v mapping.[37] His Factoring Theorem for finite regular bipartite graphs proves this theorem for the finite partitions. To prove the generalized BDT it remained to establish the Factoring Theorem for denumerable regular bipartite graphs. This, however, D. Kőnig admitted (p 463) that he was unable to prove. Nevertheless, he did prove (p 464f) that it is enough to prove the Factoring Theorem for denumerable partitions when the degree of the graph is a prime number. Hence, by his result for graphs of degree 2 (see the previous section), all regular bipartite graphs of degree 2^{μ} can be factored.

In footnotes (p 461 footnote *, 464 footnote *) D. Kőnig asserts that when passing from the partitions to the entire graph the axiom of choice is invoked. So apparently he realized his mistake in the 1923 paper.[38] But D. Kőnig stressed (p 461 footnote **) that while he had no intention to avoid AC his proof uses only the simplest set theory notions without making any use of the notion of order, thereby he repeated the idea he mentioned in the introduction to 1923 and clearly distinguished between use of well-ordering and use of AC.

22.5 Factoring Denumerable Graphs

D. Kőnig's research project came to a successful conclusion when in late 1924, in collaboration with Stephan Valkó,[39] he found a proof of the Factoring Theorem for denumerable regular bipartite graphs. The paper with this result was published in 1926. The core of the proof is now known as D. Kőnig's infinity lemma (cf. Franchella 1997 p 23, 25) because in D. Kőnig's 1926 paper it was presented as a lemma to the Factoring Theorem.[40] In his 1926 (dated May 1925) D. Kőnig

[36] The older proof was repeated because it provided a proof also for the case where the networks are denumerable. The Factoring Theorem was proved in 1916 only for finite graphs so it could not provide a proof for BDT in general.

[37] In the 1926 paper of Kőnig-Valkó (p 135 footnote 2) it is stated that the theorem was first announced in D. Kőnig's 1923 paper, but there, this form of presenting BDT is missing.

[38] The view is repeated in 1926 (p 133). Regrettably, he there says that Bernstein's BDT proof was without the axiom of choice only for $v = 2$. But Bernstein's proof can be extended to any v without the axiom of choice (see Sect. 14.3).

[39] We could not find any information on Stephan Valkó.

[40] D. Kőnig's lemma is an important result in graph theory with applications to computer science. It is usually known under the following formulation: every tree with infinitely many vertices has an infinite path (or branch).

compiled all his results emerging from his attempts to prove BDT[41] with an historical background and links to other relevant results of the period. Such results, that share with D. Kőnig work the context of BDT and the string gestalt of J. Kőnig, were obtained by members of the Polish school of logic: Sierpiński (1922, 1947a, 1947d), Banach (1924) and Kuratowski (1924), Lindenbaum-Tarski (1926), Tarski (1930 pp 246, 248, 1949b). Cf. Sect. 30.3.

We summarize the proof from the 1926 paper of Kőnig-Valko.[42] Let G be a denumerable regular bipartite graph of degree v. Let A_i, B_i be its two sets of points. Let G_n be the sub-graph of G composed of all the first n A_i, B_i points, with all the edges incident at them and all the points incident to these edges which are not among the first n points ('border points'). G_n is bipartite but generally not regular because not all edges of the border points are included. It can, however, be completed to a finite regular bipartite graph of degree v, G'_n, by duplicating it (the mirror metaphor again) and connecting the border points and their duplicates to complete the number of edges incident at them to v. G'_n is bipartite. By Theorem B, G'_n can be factored and therefore also G_n. Now G_n is a sub-graph of G_{n+1} and every factor of G_{n+1} extends a factor of G_n. So among the factors of G_n there is one that is extended by infinitely many of the factors of all G_m, $m > n$.[43] Hence one can define by induction for every n a factor S_n that is a part of infinitely many factors of the G_m, $m > n$, and extends S_{n-1}. S_0 can be taken as the empty set. The union of the S_n is a factor of G. Clearly, the axiom of choice is required to choose all the S_n. Note that while the Factoring Theorems prove that the required equivalence in BDT is related to the mappings given in the conditions of that theorem, the specific mapping for every element cannot be designated. In this the result provides less information than J. Kőnig's proof.

[41] The presentation in this paper suffers from unattractive notation, which seems to follow the structuralist vogue of the period.

[42] Valkó provided the crucial metaphor of the proof (D. Kőnig 1926 p 120 footnote 1).

[43] Some similarity to the Bolzano-Weierstrass theorem appears here. Cf. Franchella 1997 p 24 footnote 20.

Chapter 23
Jourdain's Improvements Round

Harward's criticism (see Chap. 18) effected Jourdain to produce two new papers, his 1907b and 1908a papers, where he provided correct proofs of results which he failed to properly establish in his papers of 1904 (see Chap. 17).

The 1907b paper revolves around the problem of the comparability of sets, organized in the scheme of complete disjunction, and how the axiom of choice provides a solution to case (IV). Jourdain (p 355, end of Sect. 2) correctly identified the statements of the scheme as being "more general statements than statements involving cardinal numbers", namely the comparability of sets being more general than the Comparability Theorem for cardinal numbers.[1] Since one of the cases of the scheme is CBT, Jourdain addressed the theorem again, this time providing for it a rigorous proof. In it, Jourdain constructed the mapping between the given sets from the given mappings. However, by 1907 this feature was no longer new: it was implicit in Zermelo's proof published by Poincaré (see Sect. 19.5) in 1906b, and it became explicit in Peano's 1906 proof (see Chap. 20) as well as in J. Kőnig's 1906 proof (see Sect. 21.2).

An unfortunate long-term influence of the paper was in the naming of CBT. In the paper (p 355 footnote §) Jourdain explained his preference to call CBT the Schröder-Bernstein Theorem,[2] by saying that Cantor's proof was not as general as the proofs of Schröder and Bernstein. Needless to repeat that Schröder's proof was irreparably erroneous (see Chap. 10), while Cantor, not only conceived the

[1] However, Jourdain added an obscure remark that the last statement holds "unless we define the cardinal number of M, as I do now [here a footnote is linked], as a certain class similar to M." In the footnote Jourdain promises to publish soon his new definition. He goes on to explain why the Frege-Russell definition of cardinal number is inappropriate but he gives no clue why he believes that a representation theory for cardinal numbers will by-pass his statement regarding the greater generality of the scheme. Jourdain never published the definitions (see Grattan-Guinness 1977 p 85ff).

[2] In 1904, Jourdain still used the name suggested by Whitehead-Russell (Whitehead 1902): the theorem of Schröder and Bernstein.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_23, © Springer Basel 2013

theorem, but also proved it for sets with power in his scale of number-classes (see Chap. 1).

The 1908a paper is focused on a proof of the Union Theorem. Though the theorem was proved in 1906 by Hessenberg,[3] Jourdain's proof was noteworthy and in fact, it became since a classic (cf. Levy 1979 p 96f). Possibly because it uses the given well-ordering of the ordinals while Hessenberg used the more involved metaphor of natural sums. It should be mentioned here that both these proofs are different from the proof given by Harward.

At the end of the 1908a paper (p 512 footnote **; cf. p 511 footnote **) Jourdain also mentions his gratitude to Zermelo for "repeatedly criticizing weak points in my proofs and suggesting improvements". We are not aware of any correspondence between Jourdain and Zermelo at the time but one can say that the style of the 1908a paper is much improved compared to Jourdain's style in his prior papers. Jourdain bowing towards Zermelo is interesting because it came after Jourdain, in his 1905a paper, claimed priority over Zermelo in finding the Well-Ordering Theorem. Zermelo addressed this claim in his 1908a (p 191ff).

23.1 The 1907 CBT Proof

The theorem's presentation begins thus (p 358):

> given two classes[4] M and N, a one-one representation ϕ of M on a proper part N_1 of N,[5] and a one-one representation ψ of N on a proper part M_1 of M, so that we have $\phi(M) = N_1$, $\psi(N) = M_1$, we can define, in terms of ϕ and ψ, a process of separating of what correspond to the P_ν's above which permits the application of (1). This proof that (1) is applicable and the application of (1), then, constitutes the essence of the theorem of Schröder and Bernstein that a representation θ of M on N is uniquely determined by ϕ and ψ.

What Jourdain references as the separation process (1) he described in the preceding page (p 357). It is the same metaphor that he used in his 1904a rendering of Zermelo's 1901 CBT proof, which was criticized by Harward (see Sect. 17.5). In 1907b, Jourdain presented the process as follows:

> Let, now, P be a class, and let P be divided into two subclasses P_1 and M_1, so that $P_1 + M_1 = P$; let M_1 be divided into P_2 and M_2; and so on. Thus $P = \Sigma_{\nu=1...n}P_\nu + M_n$, and this process can always be continued if M_n has any members. If our P and our process of separating out subclasses is so defined that M_n always has members, whatever the finite n may be, we have (1) $P = \Sigma_{\nu<\omega}P_\nu + M_\omega$, where ω is the first transfinite ordinal number and M_ω is a class which may or may not have members. If it has, these members are common to all the classes M_ν, where ν is a finite ordinal number; if it has not, M_ω is the first class in our series of classes M_ν that has not.

[3] At the end of the paper Jourdain noted Hessenberg's 1906 proof.

[4] In the middle of the 1907b paper (p 357 footnote *) Jourdain switched from 'aggregates' to 'classes', claiming the second term has the advantage that it is shorter.

[5] Jourdain uses the term 'representation' for 'mapping' perhaps after the German *Abbildung*. We will use Jourdain's term only when quoting from him.

We can extend this formula to the case of any transfinite ordinal ζ. If we have a process of separating out classes P_γ from P, and, if $\gamma < \zeta$, M_γ has members, then $P = \Sigma_{\gamma<\zeta}P_\nu + M_\zeta$, where M_ζ may possibly not have members.

Jourdain demonstrates in this passage a clear misunderstanding of Harward's criticism. The problem is not with consolidating separated frames but of generating equivalent frames. The application of lemma (1) in the proof of CBT (see Schoenflies' proof in Chap. 12) is hardly noticeable and it is clearly not the step that wins the proof.

Anyway, to prove CBT Jourdain moves to the equivalent single-set formulation. How the shift is contemplated, he details at the end of the proof (see below). The single-set formulation he formulated as follows:

If P is a class, P_1 is a proper part of P, P_2 a proper part of P_1, and ϕ a one-one representation of P on P_2, we can find, in terms of ϕ, a one-one representation of P on P_1.

Though the proof of the theorem can be given in a few lines we prefer to bring Jourdain' full proof because it evidences the wrestling sometimes necessary in order to obtain concise expression of mathematical ideas, even elementary ideas, and how such wrestling may end in a failure from the point of view of style.

We will here examine the construction of this θ with some minuteness, and this will bring out more clearly, what is the real nerve of the proof, that the correspondence (θ) of the whole (P) and part (P_1) is brought about by the division, brought about by ϕ, of P into two sets of $\aleph_0$ parts each; the one, consisting of $\aleph_0$ different parts of P_1, in which $\theta(x) = x$; the other, consisting of $\aleph_0$ different parts of P (making up together with the former parts, the whole of P)[6] and that part of P which is not P_1 (which we will denote P-P_1), in which $\theta(x) = \phi(x)$. The fact that the whole of P is then represented on the whole of P_1 is, then, a consequence of the equality $\aleph_0 + 1 = \aleph_0$.[7]

Further, this construction will emphasize the important fact that θ is determined *uniquely* by ϕ and P_1, an analogous process allows us to determine uniquely $\aleph_0$ members in any infinite class M, in terms of any element m_0 of M and a representation ϕ, which must exist, of M on the part of M which lacks only the member m_0 of M.[8]

We will put $P_2 = \phi(P)$, [apparently Jourdain forgot that he already defined P_2] $P_3 = \phi(P_1)$, $P_4 = \phi(P_2)$, and in general $P_{2n} = \phi^n(P)$, $P_{2n+1} = \phi^n(P_1)$, (n = 1, 2, ...),[9] where ϕ^n denotes the 'relative product'[10] of n representations ϕ.

---

[6] There seems to be a typo in the original and P_1 is written instead of P.

[7] Here Jourdain references his Zermelian 1904 proof of CBT where this equation is used (see Sect. 17.5). In the proof given here, however, this equation is not used. See footnote 17 of Chap. 19 concerning a similar remark of Fraenkel.

[8] Jourdain is correct only if 'infinite set' is Dedekind infinite. Jourdain, however, had not noticed this and he suggested (p 360f) his argument here as an improvement of Cantor's proof (1895 *Beiträge* §6) that every infinite set contains a denumerable subset, which requires the axiom of choice if such ϕ is not given.

[9] Here Jourdain obtains the gestalt of two sequences of nesting sets. He makes no explicit passage to the gestalt of frames but uses it in the definition of θ. Likewise the residue is introduced only obliquely in (3) below.

[10] Here Jourdain references Russell 1903 p 25f.

Then θ, where $\theta(P) = P_1$, is constructed as follows: If x is any member of P_1 we find $\theta(x)$ in the way described in the four cases:

1. If x is a member of P_1 and also a (P_{2n-1}-P_{2n}), we have $\theta(x) = x$;
2. If x is a member of P_1 and also a (P_{2n}-P_{2n+1}), we have $\theta(x) = \phi(x)$;
3. If x is a member of P_1 but is neither a (P_{2n}-P_{2n+1}) nor a (P_{2n-1}-P_{2n}), if such an x exists, we have $\theta(x) = x$[11];
4. If x is not P_1 we have $\theta(x) = \phi(x)$.

In this way every x of P has its correlate $\theta(x)$ in P_1, and conversely, if y is any member of P_1 it has a correlate $\breve{\theta}(y)$, where $z=\breve{\theta}(y)$ means $y = \theta(z)$, in P.[12]

At this point Jourdain explains the link of CBT in its two-set formulation to the single-set formulation:

We pass from this to the Schröder-Bernstein theorem by putting $P = M$, $P_1 = M_1$, N for a class such that $\psi_2(N) = P_1 = M_1$, and N_1 for a proper part of N such that $\psi_1(M) = N_1$[13]; so that our ϕ becomes the "relative product" $(\psi_2; \psi_1)$ of ψ_2 and ψ_1.

We leave it to the reader to determine how much clearer the proof can get.

23.2 The 1908 Proof of the Union Theorem

In his 1908a paper Jourdain came back to the Union Theorem to correct his 1904b proof criticized by Harward (1905). At the opening of the paper Jourdain declares that he wants to prove that $\aleph_\gamma \cdot \aleph_\gamma = \aleph_\gamma$ by generalizing the method of proof given by Cantor. Jourdain references three of Cantor's papers for the mentioned method: the 1874 *Eigenschaft*, the 1878 *Beitrag*, and the 1895 *Beiträge* (§6).[14]

Of these three sources, only the third appears relevant. There Cantor proved that $\aleph_0 \cdot \aleph_0 = \aleph_0$ and it is that proof that Jourdain intends to generalize in the 1908a paper. Mention of the second source is perhaps because of the proof given there that $\aleph \cdot \aleph = \aleph$, but this result is not relevant to $\aleph_\gamma \cdot \aleph_\gamma = \aleph_\gamma$. The first source contains the result mentioned from the third source but the proofs are slightly different. Jourdain, however, says that the method of proof in the third source is fundamentally the same as that in the first source. In view of Jourdain's statement that the proofs of Schröder, Borel, Schoenflies and Zermelo 1901, are all similar (Jourdain 1907b p 353 footnote ‡), his opinion is not reliable.

---

[11] Equally rigorous would be to define $\theta(x) = \phi(x)$ in this case, so we see that θ is not uniquely determined as Jourdain claimed. And there are other changes to the definition that can be made, rendering different possible θ.

[12] The definition of θ is the metaphor of the proof. It is the partition and pushdown metaphor of Peano, badly articulated. And nothing motivates the sudden introduction of the inverse of θ.

[13] There is a typo here in the original and N is written instead of N_1. In introducing the mappings ψ_1 and ψ_2 Jourdain deviates from his own notation in the presentation of the two-set formulation given above.

[14] Jourdain also mentioned Bernstein's testimony that Cantor had a general proof of the theorem, see Sect. 1.3.

As a set of power $\aleph_\gamma \cdot \aleph_\gamma$ Jourdain takes the set S = $\{(\alpha, \beta) \mid \alpha, \beta < \omega_\gamma\}$ (p 509). The pairs in this matrix are "couples with sense", namely, ordered-pairs in current terminology. However, whereas Cantor assigned to every such pair, in the case $\gamma = 0$, a finite number $\lambda = \alpha + ½(\alpha + \beta - 1)(\alpha + \beta - 2)$, Jourdain proved that S is well-ordered and that the ordinal number of every segment of S is $<\omega_\gamma$. Hence, it follows that the power of S is $\leq \aleph_\gamma$. Since obviously the power of S is $\geq \aleph_\gamma$, we have by CBT the desired result. Jourdain could assume CBT for any set but Cantor could not. Thus Jourdain is not generalizing the Union Theorem under Cantor' own terms (see Chap. 1).

The well-ordering of S is defined (p 507) as follows: $(\alpha, \beta) < (\alpha', \beta')$ when $\alpha + \beta < \alpha' + \beta'$ or $\alpha + \beta = \alpha' + \beta'$ and $\alpha < \alpha'$. Note that if $\alpha + \beta = \alpha' + \beta'$ and $\alpha = \alpha'$ then necessarily $\beta = \beta'$ for if $\beta < \beta'$ then $\alpha + \beta < \alpha + \beta + 1 \leq \alpha + \beta'$, a contradiction. This order is indeed a well-ordering of S because (p 508) if P is any subset of S, the set of all ordinals $\alpha + \beta$, where $(\alpha, \beta) \in$ P, being a set of ordinals, has a minimal member, say κ, and the set of all α such that there is a β such that $(\alpha, \beta) \in$ P and $\alpha + \beta = \kappa$ also has a minimal member, say α_0. Let then β_0 be the ordinal such that $(\alpha_0, \beta_0) \in$ P; then (α_0, β_0) is the minimal member of P.[15] Let σ be the ordinal number of S.

The proof is explicitly conducted by transfinite induction (p 509 footnote †). The theorem was proven by Cantor for $\aleph_0$ so Jourdain says that "if it holds for all alephs less than $\aleph_\gamma$, it holds for $\aleph_\gamma$; then it holds for all alephs." As induction hypothesis Jourdain takes the following (1) $\zeta < \omega_\gamma$ implies $\bar{\zeta} \cdot \bar{\zeta} = \bar{\zeta}$, where $\bar{\zeta}$ is the cardinal number of ζ.

Let S′ be a segment of S and let σ' be the ordinal of that segment. Let (α', β') be the smallest pair in S not in S′. Jourdain first proves the lemma: $\alpha' + \beta' < \omega_\gamma$ when $\alpha', \beta' < \omega_\gamma$. By the Sum Lemma (see Sect. 17.6) $\overline{\alpha' + \beta'} = \bar{\alpha}' + \bar{\beta}' \leq 2 \cdot \bar{\xi} \leq \bar{\xi} \cdot \bar{\xi}$ where ξ is the larger of α' β'.[16] By (1) we have that $\bar{\xi} \cdot \bar{\xi} = \bar{\xi} < \aleph_\gamma$ hence the lemma. Put $\lambda = \alpha' + \beta'$. Then $\bar{\sigma}' \leq \bar{\lambda} \cdot \bar{\lambda}$, since all the couples (α, β) in S′ obey $\alpha < \lambda$, $\beta < \lambda$. and by the induction hypothesis $\bar{\lambda} \cdot \bar{\lambda} < \aleph_\gamma$. So we have $\bar{\sigma}' < \aleph_\gamma$. Hence $\bar{\sigma} \leq \aleph_\gamma$, for else S would have a segment of power $\aleph_\gamma$.

After the proof, Jourdain notes that the theorem is essential to prove that all the alephs are different (the Different Alephs Theorem). He does not explain how this proof should go about but references his 1904a p 74. There it is pointed out that the Union Theorem is necessary to prove the existence of the series of alephs. The reference that Jourdain had in mind was perhaps his attempt to prove (in his 1904b) the Next-Aleph Theorem (see Sect. 17.6).

---

[15] Jourdain considered this result as a proof that the multiplicative axiom holds for S, which means, presumably, that it holds for the power-set of S, and stated that Zermelo (1904) had proved that if the multiplicative axiom can be proved for a set then this set can be well-ordered. This remark is amazing: Jourdain just proved that S is well-ordered, he concluded that in S the multiplicative axiom holds AND he discovered thereby that S is well-ordered! Even the, no doubt, clear guidance of Zermelo, could not stop Jourdain's compulsive tendency for over-statement.

[16] Adding an overline above the sign for an ordinal signifies its cardinal.

Chapter 24
Zermelo's 1908 Proof of CBT

Zermelo's aim in his 1908b paper was "starting from set theory, as it is historically given, to seek out for the principles required for establishing the foundations of this mathematical discipline" while restricting "these principles sufficiently to exclude all contradictions".[1]

Moore (1978 p 308) said that "the paradoxes were neither the sole nor the central factor motivating Zermelo's axiomatization", rather "Zermelo was primarily motivated ... by a desire to preserve his axiom of choice". But it comes out from the quoted passage in the text that Zermelo' motivation was to carve out a minimal safe set of existential (closure) statements necessary to preserve set theory. The unveiling of the axiom of choice may have suggested this research plan, which fell within Hilbert's axiomatization program, and was spurred by the antinomies.

The part of set theory that Zermelo wanted to save is presented in the second section of his paper titled "Theory of equivalence". After establishing the definition of equivalence this part contains a proof of CBT in its two formulations and proofs of Cantor's Theorem, Kőnig's inequality and the minimality of $\aleph_0$. We will focus on CBT and show how already its presentation requires axioms I–V and the grounding concepts of Zermelo's axiomatic set theory, suggesting that CBT was the context of its formation.

Zermelo first communicated his new proof of CBT to Hilbert in June 1905 (Ebbinghaus 2007 p 89), when he was developing the systematic exposition of set theory from axioms (Peckhaus 1990 p 30). Thus when Poincaré challenged Couturat (1906a p 29, published in January) to produce a proof of CBT that does not use complete induction, Zermelo was quick to react; already in January 1906 he sent Poincaré his proof that employed Dedekind's definition of a chain of a set as

[1] Incidentally, we raise a question that we will not attempt to answer here: Is axiomatization a method in the development of mathematics? Clearly, axiomatization effects the presentation of the axiomatized theory, but does it effect also the gestalt and metaphors used to obtain the theory's proofs? Cf. Yuxin 1990 p 388.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_24, © Springer Basel 2013

the intersection of all chains containing the set (see Sect. 9.1). Poincaré presented the proof (1906b published in May, see Sect. 19.5) but objected its use of impredicative definition. In 1908a Zermelo retorted (1908a §2.b), rejecting Poincaré's criticism and published the proof in his 1908b paper.

Zermelo, unlike Peano, remained loyal to the informal language used by Dedekind and Cantor in their discussions of set theory, even in his paper on axiomatic set theory. Though later, offsprings from Peano's sign language did become the standard language of presentation for mathematical formulas in mathematical logic and set theory (see Levy 1979), we will adhere in this chapter to Zermelo's language.

An important element of style that Zermelo apparently picked from Dedekind *Zahlen* concerns the numbering of passages. The advantage of such numbering is that it facilitates internal reference. In addition, with every new number a new topic begins and the connection to earlier passages is required to be logical rather than through flowing rhetoric.

We quote from the translation of Zermelo 1908b in van Heijenoort 1967 p 200–215. We will reference Zermelo's text through its numbering, using the prefix # instead of Zermelo's 'Nr.' and the translated 'No.'.

For more on Zermelo and his work see Moore 1982 and Ebbinghaus 2007.

24.1 CBT and Its Proof

We will first present Zermelo's proofs of CBT in its two formulations. In the following sections we will explicate how the notions used in the proofs are established in Zermelo's theory. We will end this chapter with a comparison of Zermelo's proof in 1908b and the proof presented by Poincaré in his 1906b.

Zermelo presented the Equivalence Theorem[2] in its two-set formulation (#27) as follows:

> If each of two sets M and N is equivalent to a subset of the other, M and N are equivalent.

In his notation here Zermelo followed Cantor who used the letters M and N in the context of comparison of sets already in his 1878 *Beitrag* and in all his later papers where CBT in its two-set formulation is mentioned.

The two-set formulation is proved by the single-set formulation:

> Let $M \sim M' \subseteq N$ and $N \sim N' \subseteq M$. Then on account of #21[3] there corresponds to the subset M' of N an equivalent set M'' such that $M'' \subseteq N' \subseteq M$, and we have $M \sim M' \sim M''$; therefore, according to the Theorem #25 [the single-set formulation], also $M \sim N' \sim N$, Q.E.D.

[2] Zermelo used this name here as well as in all his other writings about CBT.

[3] This is Dedekind's lemma.

Cantor named the corresponding sets after the containing sets and Zermelo names them after the origin sets. Zermelo's use of ′ to denote the image of a set seems to have come from Dedekind *Zahlen*. Zermelo's notation was new in the history of CBT; only Harward used a similar notation. Zermelo's notation was later adopted in Principia Mathematica (*117.2).

Zermelo presents the single-set formulation and its proof in #25 as follows:

> If a set M is equivalent to one of its parts, M′, it is also equivalent to any other part M_1, that includes M′ as component.

Cantor, in his 1895 *Beiträge* used the notation M, M_1, M_2. Zermelo followed Cantor with M_1 but not with M_2, which Zermelo denoted by M′. Zermelo followed none of his predecessors when he introduced M_1 as a "floating" set between M and M′, a point that is generally omitted in textbooks. He did it already in his 1901 paper, where in particular he pointed out that $M \sim M' + P$ where $P = M - M_1$.

The proof of the single-set formulation runs as follows (#25):

> Let $M \sim M' \subseteq M_1 \subseteq M$ and $Q = M_1\text{-}M'$.[4] Because of the equivalence $M \sim M'$ that has been assumed, there exists, according to #21, a mapping {Φ, Ψ} of M onto M′, mediated by, say, M″. If now A is an arbitrary subset of M,[5] a certain subset A′ of M′ will correspond to it under the mapping in question, and it is definite whether $A' \subseteq A$ or not. Thus all elements A of 𝔘M[6] for which we have both $Q \subseteq A$ and $A' \subseteq A$ are, according to axiom III [axiom of subsets], the elements of a certain subset T of P(M),[7] and, in particular, M is itself an element of T. The common component $A_0 = \mathfrak{D}T$ (#9)[8] of all elements of T now possesses the following properties: (1) $Q \subseteq A_0$ since Q is a common subset of all elements A of T; (2) $A'_0 \subseteq A_0$,[9] because every element x of A_0 is a common element of all elements A of T and its map $x' \in A' \subseteq A$ is thus also a common element of all A. On account of (1) and (2), therefore, also $A_0 \in T$. Finally we have (3) $A_0 = Q + A'_0$.[10] For, since $A'_0 \subseteq A_0$ and also $A'_0 \subseteq M' \subseteq M\text{-}Q$, on the one hand $A'_0 \subseteq A_0\text{-}Q$ [so $A'_0 + Q \subseteq A_0$].[11] On the other hand, however, every element r of A_0-Q is also an element of A'_0, and therefore $A_0\text{-}Q \subseteq A'_0$ [so $A_0 \subseteq A'_0 + Q$ and thus $A_0 = Q + A'_0$]. Indeed, if r is not an element of

4 The letter Q Zermelo may have picked from Poincaré's second inductive proof.

5 Zermelo seems to follow Dedekind in his use of A to denote an arbitrary subset of M and in his convention (#17, #21) to denote by A′ the image under the mapping from M to M′, of any subset A of M (1963 #21 p 51).

6 The power-set of M, which we will denote henceforth by P(M).

7 Besides this direct use of Axiom III, it is used indirectly in the definitions of intersection and equivalence.

8 𝔇T is the intersection of T, which we will denote by ∩T. The subscript 0 is used in the proof as a sign not as a number (see Sect. 21.2). Using the subscript '0' to denote the intersection of T is taken from Dedekind, though Dedekind would have denoted A_0 rather by Q_0 because it is what he called the 'chain of Q' (1963 #44).

9 Zermelo does not define A'_0. Dedekind defined the zero operator in general and proved that $(A_0)' = (A')_0$ (1963 #57).

10 This lemma is taken from *Zahlen* (1963 #58). Kanamori (2004 p 508f) points out that this lemma says that A_0 is the fixed point of the function f defined by $f(A) = Q + A'$.

11 The transitivity of the subset relation, stated in #3, is used here.

A'_0, then $A_1 = A_0$-{r} would still have A'_0, and a fortiori A'_1,[12] as a component, and, since it still includes Q, it would itself be an element of T, whereas it is in fact only a part of $A_0 = \cap T$.[13]

Therefore $M_1 = Q + M' = (Q + A'_0) + (M'\text{-}A'_0) = A_0 + (M'\text{-}A'_0)$,[14] where the summands on the right have no element in common, since Q and M′ are disjoint. But now, since A_0 is equivalent to A'_0 and M'-A'_0 is equivalent to itself, it follows according to #24 that $M_1 \sim A'_0 + (M'\text{-}A'_0) = M' \sim M$; that is $M \sim M_1$ as asserted.

Zermelo acknowledged (p 209 footnote 11) his (proof-processing) reliance on Dedekind's chain theory for his 1908b proof of CBT (as he did in his 1908a and 1901 papers). The applicability of the gestalt of chains and the pushdown metaphor to the context of CBT arises naturally because the conditions of the single-set CBT are the same as Dedekind's conditions for an infinite set, namely, that it be equivalent to one of its proper subsets (reflexive). Therefore, whatever Dedekind did in the context of his infinite sets could be associated to the context of the single-set CBT. Zermelo pointed out his indebtedness to Dedekind also in a remark that he added to Dedekind's proof of CBT, which he brought in Cantor's collected works. But even there, Zermelo seems to be blind to the CBT proof that Dedekind cached in his #63, which is identical to Zermelo's proof except that Dedekind took for Q the set M-M_1. Noether (Dedekind 1930–32 p 448), however, did make the connection. Since he was unaware of Dedekind's proof, Zermelo must have proof-processed the idea to combine two mappings on disjoint sets (#24), one of them the identity, from another source. It could have been Borel's (1898) proof of CBT or even Cantor's 1878 *Beitrag* where the idea is used (e.g., §6).

24.2 The Main Notions of Zermelo's Set Theory

Going through the theorems and proofs of CBT in the previous section, we encounter the following notions which we will explain below within Zermelo's theory (not necessarily in the following order): set, subset, equivalence, the image of a subset under equivalence (Dedekind's Lemma), transitivity of equivalence, part, component, union of two sets, difference of two sets, equality, a mapping of M onto N, set of pairs, definite, element, power-set, axiom III, the intersection of a set,

---

[12] The use of 1 in A_1 is also as a sign and not as a number. Note that A'_1 is not ambiguous; it can only mean $(A_1)'$ because 1 does not signify an operator.

[13] Peano's proof (the formal part that we omitted in Sect. 20.2) has another argument for $A_0 \subseteq A'_0 + Q$: since it was proved that $A'_0 + Q \subseteq A_0$ we have $(A'_0 + Q)' \subseteq A'_0 \subseteq A'_0 + Q$ and $Q \subseteq A'_0 + Q$ so obviously $A'_0 + Q \in T$ and hence $A_0 \subseteq A'_0 + Q$. Here the equality mentioned in footnote 9 above is necessary.

[14] Schröder may have influenced Zermelo's algebraic style in this string of equalities and equivalences. Note that the second equality is obtained by substituting $M' = A'_0 + (M'\text{-}A'_0)$ and not $M' = M' + (A'_0\text{-}A'_0) = (M' + A'_0)\text{-}A'_0$ which equals M'-A'_0 only.

transitivity of the subset relation, singleton set, the empty set, the associativity and commutativity of +, disjoint sets, combining two mappings, the identity mapping.

24.2.1 Sets and Elements

Zermelo assumes (#1) that there is a domain (*Bereich*, as in Schröder's *Denkbereich* – van Heijenoort 1967 p 228) of *Objekten*, translated 'individuals', which he calls, perhaps after Dedekind (1963 #1), *Dinge*, translated 'objects', not 'things' as in the 1963 translation of Dedekind *Zahlen*.[15] Of an object of the domain it can be said that it exists.

There is a fundamental relation called membership among the objects of the domain denoted by ∈ (#2). Membership comes with the domain and is not given by some law or association (Cantor 1932 p 204 (1), Ewald 1996 vol 2 p 916 [1], *Zahlen* #2). If a ∈ b, a is called an element or member of b and b is said to contain (*enthält*) a. An object that contains other objects is called a set. Zermelo's sets are thus given with the domain. Objects that contain no elements but are not the empty set, were later called 'urelements'.

There is one exception to the sets – the empty set: it contains no elements but it is a set and not an urelement. The existence of an empty set, denoted 0 by Zermelo, is postulated by axiom II of elementary sets (#4), together with the existence of singleton sets and pair sets. However, unlike the two other types of sets, the empty set is not obtained by an operation: it is a constant of the theory. Zermelo is confident in his use of the empty set, which Dedekind avoided (1963 #2) and Cantor never mentioned. Unlike Dedekind (see Sect. 9.2), Zermelo treats uniformly the trivial cases $M_1 = M$ or $M_1 = M'$ of CBT and thus the empty set is implicitly present in his CBT proof.

Another constant of the theory is the infinite set that consists of the sets 0, {0}, {{0}}, etc. The existence of this constant is not imposed by an axiom but derived (cf. Kanamori 2004 p 505) from axiom VII (of infinity, #13, #14) by taking the intersection of all infinite subsets of an infinite set that exists by the axiom (compare 8.1). In 1908a Zermelo considered 0, {0}, {{0}}, … as the number sequence. Later, in the mid 1910s, almost 20 years before von Neumann, he preferred the sequence 0, {0}, {0, {0}}, {0, {0}, {0, {0}}}, … – see Ebbinghaus 2007 p 133.

A class consists of the elements that satisfy a class-statement (see below). To the relation between a class and its objects, the membership terminology applies as between a set and its elements. Therewith the other notions regarding sets that are defined by the membership relation, such as subclass, disjoint classes, etc., become available to classes. Note that to a set corresponds the class of all the set's

[15] There is no apparent reason to use two names for the elements of the domain.

elements.[16] However, classes do not exist as do sets or their elements because they do not belong to the domain.

Given the domain, the axioms provide for the existence[17] of certain sets generated by certain operations from the pre-given sets. In other words, the axioms give certain closure condition on the domain (e.g., union). From these closure conditions others can be derived (e.g., intersection).

The axioms in Zermelo's system are not meant to give an implicit definition of sets, a view that is sometimes expressed. The idea of an implicit definition through axioms comes from Hilbert's structuralist program (Corry 2006 p 220, Taylor 1993 p 543). But Zermelo assumed the domain as a model of the theory, presumably after Dedekind's *Zahlen.* Thus the theory is assumed consistent, instead of proved consistent – a quest which Zermelo gave up on (1908b p 200f, 1929 p 341). Investigating the theory in relation to a model does not diminish the generality of the theory because it is not a specific model. For this reason we cannot say that the theory is complete (cf. Taylor 1993 p 544). By leaning against a model, Zermelo assumed a relativistic position[18]: his set theory does not attempt to describe a piece of reality and thus is contrary to the realistic attitudes of Kronecker, Cantor, Dedekind, Poincaré, Russell.[19] As a minor sign of Zermelo's relativism is his convention to mark results that are obtained using the axiom of choice or the axiom of infinity, allowing models not satisfying these axioms.

24.2.2 Subsets, Parts and Components; Transitivity of $\subseteq$

The subset relation and the relation of being disjointed are defined as usual by way of the membership relation (#3). They are associated perhaps because both invoke the gestalt of a member of one set belonging to another.

---

[16] The interpretation given in van Dalen-Ebbinghaus 2000 footnote 24 that Zermelo viewed the $\in$ relation between objects as the real membership relation, does not take into account this class, and it is not clear why membership in a set is more real than membership in its class.

[17] Actually Zermelo uses 'exists' (*existiert*) only for the singleton and the pair sets. For the empty set he says that it is 'given' (*es gibt*). He says that M possesses (*besitzt*) a subset (Axiom III) and that its union set and power-set correspond (*entspricht*) to it. The latter is especially strange because it implies a mapping from the entire domain into itself, clearly a non-Zermelian notion. This observation calls our attention to the fact that we could define certain classes of pairs to represent correspondences between classes. Therewith we realize that Zermelo's 1908b CBT is not the most general one and that a CBT for classes could be proved, if not by Zermelo's impredicative proof, that calls for T which would be a class of classes, at least by the inductive proof of, say, Borel, with its frames.

[18] That was, however, different from Skolem's relativism (van Dalen-Ebbinghaus 2000 p 153; cf. Moore 2009 p 827).

[19] Russell was willing to accept tentative axioms. The change brought about by Zermelo reflects perhaps a change in the philosophy of language at the turn of the twentieth century.

Zermelo uses Schröder's sign $\Subset$ for subsumption (p 201 footnote 3; the translation is 'inclusion' for the original '*Subsumptions*' p 262) which we have replaced by $\subseteq$. It is easy to see that the relation of subset is transitive.

A subset for Zermelo is not necessarily a proper subset. Here he adopted Dedekind's approach against that of Cantor. For a proper non-empty subset Zermelo uses the term 'part' (#6). If $\mathrm{M} \subseteq \mathrm{N}$ then by the class-statement '$x \notin \mathrm{M}$' the set $\mathrm{N}-\mathrm{M}$ can be separated using Axiom III and thus the difference between sets is defined.[20]

Unfortunately Zermelo uses the term *enthält*, which he defined to describe the relation between a set and its members, also for subsumption, that is, also between a set and its subsets. Thus in Axiom II (#4) it says with regard to pairs: "If a and b are any two objects of the domain, there always exists a set {a, b} containing (*enthält)* as elements a and b but no object x distinct from both". Clearly the words "as elements" in this statement were added for the ambiguity of "*enthält*". Likewise in #25 Zermelo says of a set M_1 that it "M' *als Bestandteil enthält*" but he does not mean that M_1 contains M' as member but includes it as a subset. Thus the words "*als Bestandteil*", which appear redundant in the translation ("M_1, that includes M' as component") are necessary in the German original.

The term 'Component' (*Bestandteil*) is not defined by Zermelo. In the above example it means no doubt 'subset', but usually it is used in the expression 'common component' (*gemeinsame Bestandteil*), a synonym to the term 'intersection' (#8). '*Bestandteil*' may have emerged from Cantor's 1878 *Beitrag* and his 1887 *Mitteilungen* where it is used for subset. In his *Grundlagen* and in his 1895 *Beiträge* Cantor used for subset *Teil* or *Teilmenge*, which was used by Zermelo in 1901, while in 1908a he uses *Untermenge*.

24.2.3 Equality

Zermelo introduces equality of sets in two ways. First Zermelo says (#1): "if two symbols, a and b, denote the same object, we write a = b". Then (#4) he introduces axiom I (of extensionality – literally 'determination' *Bestimmtheit*) which says: $\mathrm{M} = \mathrm{N}$ when $\mathrm{M} \subseteq \mathrm{N}$ and $\mathrm{N} \subseteq \mathrm{M}$. By extensionality there is only one empty set; extensionality does not apply to urelements.

After he introduces axiom II (of elementary sets (#4)), which warrants that to every set, its singleton exists, Zermelo reduces the first definition to the membership

[20] Zermelo defined the difference between sets only in 1908a (van Heijenoort 1967 p 183f), neither in his 1908b nor in his 1901. Still he used it in the proof of CBT above. Cf. Borel's definition of complementary subsets mentioned in Sect. 16.1.

relation by saying: a = b iff a ∈ {b}.[21] Thus if a, b denote the same set, a belongs to every set that contains b. The first notion of equality of Zermelo is considered a confusion of use and mention of symbols (Fraenkel et al. 1973 p 25–27). The right to replace one symbol with another, when the two mean the same entity, is given to us by the rules of language not by the rules of set theory. It is, however, not worthless to note that the first type of equality is used in nominal definitions, when we assign a symbol to a set generated by a certain procedure. On the other hand when we equate the symbols of defined sets we make a statement about their equality that needs to be proved by extensionality. For example, when we wrote in the proof of the single-set CBT, Q = N_1-M, we made a definition but when we wrote $A_0 = A_0' + Q$, we made an assertion which requires proof by extensionality.

24.2.4 Definiteness

Working against a model gave a semantical, rather than formal, dimension to Zermelo's object language, which provides the class-statements. Of a class-statement Zermelo demands (#4, #6) that it be 'definite', namely, such that "the fundamental relations of the domain, by means of the axioms and the universally valid laws of logic, determine without arbitrariness whether it holds or not" for any object of the domain.[22] As examples of definite assertions Zermelo gives: (#4) a ∈ b or not, M ⊆ N or not,[23] (#5) a = b or not,[24] (#13) M = {a} or not,[25] (#16) M ~ N or not,[26] mentioning in passing that it is also definite whether a set is empty or not.[27] Zermelo also defines when a class-statement $\mathfrak{E}(x)$ is definite in a set: when it is definite for each element of the set (#6). The class-statement that separates a subset from a set according to axiom III, must be definite in the set.[28] The

[21] The set {b} is the singleton of b. It contains only b. Its existence is warranted by Axiom II of elementary sets.

[22] Thus the domain is a prerequisite to the introduction of definiteness.

[23] M ⊆ N is definite because for every x in M, x ∈ N or not is definite (cf. Kanamori 2004 p 519, van Dalen-Ebbinghaus 2000 p 149, and Taylor 2002).

[24] Sets equality is clearly definite because it is established by ∈.

[25] This statement is composed of a ∈ M and $x \neq a \rightarrow x \notin M$, and all the component statements are clearly definite.

[26] If M, N are disjoint, M ~ N or not is equivalent to EQ(M, N) is not-empty or it is empty, which is definite by the next footnote. EQ(M, N) is defined in the next section. If M, N are not disjoint, then M ~ N amounts to saying that M ~ R where R is the set disjoint from both M, N and equivalent to N, defined in the next section. So again M ~ N or not, is definite.

[27] The empty set is a constant of the domain. It is directly identified and so the question whether a set is the empty set or not is directly verifiable.

[28] Since in axiom III the range of x is taken to be a set, it may appear that the notion of class is redundant in Zermelo's axiom system (cf. van Dalen-Ebbinghaus 2000 p 153). But the domain, which is clearly a class, is not a set (by Russell's Paradox) so that it appears better to leave the notion of class.

requirement for definiteness entails that assertions cannot contain terms not reducible to ∈, such as the terms of the semantical paradoxes (Kanamori 2004 p 541).

The notion of definiteness was criticized (Ebbinghaus 2007 p 88) through the years but Zermelo defended it in his later papers (cf. van Dalen-Ebbinghaus 2000 p 152). The criticism seems to have been more for reasons of vogue, the interest in problems of first-order logic, than any failure of Zermelo's approach.

24.2.5 Equivalence and Related Notions

Zermelo adopted the approach, slowly developing in mathematics through the end of the nineteenth century, that a mapping is a table but he transferred it into the context of set theory by regarding the table as a set of pairs. Thus Zermelo parted from Cantor's conception of equivalence as a mapping defined by a law (*Grundlagen* §13). Probably because Zermelo did not want to introduce the notion of order at the foundations level of set theory he refrained from using ordered-pairs. For this reason he had to define equivalence for disjoint sets first, addressing the problem of equivalence of non-disjoint sets later.

Given two sets, there are many mappings between them. Zermelo sought not to establish one specific mapping but the existence of the set of all these mappings, from which a desired mapping can be chosen in arbitrary.[29] Zermelo approached the desired set by a top-down separation process using his axiom III of subsets or separation.

Zermelo begins with two disjoint sets M, N. The set {M, N} exists by axiom II of elementary sets which warrants, for any two different sets, the existence of the set that contains both sets and only them. The existence of the union of {M, N}, denoted by M + N,[30] is provided by axiom V of the union. The union has as its elements all elements contained in either M or N. Next Zermelo takes the power-set of the union, provided by axiom IV of the power-set. The class-statement 'the intersection of x and M is a singleton and the intersection of x and N is a singleton' or not, is definite for every x in P(M + N) (#13). By axiom III the subset of P(M + N) of all such x exists; Zermelo denotes it by MN and calls it the product of M, N or the connection set of {M, N}.[31]

---

[29] Prior to making such a choice one must show that the set is not empty. For this task CBT is often handy. Like the set of mappings, the set of selectors (from a family of sets) can be defined. But there is no similar device to prove that it is not empty so the axiom of choice must often be involved. Perhaps Hilbert's proof in the theory of invariants gave Zermelo the pattern of his top-down metaphor: first define the set of all, then prove that it is not empty, then choose a representative in arbitrary.

[30] The sign + for union Zermelo probably took from Schröder (see Sect. 9.2.2).

[31] It is easy to verify that MN = NM.

Let $\Phi \subseteq$ MN. For all $x \in$ M + N it is definite whether $(\cup\Phi)\cap\{x\} = \{x\}$ or not.[32] Let Φ(M, N) be separated from M + N by this class-statement. It is now definite if Φ(M, N) = M + N or not and so we can separate from MN the set of all Φ such that Φ(M, N) = M + N. We denote this set by EQ(M, N). It is the set of all (equivalence) mappings between M and N.

We say that M, N are immediately equivalent (#15), M ~ N, if M, N are disjoint and EQ(M, N) is not empty. A member of EQ(M, N) Zermelo called (p 205) "a mapping of M onto N".[33] We say that two sets are mediately equivalent (#21) if there exists a set R disjoint from both such that M ~ R and R ~ N. When two sets M, N are mediately equivalent, there is a pair of mappings {Φ, Ψ} such that Φ is a mapping from M onto the mediating set R and Ψ a mapping from R to N. In short Zermelo says that {Φ, Ψ} is a mapping of M onto N. Both immediate and mediate equivalences Zermelo denotes by ~.

Given two sets M, N there is always a set R that is disjoint from M, N and equivalent to M (#19). Let S = ∪(M + N)[34] and r a set that does not belong to S, e.g., its Russell's set: the set of all elements of S that do not contain themselves. This set can be separated from S because the statement $x \in x$ or not is definite (#10). The connection set R = M{r} is disjoint from both M and N and is equivalent to M. We can have R disjoint from any finite number of sets and immediately equivalent to one of them. In this way if M is mediately equivalent to N by R′ and N mediately equivalent to P by R″, if we take R disjoint from all five sets and immediately equivalent to P, then R is immediately equivalent to each of the five sets and thus M is mediately equivalent to P, so mediate equivalence is transitive. Two sets that are immediately equivalent are also mediately equivalent. Thus, the transitivity of immediately equivalence is provided by the transitivity of mediate equivalence.

By the way mappings are defined, there is no identity mapping because its pairs degenerate to singletons. We can make a special provision for this case, allowing a mapping composed of singletons or agree that a set M is only mediately equivalent to itself. If M ~ N and P ~ Q and all sets are disjoint, then for every mapping Φ from M onto N and mapping Ψ from P onto Q, Φ + Ψ is a mapping from M + P onto N + Q. With Ψ the identity, this is how the final step of the CBT proof was obtained.

---

[32] Zermelo denotes the union of a set T, which exists by Axiom V, by 𝔖T. We use ∪T. It is easy to see from the definition of the union that it is commutative and associative.

[33] This formulation appears unfortunate because we have no means to distinguish with regard to a member of EQ(M, N) if it is from M onto N. A member of EQ(M, N) is a mapping between M and N. Despite our reservation, we stick to Zermelo's language.

[34] The brackets here are regular brackets to distinguish the expression from ∪M + N. In van Heijenoort English translation the expression is rendered as ∪{M + N} which is simply M + N. S is said there to contain the elements of M + N while in the original it says that S contains the elements of the elements of M + N.

Dedekind's Lemma is the observation that when M is equivalent to N, to a subset of M corresponds a subset of N (see Sect. 10.1). Here is the proof: If M, N are immediately equivalent there is a mapping Φ of M onto N. Let Φ' be the subset of Φ that contains all $\{m, n\} \in \Phi$ such that $m \in M' \subseteq M$. The existence of Φ' by the subset axiom rests on the fact that $\{m, n\} \in \Phi$ or not and $m \in M'$ or not, are both definite. Let now N' be the intersection of $\cup\Phi'$with N. If M, N, are mediately equivalent by R then for every $M' \subseteq M$ there is an $R' \subseteq R$ such that $M' \sim R'$ and for R' there is $N' \subseteq N$ such that $R' \sim N'$. By the transitivity of equivalence $M' \sim N'$.

24.2.6 Intersection

The closure of the domain under intersection, which Zermelo used in the single-set formulation CBT to define A_0, and in the proof that a constant infinite set exists in the domain, he proves in two steps. If M and N are two sets then for every a in M it is definite if $a \in N$ or not. So by axiom III, *a* subset of N can be separated (it may be empty) that contains all members of M that belong to N. This is the intersection of M and N (#8).

This definition can be extended to any set T that is not empty (#9): To any object a it is definite for every $N \in T$ whether $a \in N$ or not and so a subset of T can be separated, denote it by T_a, of all N such that $a \in N$. For any subset S of T it is definite if $T_a = S$. So the class-statement 'x is such that $T_x = T$' is definite and if A is an arbitrary element of T, a subset of A can be separated of all the elements of A that fulfill this statement. This subset is the intersection of T, which we denote by $\cap T$. It is to enable the choice of A that Zermelo notes that T is not empty.

Note that the set of all T_a is not necessary for the definition of $\cap T$, nor is it necessary to have a universal quantifier in the class-statement applied, such as 'x is a member of each member of T' (Fraenkel 1966 p 20, Fraenkel et al. 1973 p 36–39), though Zermelo later permitted quantifiers in definite statements (1929 p 343).

24.3 Comparison with the Proof in Poincaré 1906b

The notation of Zermelo's proof presented by Poincaré (see Sect. 19.5) is probably not that of Zermelo's letter (which we haven't seen) for it appears streamlined with another proof of CBT which Poincaré presented just prior to Zermelo's proof (see Sect. 19.2). R of that proof is A_0 here. But there it was the chain generated by H_0, not Q_0, that was "pushed down".

Poincaré does not mention T (see Sect. 24.1) and its relation to the axiom of subsets. Instead Poincaré calls 'B set' any set which satisfies the conditions of belonging to T. When Zermelo wrote his letter to Poincaré he had already conceptualized the axiom of subsets (Ebbinghaus 2007 p 82f), but it makes sense that Zermelo did not want to burden his letter with a discussion on the axioms and he used a language similar to the one he used in his first proof of the Well-Ordering Theorem ('γ-set').

Poincaré's rendering of Zermelo's proof is missing also other details that appear in the 1908b paper. Instead of the argument in Zermelo's proof to justify $A_0 = Q + A'_0$, Poincaré is satisfied in saying (p 314) that otherwise "$Q + A'_0$ would be a B set that does not contain A_0, because on the contrary it would be its part", implicitly maintaining that this is a contradiction as A_0 must be contained in every B set. Poincaré also omits the point that the identity is combined to the $'$ mapping (which Poincaré denotes by φ) in order to obtain the equivalence of M_1 and M'.[35] The reason seems to be Poincaré's haste to get to the point of impredicativity, which, with regard to Zermelo's proof, Poincaré finds in the definition of A_0: because A_0 belongs to T there seems to be is a vicious circle in its definition.

Following the proof of CBT Zermelo presents a simple corollary (#26):

> If a set M is equivalent to one of its parts it is also equivalent to any set M_1 that is obtained from M when a single element is removed or added. [We skip the proof, which is not difficult.]

The reason that Zermelo brought this corollary seems to be pinned in Poincaré 1906b. There (p 308), Poincaré defined 'finite number' to be the cardinal number of a set that is not equivalent to any of its subsets. He also defined "finite integer" to be a cardinal number n that is not equal to n-1. It is obvious, Poincaré noted, that every finite number is also a finite integer. However, the reciprocal result requires CBT. So we see that with this corollary Zermelo provides the argument for Poincaré's contention. Zermelo may have devoted attention to this issue because at the time he wrote his 1908b he also wrote his 1909, which demonstrated how the theory of finite numbers and complete induction could be established in set theory, contrary to Poincaré's views (1909). The above corollary can be extended to any set B that is equivalent to a subset of M-M$'$ (certain restrictions apply in the subtraction case). Note the convex-concave aspect of this result (see Sect. 13.3).

24.4 Final Words

When first published by Poincaré, Zermelo's proof was certainly grasped as beautiful and novel because of the way it avoided use of the natural numbers. The fact that a similar proof was presented by Peano may have only added to the

[35] The same omission is made by Poincaré with regard to the preceding, Whitehead-Russell proof.

strength of the chain gestalt employed by both proofs. In the 1908b version of the proof Zermelo added the analysis of what set theoretic notions are necessary for the proof, which avoids the use of natural numbers and complete induction. It was not only the impredicative nature of the proof, this criticism by Poincaré was handled successfully by Zermelo and Peano, but its use of such heavy machinery as the power-set axiom, the union axiom and the separation axiom that clarified how much mathematics is necessary to by-pass the natural numbers. Still it left outside CBT the classes that could be handled by the inductive proofs.

Chapter 25
Korselt's Proof of CBT

In 1911, A. Korselt[1] published in *Mathematische Annalen* a paper (dated April 1911) showing that Schröder's 1898 proof of CBT is erroneous (see Sect. 10.2). In his paper Korselt brought a proof of CBT (see the next section), which, he says, he had communicated to Schröder, and "in somewhat different form" (p 296), sent to *Mathematische Annalen* on May 30, 1902. Korselt does not say if his 1902 paper included his criticism of Schröder's proof and he does not explain why his 1902 article was not published. If the criticism of Schröder appeared in that paper we can perhaps assume that the death of Schröder (all the more if it was rumored that the death was linked to the mistake found in his proof of CBT) affected both the journal and Korselt in a way to delay a critical paper of a respected colleague. Otherwise, if the paper only contained the proof of CBT, it could be that it was not considered worthy of publication in the days preceding Poincaré's challenge. We have asked the archivist of the *Mathematische Annalen* if Korselt's 1902 paper could be located and were answered in the negative.

After mentioning his 1902 paper Korselt stated that "later Zermelo and Peano found similar proofs. They all have in common that they do not presume the integers and complete induction." Korselt refers, no doubt, to the proofs provided by Zermelo (see Sect. 19.5 and Chap. 23) and Peano (see Sect. 20.2) that emerged in response to Poincaré's challenge. Korselt's statement implies that he thought that his proof too avoids these notions. Maybe this statement was intended to establish both the independence of his proof and its priority over the proofs of his famous colleagues, especially against the background of Poincaré's challenge. It would indeed be interesting to learn that Korselt has some priority over Zermelo and Peano for their 1906 CBT proofs, which had historical significance. However, Korselt's claim is not justified by the proof that he provides. Though it seems that

[1] On the biography of Korselt see Kreiser 1995, where Peckhaus 1986 is mentioned, a work that we could not locate. Korselt was a high-school teacher and obtained his doctorate when he was nearly 40. It seems that he corresponded with Schröder before 1902; he also corresponded with Frege (see Frege 1971).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_25, © Springer Basel 2013

Korselt had in mind use of a chain of frames, his proof does not indicate how the chain is obtained, whether by induction or by an impredicative definition through the use of intersection. Our impression is that Korselt's proof is closer to J. Kőnig's 1906 proof of CBT, who is not mentioned by Korselt.[2]

Korselt refers to CBT as the Equivalence Theorem, like the veterans, Bernstein and Zermelo, who followed Cantor. Korselt gave his proof for the single-set formulation of the theorem, about which he said that it is equally valid (*gleichgeltenden*) with the Equivalence Theorem, which is a somewhat awkward way to say that the two formulations are equivalent.

In the presentation of the theorem and proof Korselt uses Schröder's notation from his 1898 paper; the sign $\Subset$ used by Schröder and Korselt for both subsumption and implication we replace by $\subseteq$ and $\rightarrow$, respectively.

Korselt also provided an application of the method of his proof of CBT to answer a problem raised by Schröder (1898 p 326). We will present this result too, covering as well similar results obtained by members of the Polish school of mathematics in the 1920s.

25.1 Theorem and Proof

We bring Korselt's proof in its entirety with our comments in the footnotes:

> $(a_2 \subseteq a_1 \subseteq a_0 \sim a_2) \rightarrow (a_0 \sim a_1)$.
>
> By the given assumption, let φ be a one-one mapping (*deutliche Abbildung*) from a_0 in[3] a_2, namely $\varphi(a_0) = a_2$, $\breve{\varphi}$ the reverse of the binary relation φ,[4] a_0' be the set of all those (initial-) elements of a_0, for which $\breve{\varphi}(a_0') = 0$.[5] To every element or subset m of a_0 belongs through φ a "φ-following" m_φ (comprised of finite or infinite "cycles" or denumerable infinite "chains").[6] The φ-following of different elements are disjoint due to the 1–1 nature

[2] Medvedev (1966 p 236) thought that Korselt's proof relied indeed on Dedekind's chain theory of *Zahlen*, as did the proofs of Zermelo and Peano. He also accepted Korselt statement that he had the proof since 1902, thus giving him priority over both.

[3] Should be on.

[4] Why Korselt switches here to Schröder's language of binary relations is not clear. We may use φ^{-1} instead of $\breve{\varphi}$.

[5] 0 surely denotes the empty set but it is not explicitly introduced by Korselt. The elements of a_0' are the elements that are not an image under φ of another element in a_0, so they are properly called the initial elements of a_0. Simply put, $a_0' = a_0 - a_2$, but it seems that Korselt, like Dedekind and Schröder, avoided use of the difference operation.

[6] The φ-following of *m* contains not just the image of *m* but also the image of the image, and so on. This comes out from the language in the brackets. Moreover, the mention of infinite cycles, surely J. Kőnig's left-extendible strings, suggests that φ-followings include images under φ^{-1}. The φ-following of a set is comprised of the φ-followings of its elements but the latter need not be of the same type as the former. Korselt makes no use of the φ-following of a set in the proof. The φ-followings of elements (sets) may be contained in the followings of other elements (sets). The φ-following of a set may contain just the set; this happens when the φ-followings of the elements of the set are all cycles. Korselt's φ-following combines into one gestalt the chain gestalt of Dedekind and the string gestalt of J. Kőnig (in a setting of the single-set formulation). Korselt did not explain how the φ-followings are conceived: by induction or by intersection.

(1–1-*Deutigkeit*) of φ.[7] So, every element of a_0 specifies either an infinite φ-chain (as an initial member of φ) or a φ-cycle (as an inner member of φ).[8]

There exists then a most comprehensive subset d of a_0 (it could be empty) that is partitioned by φ into cycles, and $d \subseteq a_2 \subseteq a_1 \subseteq a_0$.[9] The set a_2 consists therefore of the inner chain-elements and the elements of the cycles, and these together with certain initial members of the φ-chains build the set a_1.[10] With help of the presumed mapping φ, the 1–1 mapping φ′ from a_1 onto a_0 would be thus compounded:

All members of d will correspond to themselves, likewise any member of a chain, which has an a_1 [a member of a_1] as its starting member. However, to every member of a different chain (in it only the inner members are elements of a_1) the following element will be assigned.[11] By binary relations this can be expressed much shorter.[12]

Korselt's proof uses J. Kőnig's strings gestalt and the 'partition and pushdown the chain' metaphor published by Zermelo and Peano. But Whether Korselt had priority in his use of these descriptors, remains an unanswered question. Obviously, by 1911, his proof provided no new insight on CBT.

25.2 An Application

Following his CBT proof, Korselt comments that by partitioning into chains, a question from Schröder's paper (p 326) can be answered. It concerns the following theorem (p 317): (21) $(a \sim d \subseteq b) \rightarrow \exists c(a \subseteq c \sim b)$, which is a dual to Dedekind's Lemma (19) $(a \subseteq c \sim b) \rightarrow \exists d(a \sim d \subseteq b)$. Each of these theorems is obtained from the other by interchanging the ~ and ⊆. Schröder's problem with (21) was, apparently, to obtain c without adding new elements not in a + b. Korselt suggested to take $c = a + b$, and the equivalence ψ from c onto b he defined as equal to φ, the equivalence from a onto d, for all members of the φ-followings of members in a-b

[7] As stated, this lemma is false, for different elements of a_0 may belong to the same φ-following. A correct formulation would be that two φ-followings are either disjoint or one is contained in the other. The proof of this lemma seems to require complete induction unless the φ-following is defined impredicatively. So Korselt's contention that he does not use complete induction must be doubted, at least with regard to his 1902 proof; by 1911, Korselt was surely aware of the impredicative definition option.

[8] This assertion is not precise for an element of a_0 can belong to a chain without being its first element. Korselt does not define 'inner element', which are, no doubt, φ images.

[9] d is the union of all φ-cycles. d is obtained from a_0 after all φ-chains are removed.

[10] These initial elements are the members of a_1-a_2.

[11] This definition is for a mapping from a_0 onto a_1 and not the other way around as declared by Korselt with regard to φ'. The mapping defined is the same mapping defined in Dedekind's proof (see Sect. 9.2).

[12] In Schröder 1898 a binary relation is a set of pairs. Korselt could have meant the binary relation that is constructed by having the pairs of the φ-following of members of a_0-a_1 and for all other members of a_0 the identity pair.

and the identity on all other members. Let us call Korselt result, the Schröder-Korselt Theorem.[13]

The Schröder-Korselt Theorem was revisited in Lindenbaum-Tarski 1926. There the following Theorem 14 (p 302) is stated: When $A \subseteq B \subseteq C$, $A_1 \subseteq C$ and $A \sim A_1$, there exists a set B_1 such that $A_1 \subseteq B_1 \subseteq C$ and $B \sim B_1$. In a footnote, which appears to have been added in proof,[14] Korselt's proof of Schröder's question is referenced as a sketch of a proof for Theorem 14. Indeed, if in the Schröder-Korselt result we place A_1, A, B, C, B_1, for a, d, b, a + b, c, respectively, the Schröder-Korselt result proves Lindenbaum-Tarski's Theorem 14.

In Lindenbaum-Tarski a proof of Theorem 14 is attributed to Lindenbaum.[15] The proof is said to be not by way of CBT, hinting perhaps at the connection between Korselt's proof and CBT, but based directly on properties of 1–1 mappings such as Theorem 3 of §2 of the same paper (p 318), which is related to Banach's Partitioning Theorem (1924, see Chap. 29). A "dual" of Theorem 14, Theorem 14 bis, is also mentioned in Lindenbaum-Tarski 1926; in it the relations of subsumption are reversed, giving: $C \subseteq B \subseteq A$, $C \subseteq A_1 \sim A$ entail that there is B_1 such that $C \subseteq B_1 \subseteq A_1$ and $B_1 \sim B$. It is remarked that this theorem generalizes CBT. Indeed, putting $A_1 = C$ we get $B_1 = A_1$, hence that $B \sim A$.[16]

Theorem 14 and Theorem 14 bis, were generalized in Theorem 15 of Lindenbaum-Tarski 1926, marked as proved by Tarski; it states: When $A \subseteq B \subseteq C$, $A_1 \subseteq C_1$, $A \sim A_1$, $C \sim C_1$, there exists $A_1 \subseteq B_1 \subseteq C_1$ such that $B \sim B_1$. Using Dedekind's Lemma (Schröder's Theorem (19) mentioned above) it is easy to prove Theorem 15 from Theorem 14. Cf. Tarski 1949a p 16 footnote 4, p 236 footnote 27. This theorem was named in Tarski 1930 (p 248) as the Mean-Value Theorem [*Mittelwertsatz*] (see Sect. 35.3).

In 1947, Sierpiński (1947b) proved the following theorem: if M is a set of power $\mathfrak{m}$ and P a subset of power $\mathfrak{p}$ and if n a cardinal number such that $\mathfrak{m} \geq \mathfrak{n} \geq \mathfrak{p}$ then there exists a set N of power n such that $P \subseteq N \subseteq M$. Sierpiński's proof follows the proof pattern of Zermelo's CBT proof of 1908. An alternative is to apply J. Kőnig's string gestalt used by Korselt as follows: Let Q be a set of power n, φ a 1–1 mapping from P into Q and ψ a 1–1 mapping from Q into M. These mappings exist because $\mathfrak{m} \geq \mathfrak{n} \geq \mathfrak{p}$. Then if ψ(Q) is disjoint from P we can interchange in ψ(Q) between $\psi\varphi$(P) and P and obtain the desired result. Otherwise, we will change ψ as follows: for $x \in P$, if for some finite n, $(\psi\varphi)^n(x)$ is not in P we interchange $\psi\varphi(x)$ with $(\psi\varphi)^n(x)$. Let D be the set of all x in P for which $(\psi\varphi)^n(x)$ is always in P. Then $\psi\varphi(D) = D$. Now we can replace in ψ(Q-φ(D)) the part $\psi\varphi$(P-D) with P-D. The set resulting from ψ(Q) in this way is the desired N.

[13] This simple theorem is a good example of proof-processing: Korselt is not relying on a theorem to obtain the result but on a technique, which he applied in his CBT proof. The conditions of (21) are reminiscent of the conditions of the single-set formulation of CBT and so the technique applied for CBT is recalled to obtain a proof of (21).

[14] Apparently, Lindenbaum proved theorem 14 before he became aware of Korselt.

[15] In Lindenbaum-Tarski 1926 no proofs were provided.

[16] Cf. Kuratowski-Mostowski 1968 p 191ff.

Sierpiński tells that a comment by Mostowski made him aware that the theorem is an immediate consequence of Theorem 15. Therefore, we can actually regard Sierpiński's proof to be an alternative proof of Theorem 15, by direct means as attributed to Lindenbaum's proof. Both Sierpiński's proof and the proof given here are examples of proof-processing, because in them the gestalt and metaphor acquired in one context (CBT) is applied to another.

Chapter 26
Proofs of CBT in Principia Mathematica

In Principia Mathematica (PM, 1910–1913), the monumental treatise of A.N. Whitehead and B. Russell (together 'WR'), the logicist movement obtained its most clear statement. The treatise contains four formulations of CBT[1] with their proofs, which we will review below. The first two formulations appear in *73:[2] In *73.85 the single-set formulation is given and in *73.88 the two-set[3] formulation follows. The second formulation is proved by the first, in a standard fashion; the proof of first formulation is adapted from Zermelo's proof in his paper on axiomatic set theory (1908b, see Chap. 23), which WR reference. It was, no doubt, because this proof avoids the notion of number that WR preferred it, for by *73 the notion of number was not yet introduced in PM. We present the first two formulations in Sect. 26.1 and their proof of *73.85 in Sect. 26.2.

Because of Poincaré's criticism that Zermelo's proof is impredicative we have so dubbed PM's proof of *73.85, even though this name is not very appropriate in the context of PM: In PM the intersection used in Zermelo's proof, is of a different type from the type of the members of the class intersected, and so cannot belong to it and no impredicativity arises. Still, in the intersected set, a set with the same extension as the intersection exists, due to the axiom of reducibility, so Zermelo's proof stands. The situation, however, changed in the second edition of PM (1925–1927), when the axiom of reducibility was dropped. So there, in the introduction, the impredicative proof was changed (vol I, pp. xxxix–xlii), to remove its dependence on the axiom, see Sect. 26.3.

[1] WR name it the 'Schröder-Bernstein theorem'.

[2] Sections are numbered in PM with a preceding * and theorems within sections are numbered by the section number followed by a dot followed by the theorem's number.

[3] We generally follow WR in using 'class' for Cantor's 'set', disregarding Cantor's distinction between consistent and inconsistent sets, the latter are commonly now called classes. Still, when using a common expression in which 'set' appears, we maintain the expression and avoid translating it to the language of classes.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_26, © Springer Basel 2013

The third formulation appears as an endnote to section *95 and is not numbered. Its proof, scattered in *93, *94, *95, is based on Borel's proof (1898, referenced by WR), which emerged from a proof of Bernstein (see Chap. 11). Like Borel's proof, PM uses complete induction but, as this proof too is placed in PM before numbers are introduced, it is not the standard complete induction that is used but an original procedure based on hereditary classes and ancestral relations. For its use of inductive reasoning we call the proof of the third formulation the inductive proof.

After providing a proof of CBT within their system, the impredicative proof, one may wonder why WR bothered to provide yet another proof. The answer we suggest is that the inductive proof gave WR the opportunity to demonstrate the power of their logic of relations, especially the notions of hereditary classes and ancestral relations, developed by Russell since his 1901 paper (1956 p. 16). For this reason it seems to us that the inductive proof is the focal point of the first volume of PM. Cf. our remark on the place of CBT in Zermelo's axiom system (Chap. 24). We thus disagree with Kanamori (2004, p. 511) regarding the "prolonged and gratuitous labor" that was put into it.

We bring the inductive proof in Sect. 26.4. We do not discuss the effect of the removal of the reducibility axiom on the third proof, since WR did not address this question directly. Thus we avoid the renewed discussion of induction contained in Appendix B of the second edition (cf. Urquart 2003, p. 298).

In *94 WR give an informal discussion of the impredicative and inductive proofs, accompanied by drawings – a rarity in the voluminous PM. We cover this discussion in Sect. 26.5 where we add to PM's drawing variations of our own.

The fourth formulation of CBT in PM is *117.23. Earlier, it was only Zermelo (not Cantor), in his 1901 paper, who gave CBT in the language of cardinal numbers. However, Zermelo used Cantor's definition of cardinal numbers, by abstraction, which ultimately always requires regression to the language of sets and mappings (classes and relations in the language of PM) for its basic applications. WR use Frege's definition of cardinal numbers as classes of equivalent classes adapted to their type-theory. They are thus able to deduce the properties of cardinal numbers without regressing to the language of classes and relations because for them cardinal numbers are classes to begin with. Therefore, instead of mappings, or relations, WR use intersection. We describe the proof of this cardinal version in Sect. 26.6. It seems to us that the cardinal version is not affected directly by the elimination of reducibility.

In Sect. 26.7 we compare PM's CBT proofs to the previous proofs, covered in earlier chapters.

PM is a very elaborate and rigorous[4] product but its formalism can be tiring: 16 lemmas of section *73 are used for the impredicative proof and several dozens more from previous sections; even more numerous is the number of lemmas used for the inductive proof. For this reason, though it is our aim to present the proofs of CBT in detail, we will present the arguments in PM only up to lemmas that we deem intuitive.

[4] Notwithstanding Gödel's correct criticism of PM's lack of "formal precision in the foundations" (Gödel 1944 p 448).

26.1 The First Two Formulations

In *73, titled “similarity of classes”, CBT is presented as:

*73.85 $\vdash : \mathrm{R} \in 1\rightarrow 1 \,.\, \text{ɑ‘R} \;\mathbf{C}\; \beta \,.\, \beta \;\mathbf{C}\; \text{D‘R} \,.\supset.\, \beta \text{ sm } \text{ɑ‘R} \,.\, \beta \text{ sm } \text{D‘R}$

*73.88 $\vdash : \alpha \text{ sm } \gamma \,.\, \beta \text{ sm } \delta \,.\, \gamma \subseteq \beta \,.\, \delta \subseteq \alpha \,.\supset.\, \alpha \text{ sm } \beta$

‘$\vdash$’ is the assertion sign used in PM, following Frege, to signify an assertion, whereas definition has ‘Df’ at the end. Here small Greek letters refer to classes (*20; cf. vol. I, pp. 23, 47, 199); in *100 and after, these may refer in particular to cardinal numbers. A class consists of all the terms that satisfy a propositional function (p. 38), namely, that turn it into a true proposition. Capital Latin letters denote relations. A relation can be regarded as a class of ordered-pairs (Wiener 1914; Grattan-Guinness 1975). When (x, y) is in R, which is also denoted by $x\mathrm{R}y$, we say that x is a referent to the relata y. ‘1→1’ denotes the class of all 1–1 relations (*71.03). These are relations in which to every referent there is one and only one relata and to every relata one and only one referent. D‘R is the domain of R (read ‘as ‘of’; cf. *30 and Sect. 26.6 below); it consists of all terms that appear as referents, on the left side of R. ɑ‘R is the range (WR’s converse domain) of R, ɑ‘R consists of all terms that appear as relata, on the right side of R. ‘sm’ (*73.04) stands for ‘similarity’, namely, α sm β holds when there is a 1–1 relation, the domain of which is α and the range β. ‘Similarity’ is what Cantor called ‘equivalence’ and denoted by ‘ ~ ’, a symbol that WR use for negation. The dots in *73.85, 88 are Peano’s signs for ‘and’ and brackets; ‘.⊃.’ denotes implication. **C** is PM’s inclusion sign (*22.01). When the range of a relation is a subset of its domain, as in *73.85 when there exists a β that satisfies the hypothesis, R is called a reflection (vol. I, p. 589). The domain and range of a reflection are similar, so in *73.85 one of the conjuncts of the thesis is redundant, sm being transitive (*73.32).

We will use sm and ~ as WR do but in general we will use prevailing notations for the logical connectives and set theoretic notions. Thus *73.85, 88 are in our notation:

*73.85 $\vdash \mathrm{R} \in 1\rightarrow 1 \;\wedge\; \beta \subseteq \text{D‘R} \;\wedge\; \text{ɑ‘R} \subseteq \beta \rightarrow \beta \text{ sm } \text{D‘R} \;\wedge\; \beta \text{ sm } \text{ɑ‘R}$;

*73.88 $\vdash \alpha \text{ sm } \gamma \;\wedge\; \beta \text{ sm } \delta \;\wedge\; \gamma \subseteq \beta \;\wedge\; \delta \subseteq \alpha \rightarrow \alpha \text{ sm } \beta$.

*73.88 resembles CBT’s two-set formulation. It uses classes and the relations of similarity while it does not provide the relations that carry the similarities. Fortunately it avoids the hybrid language used by WR in 1902 paper, which contained cardinal numbers. *73.85, on the contrary, does not resemble the single-set formulation, which is stated with sets instead of the domain and range of a relation used here.

WR say that the proof for *73.88 is due to Zermelo (PM p. 478) and they reference (p. 487) his 1908b paper. Actually it is *73.85 that is proved by Zermelo’s method; *73.88 is proved in the familiar way by shifting it to *73.85. WR do not mention, not even in *94 where an informal discussion of the proofs is presented, the reason for bringing Zermelo’s proof: that it avoids the notion of number. Note that while Zermelo acknowledged his indebtedness to Dedekind (1908b, p. 209 footnote), WR failed to make a similar mention.

26.2 The Impredicative Proof

Since the derivation of the two-set formulation *73.88 from the single-set formulation *73.85 is standard (see Sect. 24.1) we will concentrate on the proof of the latter. As first lemma in this proof WR give:

*73.8 ⊢ Cl‘R ⊆ β ∧ β ⊆ D‘R ∧ κ = {α | α ⊆ D‘R ∧ (β-Cl‘R) ⊆ α ∧$\breve{\text{R}}$“α ⊆ α}
→ D‘ R ∈ κ ∧ p‘κ ⊆ D‘R

$\breve{\text{R}}$ signifies the converse relation of R (*31.01); we may use R^{-1} for $\breve{\text{R}}$. $\breve{\text{R}}$ “α (*37) signifies the image of α under R, namely, {y | ∃x(x ∈ α ∧ xRy)}, which we denote by R(α) and call it the ‘image of α under R’. Note that R is not required to be 1–1 but for simplicity we assume this condition to hold. Note that the third conjunct of the hypothesis is in fact a definition rather than a condition: the definition of κ. We use the notation {α|. . .} instead of WR’s $\hat{\alpha}$(. . .). p‘κ is the intersection of the members of κ which we denote by ∩κ. Note that ∩κ is in Dedekind’s terminology the chain of (β-Cl‘R) (1963 #44).

Note that in the definition of κ, the first conjunct is only necessary to warrant that κ is a set by the subsets axiom, because it is a subset of D‘R. Therefore, while it was essential to Zermelo to state it, it is unnatural in the context of PM where the use of definitions under hypothesis that constrain variables, is avoided (Grattan-Guinness 1977, p. 30f and PM p. 45). Erasing this unnecessary clause does not turn WR’s proof to hold for any class because classes cannot be elements of classes and so the members of κ cannot be classes that are not sets (see footnote to ‘existence’ in Sect. 24.2.1). Omission of the first conjunct of κ does not effect the proof of *73.8: The first conjunct of the conclusion is obvious by the first two conjuncts of the hypothesis and the second conjunct of the conclusion follows from the first and the nature of the intersection that it is a subset of each member of the intersected class (*40.12). The proof of *73.8 is intuitive so we omit review of the lemmas which WR reference in order to establish it.

The other main lemmas to the proof of *73.85 are (in our notation):

*73.81 ⊢ Hp *73.8 → ∩κ ∈ κ

*73.83 ⊢ Hp *73.8 → ∩κ-(β-Cl‘R) = R(∩κ) ∧ ∩κ = (β-Cl‘R)∪R(∩κ)

*73.84 ⊢ Hp *73.8 → β = ∩κ∪(Cl‘R-R(∩κ))

‘Hp *73.8’ stands for the hypothesis of *73.8 (see comment to *73.801). The proof of *73.81 runs as follows: It is given by *73.8 that ∩κ ⊆ D‘R and as for every α ∈ κ, (β-Cl‘R) ⊆ α, also (β-Cl‘R) ⊆ ∩κ (*73.801); now for every α, R(α) ⊆ α so every member of ∩κ has its relata under R in every α and so it too belongs to ∩κ (*73.802); hence all the conditions for belonging to κ apply for ∩κ. It is because of *73.81 that the definition of ∩κ is impredicative.

With regard to *73.83, note that it is redundant to state the two conjuncts in the conclusion because they are equivalent (*25.47). The proof, however, does make use of the two conjuncts and of the equality of classes by extensionality (*20.43): it is proved that R(∩κ) ⊆ ∩κ-(β-Cl‘R); then that ∩κ ⊆ (β-Cl‘R)∪R(∩κ). From the two results the two conjuncts of the conclusion follow: because (*β*-Ɑ‘R) and R(∩κ) are obviously disjoint, (*β*-Ɑ‘R) can be moved between the two sides of the obtained subsumption equations to give the two conjuncts of the conclusion of the theorem. The first part of the proof follows from the proof of *73.81 where it was established that R(∩κ) ⊆ ∩κ. As (*β*-Ɑ‘R) is disjoint from D‘R it is also disjoint from R(∩κ) and hence R(∩κ) ⊆ ∩κ-(*β*-Ɑ‘R) (*73.811). Actually from *73.81 it could be deduced that (*β*-Ɑ‘R)∪R(∩κ) ⊆ ∩κ making the second conjunct of *73.83 indeed unnecessary.

Anyway, for the second part it is assumed that *x* is not a member of (*β*-Ɑ‘R)∪R(∩κ). Again, since R(∩κ) ⊆ ∩κ and *x* is not in R(∩κ), it follows that R(∩κ) ⊆ ∩κ-{*x*}.[5] Since obviously R(∩κ-{*x*}) ⊆ R(∩κ), R(∩κ-{*x*}) ⊆ ∩κ-{*x*} (*73.812). By another observation made in the proof of *73.81 and the assumption on *x*, (*β*-Ɑ‘R) ⊆ ∩κ-{*x*} ((1) of *73.82). By yet a third observation from the proof of *73.81 ((2) of *73.82), as ∩κ ⊆ D‘R, also ∩κ-{*x*} ⊆ D‘R.[6] These three results entail that ∩κ-{*x*} ∈ κ and hence (*40.12) ∩κ ⊆ ∩κ-{*x*} so that *x* is not in ∩κ (*73.82). By *2.17 (⊢ (~*p* → ~*q*) → (*q* → *p*)) it follows that for every member *x* of ∩κ, *x* belongs to (*β*-Ɑ‘R)∪R(∩κ) so ∩κ ⊆ (*β*-Ɑ‘R)∪R(∩κ) which provides the second step for *73.83. It seems to us that WR used this odd way of reasoning (by *2.17) to avoid the language pattern of a *reductio* argument used in Zermelo’s proof of 1908b.

The proof of *73.84 is simple (using in the last step *73.83):
β = (*β*-Ɑ‘R)∪Ɑ‘R = (*β*-Ɑ‘R)∪R(∩κ)∪(Ɑ‘R-R(∩κ)) = ∩κ∪(Ɑ‘R-R(∩κ)).

The proof of *73.85, that *β* is similar to Ɑ‘R, now follows easily: take R on ∩κ and the identity on Ɑ‘R-R(∩κ). The definition is valid if the two classes are disjoint (*73.69). Indeed, from *73.83 we learn that ∩κ is composed of two parts: *β*-Ɑ‘R, which is trivially disjoint from Ɑ‘R, so also from Ɑ‘R-R(∩κ), and R(∩κ) which is also trivially disjoint from Ɑ‘R-R(∩κ) (*73.841).

26.3 Without the Reducibility Axiom

Let us explicate now how WR’s doctrine of types, the basic gestalt of PM. According to the doctrine of types classes are composed of elements (terms) of the same type. Thus the elements of each of the following classes of *73.8 have all the same type: *β*, D‘R, Ɑ‘R, α, R(α). This we know because relations of inclusion hold between these classes, and inclusion implies that the elements of the related classes are of the same type. Let us denote this type by t_0. Now, according to the doctrine of types, also the terms for which it is false to say that they belong to one of these classes are of type t_0. These two populations of terms together form the

[5] We denote by {*x*} what in PM is denoted by *ι*‘*x*.

[6] The third observation is redundant if the first conjunct of κ is omitted.

type t_0, which is the range of significance for the propositional function $x \in c$, where c is any of the above classes.

Types get a numerical expression, which is their order that represents the way they are constructed. With regard to the types of the classes mentioned above, they are of the order next above the order of their elements (*63). Though in general relations may have domain and range with elements of different types (vol. II, p. 5), in the situation of *73.85, the elements of the domain and range of R are of the same type. The order of the type of κ is, however, the next above the order of the type of α and the other above mentioned classes. The order of the type of ∩κ is the same as the order of the type of κ because it is defined by the propositional function $\cap\kappa = \{x \mid \forall \alpha(\alpha \in \kappa \rightarrow x \in \alpha)\}$ (*40.01, our notation) which contains a bounded variable α which ranges over the type of the elements of κ (cf. Sect. 26.6 below). However, by the axiom of reducibility, which was part of the doctrine in the first edition of PM, it can be assumed that a class exists with the same extension as ∩κ with type of order the next above the type of its member, namely, of the order of the members of κ, to which it thus belongs. Finally, for x in the range of significance of $x \in (\beta\text{-Ɑ‘R})\cup\text{R}(\cap\kappa)$, x is of type t_0 and $\{x\}$ is of the type of order the next above. Thus $\cap\kappa\text{-}\{x\}$, by the axiom of reducibility, can be assumed to be of the type of the members of κ and the proof of the second part of *73.83 holds.

However, in the second edition of PM, WR dropped the axiom of reducibility (vol. I, p. xiv) and so the assumption that $\cap\kappa\text{-}\{x\} \in \kappa$ could not be maintained[7] and the lemma $\cap\kappa \subseteq \cap\kappa\text{-}\{x\}$ could not be applied. So, to save *73.85, Russell had to offer a new proof for this lemma: Take any $x \in \cap\kappa$; then for every $\alpha \in \kappa$, $\alpha\text{-}\{x\}$ does not belong to κ because otherwise x would not belong to ∩κ; so one of the conditions that $\alpha\text{-}\{x\}$ has to fulfill in order to belong to κ fails. Since obviously $\alpha\text{-}\{x\} \subseteq \text{D‘R}$,[8] because $\alpha \subseteq \text{D‘R}$, we have either that (β-Ɑ‘R) is not contained in $\alpha\text{-}\{x\}$, in which case $x \in (\beta\text{-Ɑ‘R})$, or $\text{R}(\alpha\text{-}\{x\})$ is not contained in $\alpha\text{-}\{x\}$, in which case x belongs to R(α) for every α. Now, by *72.34, which basically says that $\cap\text{R}(\kappa) = \text{R}(\cap\kappa)$ (this is a crucial metaphor; it will be proven in the next section), where $\text{R}(\kappa) = \{\text{R}(\alpha) \mid \alpha \in \kappa\}$ by definition (*37.04, using different notation), $x \in \text{R}(\cap\kappa)$ so the second step of *73.83 is proved.

The use of $\alpha\text{-}\{x\}$ in the new proof requires the following observation made by Russell:

> We assume that α-{x} is of no higher order than α; this can be secured by taking α to be of at least the second order, since {x}, and therefore -{x} [the complement of {x} in the type of y ∈ {x}], is of the second order. We may always assume our classes raised to a given order[9] but not raised indefinitely.

---

[7] Because $\cap\kappa\text{-}\{x\}$ still fulfills the three conditions set for belonging to κ, WR say that "in a limited sense" $\cap\,\kappa\text{-}\{x\} \in \kappa$ (PM vol I p xl, xli). This is an example of concept stretching discussed by Lakatos (1976).

[8] There is no problem in stating $\alpha \subseteq \beta$ when α and β are of different types, as long as their members are of the same type. Obviously this lemma is unnecessary if our omission assumption re the definition of κ is maintained.

[9] This is a technical metaphor peculiar to the theory of types.

This remark seems to reflect the possibility that α be a class not defined by a propositional function but pre-given among the individuals of the language (cf. Urquart 2003, p. 287). In this case the class defined by $x \in \alpha$ is of a higher type. In general, if $\varphi(x)$ defines a class, then $\varphi(x) \wedge x \in \{x\}$ gives a class of type of at least the second order. Having closed this loophole WR exclaim: “Thus the Schröder-Bernstein theorem survives.”

26.4 The Inductive Proof

The inductive proof is composed of several steps scattered in *93, *94, *95 of section E. The theorem itself, though it is referred-to often in section E (as the Schröder-Bernstein Theorem), appears only after all its lemmas are proven, at the very end of *95. As it is not numbered we mark it by (CBT-I):

(CBT-I) R,S ∈ 1→1 ∧ Ɑ‘R ⊆ D‘S ∧ Ɑ‘S ⊆ D‘R → C‘(R|S) sm C‘(S|R)

For a relation P, C‘P is the field of P, C‘P = D‘P∪Ɑ‘P (*33.03). The stroke operator “|” between relations, to be distinguished from the stroke operator between proposition (due to Sheffer, second edition vol. I, p. xiii), is an operation sign used by WR for the relative product of relations: If R, S are two relations, their relative product R|S is the relation which holds between x, z, when there is a y such that xRy and ySz (*34). The y must belong to Ɑ‘R∩D‘S and thus if the application of | does not produce the empty relation it is implied that the range of R and domain of S have their members taken from the same type. The relative product is associative (*34.21). The domain and range of R|S are subclasses, possibly proper subclasses, of the domain of R and the range of S, respectively. Under the hypothesis of (CBT-I), it is pointed out next to the theorem, C‘(R|S) = D‘R and C‘(S|R) = D‘S. This remark gives a more familiar form to the theorem – that of the two-set formulation.

The proof of (CBT-I) rests on the gestalt that the field of a reflection (vol. I, p. 589) is naturally partitioned into two classes. This gestalt applies here because R| S and S|R are reflections. The metaphor of the proof is that the corresponding partitions of the fields of R|S and S|R are similar.

For the definition of the two partitions of the field of a reflection relation R,[10] WR introduce two notions. A hereditary class, with respect to R (vol. I, p. 544, comment to *90.01), is any class μ for which R(μ) ⊆ μ. Thus μ contains with every element also its successor (descendent) with respect to R. In *73.8 all the members of κ are hereditary classes. For a relation R the ‘ancestral relation’, denoted R_*, is defined (vol. I, p. 544, *90.01) by: $aR_*z \equiv a \in$ C‘R $\wedge\, \forall\mu(a \in \mu \wedge R(\mu) \subseteq \mu \rightarrow z \in \mu)$. Thus aR_*z when z belongs to every hereditary class to which a belongs. When aR_*z, a is said to be an ancestor of z and z a descendent of a (with respect to R). R_* is reflexive (vol. I, p. 544, *90.12) and transitive but generally not 1–1 even if R is. An alternative formulation of the ancestral relation uses hereditary classes with respect

[10] We assume that the relations mentioned in this section are 1–1 reflections, though WR prove some of the lemmas leading to (CBT-I) for more general relations.

to R^{-1}, it is: $aR_*z \equiv a \in C\text{‘}R \wedge \forall\mu(z \in \mu \wedge R^{-1}(\mu) \subseteq \mu \rightarrow a \in \mu)$ (*90.11). The importance of the ancestral relation is that it permits proof by inductive reasoning: to prove that all the ancestors (descendents) of an element *z* with respect to an ancestral relation R_* have a certain property, it is sufficient to prove that *z* has the property and that if an ancestor (descendent) of *z* has the property also its ancestor (descendent) has that property. For then, the class K of all members from the class of ancestors (descendents) that have the property is hereditary and all ancestors (descendents) of *z* must therefore belong to it (*90.112).[11]

Hereditary classes are what Dedekind called ‘chains’ (1963 #37). R_*(D‘R-Ɑ‘R) is the chain of D‘R-Ɑ‘R. Again, WR do not link their work to Dedekind’s chain theory. They do attribute to Frege (1879, p. 57) the notion of ancestral relation (vol. I, p. 548 footnote; cf. I. Grattan-Guinness 1977, p. 64f, 120 footnote 3, 181 6). Presumably WR regarded Dedekind as a later source but we now know that Dedekind had obtained his theory before 1878, as it appears in Dedekind’s draft of 1872–78 (Dugac 1976).

The first of the two partitions mentioned above is R_*(D‘R-Ɑ‘R), the class D‘R-Ɑ‘R is called by WR the beginning of R, and the second is the residue left of D‘R after removal of the first partition. To obtain the structure of the first partition, WR take up the following considerations: Let |R (*43) be the relation of P|R, for any relation P, to P. If μ is hereditary with respect to |R[12] then for every S $\in \mu$, S|R must also belong to μ (here the alternative formulation of the ancestral relation is applied). Thus every hereditary class that contains R also contains R^2 (R|R), R^3 (R|R^2), etc. So R is the descendent of every relation of the form R^ν, ν a finite number and R is the ancestor of *I*, the identity relation in C‘R, which can be regarded as R^0 (vol. I, p. 545). WR denote by Potid‘R[13] the class of all ancestors of *I*, which includes *I*, (vol. I, p. 545, *91.04, *91.13, *90.12).[14] Now the class of the ranges of all relations in Potid‘R is denoted (vol. I, p. 545, *93) by Ɑ“Potid‘R (*30, *37) and the class of the differences between two consecutive domains is denoted by $\overrightarrow{\min_R}$“Ɑ”Potid‘R and abridged to gen‘R – the generations of R (*93.02, *93.03). It is the class of all images of the beginning of R by the members of Potid‘R; so its members are disjoint.

To prove (CBT-I) the plan is first to prove that the generations of R|S and S|R are similar and the first step for that proof is to prove the similarity of the beginnings of R|S and S|R: D‘(R|S)-Ɑ‘(R|S) sm D‘(S|R)-Ɑ‘(S|R) (*95.7). For the proof note that (*24.412, which is intuitive):

D‘(R|S)-Ɑ‘(R|S) = (D‘R-Ɑ‘S)∪(Ɑ‘S-S(Ɑ‘R)) and
D‘(S|R)-Ɑ‘(S|R) = (D‘S-Ɑ‘R)∪(Ɑ‘R-R(Ɑ‘S)).

---

[11] As Couturat claimed against Poincaré, the induction principle is here defined rather than intuitively assumed or postulated.

[12] The ideas of hereditary set and ancestral relation are applied to a relation of relations.

[13] WR also use Pot‘R which is the same as Potid‘R without *I*. We find this extra symbol unnecessary for our interest.

[14] $I \in C\text{‘}(|R)$ because $I = R^{-1}|R$.

Now, because R(D‘R-ɑ‘S) = ɑ‘R-R(ɑ‘S) and S(D‘S-ɑ‘R) = ɑ‘S-S(ɑ‘R) (*71.38, intuitive), D‘R-ɑ‘S sm ɑ‘R-R(ɑ‘S) and ɑ‘S-S(ɑ‘R) sm D‘S-ɑ‘R. Hence (*73.71) the required result. Let us denote by Φ the relation that provides the similarity of the beginnings of R|S and S|R.

The proof idea for the second step of the proof that the generations of R|S and S|R are similar is to match the corresponding generations of R|S and S|R. This can be achieved, without any reference to numbers, by the following definitions: Let P||Q (*43) denote the relation between P|R|Q and R, for any R, and let (P_*Q)‘R (summary of *95) be the class of all relations that are ancestors of R in the ancestral relation $(P||Q)_*$.[15] It includes R (*95.13). In our context we are looking at $((R|S)^{-1}{}_*(S|R))$‘Φ. The relations in this class are between members of corresponding generations of D‘R and D‘S and thus their union is the required relation between the first partitions of these two domains (summary to *95, *95.52, *95.61).

With regard to the second partition, WR first prove that it is equal to ∩ɑ“Potid‘R for any reflection R (*93.27 and its corollaries). Let *x* be a member of the domain of R not in ∪gen‘R and K the class of all relations from Potid‘R which contain *x* in their range. *I* is in K and if a relation T is in K then T|R is also in K because otherwise *x* would belong to the generation corresponding to T contrary to the assumption that it is not in any generation. Thus K must contain all relations from Potid‘R. This proof is an example of the use of an inductive argument.

It can now be proved that ∩ɑ“Potid‘(R|S) sm ∩ɑ“Potid‘(S|R), either by R or by S^{-1} (or their reciprocals), which settles (CBT-I) (*94.53, *94.54). For example, consider R (∩ɑ“Potid‘(R|S)) (*94.51). The first step is to switch between R() and ∩. A lemma is necessary: *72.34 ⊦ R(∩κ) = ∩R(κ). To prove the lemma consider $y \in$ R(∩κ); then there is some $x \in \cap\kappa$ such that *y*R*x*. Then for every $\beta \in \kappa$, $x \in \beta$ and thus $y \in$ R(β) for every $\beta \in \kappa$, so $y \in$ ∩R(β) = ∩R(κ) and R(∩κ) ⊆ ∩R(κ). The proof can be reversed, bearing in mind that R is 1–1, so that *72.34 is obtained.[16]

With *72.34, R(∩ɑ“Potid‘(R|S)) is transferred to ∩R(ɑ“Potid‘(R|S)), so it is desired to interchange R and ɑ“. This is achieved by Lemma
*43.411 ⊦ R(ɑ“λ) = ɑ“|R“λ. The symbol |R“λ, where λ is a class of relations, signifies the class of relations P|R where P ∈ λ (summary of *43). Similarly, R|“λ signifies the class of relations R|P where P ∈ λ. To prove *43.411 we first use the definition of R(κ) to obtain R(ɑ“λ) = {R(ɑ‘P)| P ∈ λ}. By Lemma
*43.401 ⊦ R(ɑ‘P) = ɑ‘P|R we get R(ɑ“λ) = {ɑ‘(P|R)| P ∈ λ} which is denoted by ɑ“{(P|R)| P ∈ λ} = ɑ“|R“λ. The proof of *43.411 thus turns out to be only a play of substitutions of notations, except for *43.401 which is intuitive from the notion of range.

---

[15] Again we note that the metaphor of the proof is to use the ideas of hereditary class and ancestral relation on a relation of relations.

[16] We short-cut PM's proof that uses a similar argument through 9 auxiliary lemmas. We have also by-passed *94.42.

Thus far we obtained *94.442 ⊦ R(∩Ɑ“Potid‘(R|S)) = ∩Ɑ“|R“Potid‘(R|S). Now we wish to get from a statement on R|S to a statement on S|R. This is Lemma *94.14 ⊦ |R“Potid‘(R|S) = R|“Potid‘(S|R). This lemma follows easily from Lemmas *94.12 ⊦ P ∈ Potid‘(R|S) → ∃T(T ∈ Potid‘(S|R) ∧ P|R = R|T) and its “conjugate” *94.13 ⊦ T ∈ Potid‘(S|R) → ∃P(P ∈ Potid‘(R|S) ∧ P|R = R|T). The proof of *94.12 (and likewise that of *94.13) is by an inductive argument. Let K be the class of P ∈ Potid‘(R|S) for which the lemma holds. *I* ∈ K because R|*I* = *I*|R. If for P ∈ Potid‘(R|S) there is a T ∈ Potid‘(S|R) such that P|R = R|T, then P|R|S|R = R|T|S|R so P|R|S ∈ K. Hence for every P ∈ Potid‘(R|S), P ∈ K and *94.12 is proved. Thus we have finally proved *94.42:
⊦ R(∩Ɑ“Potid‘(R|S)) = Ɑ“R|“Potid‘(S|R).

To complete the proof of *94.51, we need to prove *94.5:
⊦ ∩Ɑ“Potid‘(S|R) = ∩Ɑ“R|“Potid‘(S|R). It follows by extensionality from ∩Ɑ“S|“R|“Potid‘(S|R) ⊆ ∩Ɑ“R|“Potid‘(S|R) ⊆ ∩Ɑ“Potid‘(S|R), which rests on *94.402 ⊦ ∩Ɑ“R|“λ ⊆ ∩Ɑ“λ, and by *94.401:
⊦ ∩Ɑ“Potid‘(S|R) = ∩Ɑ“S|“R|“Potid‘(S|R).[17]

For *94.402 observe that Ɑ‘(R|P) ⊆ Ɑ‘P (*34.36) so that if y ∈ ∩Ɑ“R|“λ then y belongs to every Ɑ‘(P|R), $P \in \lambda$, and thus to every Ɑ‘P so that it belongs to ∩Ɑ“λ.

For *94.401 note that for every T ∈ Potid‘P, T|P ∈ Potid‘P (*93.431), so |(S|R)“Potid‘(S|R) ⊆ Potid‘(S|R) and ∩Ɑ“Potid‘(S|R) = ∩Ɑ“|(S|R)“Potid‘(S|R). By *91.3 ⊦ P ∈ Potid‘R → R|P = P|R, |(S|R) can be replaced by (S|R)| to obtain:

(S|R)|“Potid‘(S|R) = {S|R|P | P ∈ Potid‘(S|R)} = S|“{R|P | P ∈ Potid‘(S|R)}= S|“R|“Potid‘(S|R) (*43.2).

*91.3 is proved by an inductive argument: Let K be the class of all members P of Potid‘R for which R|P = P|R. Then clearly *I* is in K and if P ∈ K
R|(P|R) = R|P|R = (R|P)|R = (P|R)|R so P|R ∈ K and thus for all P ∈ Potid‘R, P ∈ K.

This completes the proof of *94.51 and thus of (CBT-I). Dozens of more lemmas are necessary for the completely formal proof of PM, which we avoided using intuitive arguments.

26.5 The Drawings

In the introduction to *94 WR gave two drawings: one for the impredicative proof and another for the inductive proof. In the voluminous PM treatise, drawings are rare and only CBT received such extended heuristic attention.[18] The reason could be that CBT was the laboratory where much of PM’s technique was developed and exemplified.

The drawing for the impredicative proof is the following (Fig. 26.1):

[17] In PM, *94.401 is given with the roles of R and S interchanged.

[18] No doubt the use of drawings by Russell goes against his point of departure on his logicism trip: to prove that Kant was wrong in stating that drawings are essential for every geometric proof (Bechler 1999 p 219).

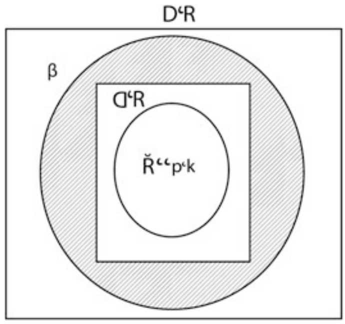


Fig. 26.1 PM’s drawing for the impredicative proof

The inner oval, denoted by R“p‘k which we denote by R(κ), is the image-chain of the shaded frame of β. Unfortunately the drawing does not suggest the structure of the image-chain.

The drawing for the inductive proof is the following (Fig. 26.2):

WR depict the original domains by rectangles, their ranges by ovals and their ranges by rectangles again, which makes it difficult to grasp immediately which figures correspond. The drawing exhibits the correspondence of points from the first frame of each domain and thus the similarity of the beginnings, but following the image of these points one step further makes the drawing lose its clarity. The similarity of the generations does not come out of the drawing. They are represented by areas between consecutive rectangles, which are the ranges of the members of Potid‘(R|S) and Potid‘(S|R). The drawing ignores the residues obtained from each

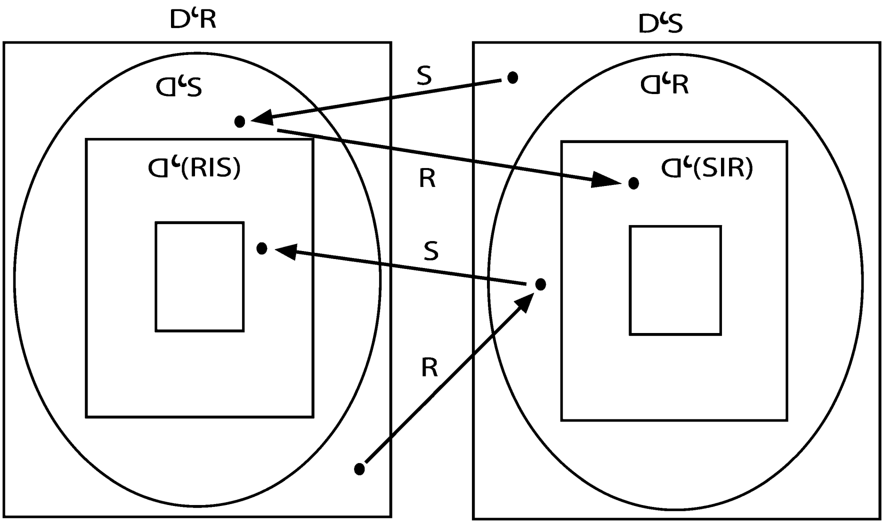


Fig. 26.2 PM’s drawing for the inductive proof

domain after removal of the frames. However, in this WR are not alone: neither are the residues represented in Borel's drawing. Residues are represented in Schröder's (see Sect. 10.1) drawing, but this was an unfortunate heuristic as it led him to the errors in his proof of CBT (see Sect. 10.2). A good representation of the residue appears in Fraenkel's drawing (see Sect. 19.2), but there the nesting of the images are not clear at first sight.

26.6 The Cardinal Version

In *117, titled "Greater and less", WR presented the fourth, cardinal number, version of CBT:
*117.23 $\vdash$ Nc'$\alpha \geq$ Nc'$\beta \wedge$ Nc'$\beta \geq$ Nc'$\alpha \equiv$ Nc'$\alpha =$ Nc'β' $\equiv$ indicates the equivalence connective between propositions ($\leftrightarrow$). Nc'α, read "the cardinal number of class α", is defined as the class of all classes that are similar to α. Informally the definition is given in PM vol. II, p. 4 "summary of section A", where it is recalled that the definition is due to Frege, and formally in *100. The definitions for the equality and inequality relations between cardinal numbers will be presented below. Here we note that in the summary to *117 WR pointed out that $>$ is transitive and antisymmetrical and that the comparability of any two cardinal numbers can be proved only by way of the Well-Ordering Theorem which assumes the multiplicative axiom, namely, the axiom of choice.

The title of *117 is reminiscent of the title of §2 of Cantor's 1895 *Beiträge*: "'Greater' and 'less' with powers". The association is not superficial for in the 1895 *Beiträge* Cantor also defined the $>$ relation, proved its transitivity and antisymmetry, and stated CBT; but there are differences between the two texts:

- Cantor never defined the relation $\geq$ between cardinal numbers or powers. In fact, as Jourdain points out, the definition of $\geq$ between cardinal numbers assumes CBT (see Sect. 17.2).
- Cantor never presented CBT for cardinal numbers. He did state the Comparability Theorem for cardinal numbers but CBT he stated only in the language of sets and mappings. The first to state CBT for cardinal numbers was Zermelo (1901).
- Cantor never presented in full a direct proof of CBT. He hinted at one in his *Grundlagen* (see Chap. 1) and presented CBT as corollary of the Comparability Theorem in the 1895 *Beiträge*.
- Though Cantor had a definition of the cardinal number of a set – by abstraction (1895 *Beiträge* §1) – he did not use this definition in his definition of the order relation $>$ between cardinal numbers. WR did use their definition of cardinal numbers in their definition of $>$ as will be pointed out below.

On *117.23 WR say in the summary to *117 (vol. II, p. 166):

> This proposition may be called the Schröder-Bernstein theorem with as much propriety as 73.88 [the two-set version]; the two are scarcely different.

This surely is a pun in understatement, for the propositions *117.23 and *73.88 appear very different. The statement must, therefore, be interpreted either as based on a proof that the two statements are equivalent, which WR do not provide but which can be supplied, as we will indicate, or as relating to the informal meaning of the two propositions, the construction of which is never explicated by WR. The second possibility is intriguing for it exemplifies the centrality of (informal) meaning in the context of formal reasoning. This issue is connected, so we believe, with the essentialist nature of mathematical thinking.

To prove that *117.23 is entailed by *73.88, WR restate the latter as follows: *117.2 $\vdash (\alpha \text{ sm } \alpha' \wedge \beta \text{ sm } \beta' \wedge \alpha' \subseteq \beta \wedge \beta' \subseteq \alpha) \rightarrow \alpha \text{ sm } \beta$.

α, β, α', β' are used instead of α, β, γ, δ of *73.88; WR use the convention to name the image by the same letter as the source. Incidentally, in *117.2 the implication can trivially be upgraded to equivalence so that it appears more analogous to *117.23.

The proof of *117.23 is established by two lemmas:

*117.211 $\vdash \exists!(\text{Cl}`\alpha \cap \text{Nc}`\beta) \wedge \exists!(\text{Cl}`\beta \cap \text{Nc}`\alpha) \equiv \text{Nc}`\alpha = \text{Nc}`\beta$
*117.22 $\vdash \exists!(\text{Cl}`\alpha \cap \text{Nc}`\beta) \equiv \text{Nc}`\alpha \geq \text{Nc}`\beta$

The sign Cl‘α signifies the class of all subclasses of α. It is introduced informally in the summary to section B, PM vol. I, p. 386. It is not introduced formally because formally it is combined of two signs: Cl (*60.01), which is a relation that relates to every class the class of its subclasses, and the operator ‘ which turns a relation R into a descriptive function R‘ (*30.01) “‘the R of’ which, when applied to some relata *x*, as in R‘*x*, reads ‘the term [referent] which stands in the relation R to *x*’” (summary to *30). Obviously, the ‘ operator can be applied only to a relation which is one-one or one-many. The sign ∃! (*24.03) signifies that the class on its right (the brackets are not in PM) is not empty. *117.23 is obtained from *117.211 by substitution from *117.22.

The proofs of the lemmas we will again roll backwards up to the point where we feel that what we leave out is intuitively clear. The left to right implication of the equivalence in *117.211 follows from:

*117.21 $\vdash \exists!(\text{Cl}`\alpha \cap \text{Nc}`\beta) \wedge \exists!(\text{Cl}`\beta \cap \text{Nc}`\alpha) \rightarrow \text{Nc}`\alpha = \text{Nc}`\beta$

Which in turn follows from *117.2 (CBT; the derivation of the hypothesis of *117.2 from that of *117.21 is intuitively clear and PM details it no further) and *100.321 $\vdash \alpha \text{ sm } \beta \rightarrow \text{Nc}`\alpha = \text{Nc}`\beta$. The latter lemma, which is the crucial bridge between the language of *117.2 to that of *117.23, is proved by *73.37 $\vdash \alpha \text{ sm } \beta \rightarrow \gamma \text{ sm } \alpha \equiv \gamma \text{ sm } \beta$, which implies the equality of Nc‘α and Nc‘β by extensionality. The proof of *73.37 is straightforward and we will not pursue this branch of our proof any further.

The right to left implication in *117.211 follows from the definition of ∃! and the two lemmas:*60.34 $\vdash \alpha \in \text{Cl}`\alpha$ and *100.3 $\vdash \alpha \in \text{Nc}`\alpha$; which also seem intuitively straightforward.

With regard to *117.22 the proof follows from the following lemmas:
*117.108 $\vdash$ $\text{Nc}`\alpha \geq \text{Nc}`\beta \equiv \text{Nc}`\alpha > \text{Nc}`\beta \vee \text{Nc}`\alpha = \text{Nc}`\beta$
*117.13 $\vdash$ $\text{Nc}`\alpha > \text{Nc}`\beta \equiv \exists!(\text{Cl}`\alpha \cap \text{Nc}`\beta) \wedge \sim\exists!(\text{Cl}`\beta \cap \text{Nc}`\alpha)$
and *117.211. Indeed:

$\exists!(\text{Cl}`\alpha\cap\text{Nc}`\beta) \equiv \exists!(\text{Cl}`\alpha\cap\text{Nc}`\beta) \wedge (\exists!(\text{Cl}`\beta\cap\text{Nc}`\alpha) \vee \sim \exists!(\text{Cl}`\beta\cap\text{Nc}`\alpha))$[19] $\equiv$ $(\exists!(\text{Cl}`\alpha\cap\text{Nc}`\beta) \wedge \exists!(\text{Cl}`\beta\cap\text{Nc}`\alpha)) \vee (\exists!(\text{Cl}`\alpha\cap\text{Nc}`\beta) \wedge \sim \exists!(\text{Cl}`\beta\cap\text{Nc}`\alpha)) \equiv$ $\text{Nc}`\alpha = \text{Nc}`\beta \vee \text{Nc}`\alpha > \text{Nc}`\beta$.

At this point our journey we need to come back to the notion of relative types (cf. Grattan and Guinness 2000 Sect. 7.9). Recall that a class, say α, as introduced in *20 (cf. pp. 23, 47, 199), consists of all the terms that satisfy a propositional function (p. 38), namely, that turn it into a true proposition. The class of all terms that satisfy the negation of the same propositional function is denoted by -α (*20.06, *22.04). The union of α and -α is the class of the terms for which the propositional function is significant. Such a class is called the 'type of the members of α' and is denoted by $t_0`\alpha$ (cf. *63). $\alpha$ is contained in $t_0`\alpha$. Now, for every term x, the class $\{x\}$ consists of all the terms which satisfy the propositional function $y \in \{x\}$. -$\{x\}$ consists of all terms y satisfying $\sim y \in \{x\}$. Necessarily, these are such y's for which $y \in \{x\}$ is significant. Thus, if the propositional function 'x was a teacher of Plato' is denoted by $\varphi(x)$, then Socrates satisfies $\varphi(x)$, Pliny falsifies it 'all men are mortal' is insignificant for it. The union of $\{x\}$ and -$\{x\}$ is called the 'type of x' and is denoted by $t`x$. Obviously $x$ belongs to $t`x$ and $t`x = t_0`\{x\}$. WR provide the following theorems:

*63.11 $\vdash$ $x \in t_0`\alpha \rightarrow t`x = t_0`\alpha$;
*63.16 $\vdash$ $x \in t`y \equiv y \in t`x \equiv \exists!(t`x \cap t`y) \equiv t`x = t`y$.

According to the above definitions all members of a class belong to the same type, but then to which type belong the members of Nc'α? In fact, there is no specific type for the members of Nc'α; WR say that Nc'α is typically ambiguous (vol. II, p. 5). The type of the members of Nc'α becomes definite, at least relative to some other term, by the context in which Nc'α is used, especially when it is equated to a term of a definite type. Because of its context dependence Nc'α may in certain cases be empty; for instance, Nc't'α, in the type of α, is empty as a result of Cantor's Theorem (*102.73), because t'α contains all the subsets of α and its cardinal number is not of that type. This is in spite of *100.3 $\vdash$ $\alpha \in \text{Nc}`\alpha$ because here the type of Nc'α is set to the next above that of α and so it is not empty. To avoid the need to pay attention to these cases WR introduced certain conventions (prefatory statement to vol. II) to the effect that whenever it is needed in a context, $\exists!\text{Nc}`\alpha$ is tacitly assumed.

As we shall shortly see, the typical ambiguity of Nc'α is instrumental in certain instances, but at other times there is a need for cardinal numbers of definite types. So WR introduced (vol. II, p. 8 and *103) the notion of 'homogeneous cardinals',

[19] Conjunction with a tautology maintains truth value.

denoted by $N_0c`\alpha$, which is the cardinal number of $\alpha$ the members of which are of the same type as $\alpha$. Thus we have that $N_0c`\alpha = Nc`\alpha \cap t`\alpha$ and $N_0c`\alpha$ is not empty by *100.3 $\vdash \alpha \in Nc`\alpha$.

Now the definition of the relation $>$ between cardinal numbers is enabled:
*117.01 $\mu > \nu = (\exists\alpha,\beta)(\mu = N_0c`\alpha \wedge \nu = N_0c`\beta \wedge \exists!(Cl`\alpha \cap Nc`\beta) \wedge {\sim}\exists!(Cl`\beta \cap Nc`\alpha))$.

This definition, because of its use of Nc‘, allows for the inequality of cardinals of different types. Replacing μ, v in *117.01 by their N_0c‘ we get:
$N_0c`\alpha > N_0c`\beta \equiv \exists!(Cl`\alpha \cap Nc`\beta) \wedge {\sim}\exists!(Cl`\beta \cap Nc`\alpha))$ and as WR define (*117.02, *117.03) $Nc`\alpha > Nc`\beta$ to mean $N_0c`\alpha > N_0c`\beta$, *117.13 is obtained. (In 117.12 it is proved that the relation $N_0c`\alpha > N_0c`\beta$ is independent on which α, β are taken from the classes $N_0c`\alpha$, $N_0c`\beta$. Indeed, $>$ is a relation between classes and not between any notational representations of them.)

With regard to $\geq$ and *117.108, WR note (vol. II, p. 165) that it cannot be defined by $\mu \geq v \equiv \mu > v \ \vee \ \mu = v$ because:

> $\mu = \nu$ restricts μ, ν too much by requiring that they should be of the same type [otherwise equating them is insignificant] and restricts them too little by not requiring that they should both be existent [not empty] cardinals.

So instead WR define $\geq$ by:
*117.05 $\mu \geq v = \mu > v \ \vee \ (\mu, v \in N_0C \ \wedge \ \mu = sm``v)$. N_0C is the class of all homogeneous cardinals (*103.02); sm“v (*37) is the class of all classes similar to any member of v, which is, by convention, a cardinal number. sm“v is typically ambiguous (*100) but when equated to μ, a cardinal number, it assumes the type of μ to the effect that it becomes the class of all classes of $t_0`\mu$ that are similar to members of v. As all the members of ν are equivalent, so are all the members of sm“v and when the latter is equated to μ it means that μ and v are the same cardinal in their respective types. Replacing μ, v in *117.05 by their N_0c‘ expressions we get:

$N_0c`\alpha \geq N_0c`\beta \equiv N_0c`\alpha > N_0c`\beta \vee N_0c`\alpha = sm``N_0c`\beta$. *117.108 is now easily obtained through substitutions in this equivalence using the definition of $>$ between Nc‘ (*117.02, *117.03) and: *103.4 $\vdash sm``N_0c`\alpha = Nc`\alpha$, with *103.16 $\vdash N_0c`\alpha = Nc`\beta \equiv Nc`\alpha = Nc`\beta$.

Finally, *103.4 holds because the two equated expressions say the same thing for any specific type. *103.16 is based on $\alpha \in Nc`\beta$ which holds under the premise of either side of the equivalence.

Note that from *117.23, by a reverse substitution using *117.22, *117.21 can be obtained and from it, with the inverse of *100.321, we can obtain back *117.2. Thus the equivalence of the impredicative and cardinal numbers versions of CBT can be established, as anticipated following the remark of WR (vol. II, p. 166) noted above.

26.7 Comparisons with Earlier Proofs

Prior to PM, impredicative proofs were given by Dedekind (Sect. 9.2), Zermelo (Sects. 19.5 and 24.1) and Peano (Sect. 20.2). All these proofs share the following scheme:

- They are all for CBT in the single-set formulation, say $M' \subseteq M_1 \subseteq M \sim_{\varphi} M'$.
- The given set M is partitioned into two partitions (gestalt).
  - The first partition is the intersection Q_0 of all chains in M that contain Q where Q is either M-M_1 (Dedekind, Peano) or M_1-M′ (Zermelo, PM).
  - The second partition is the complement of the first in M.
  - Combining φ on the first partition (pushdown the chain metaphor) and the identity on the second gives a mapping of M on M_1 (Dedekind, Peano) or M_1 on M′ (Zermelo, PM).
  - All proofs claim that $Q_0 = Q + Q'_0$ which they prove using extensionality.
  - They all first prove that $Q + Q'_0 \subseteq Q_0$: $Q \subseteq Q_0$ by the definition of Q_0 and $Q'_0 \subseteq Q_0$ because Q_0 is a chain.
  - Then Dedekind (#58) and Peano prove that $Q + Q'_0$ is a chain and by the minimality of Q_0, $Q_0 \subseteq Q + Q'_0$.[20] The proof that $Q + Q'_0$ is a chain runs as follows: From the previous bullet $Q + Q'_0 \subseteq Q_0$ so, as φ is 1–1, $(Q + Q'_0)' \subseteq Q'_0$. But obviously, $Q'_0 \subseteq Q + Q'_0$ so $(Q + Q'_0)' \subseteq Q + Q'_0$. So $Q + Q'_0$ is a chain that contains Q. So $Q_0 \subseteq Q + Q'_0$.
  - Zermelo 1908b and PM (with the reducibility axiom) both obtain that $r \in Q_0$ entails $r \in Q + Q'_0$. They both argue that for a certain r, Q_0-$\{r\}$ is a chain that contains Q. But Zermelo takes $r \in Q_0$-Q and not in Q'_0, to obtain a contradiction, while PM take $r \notin Q + Q'_0$ and obtain that $r \notin Q_0$ to obtain the desired result without a contradiction using *2.17.
    - Zermelo: Let $Q_1 = Q_0$-$\{r\}$; then we would still have $Q'_0 \subseteq Q_1$ and since $Q_1 \subseteq Q_0$ we have $Q'_1 \subseteq Q'_0$ so $Q'_1 \subseteq Q_1$ so Q_1 is a chain and it contains Q and we should have $Q_0 \subseteq Q_1$, a contradiction. Hence if $r \in Q_0$-Q, $r \in Q'_0$ so that Q_0-$Q \subseteq Q'_0$ or $Q_0 \subseteq Q + Q'_0$.
    - PM: $Q'_0 \subseteq Q_0$-$\{r\}$ so $(Q_0$-$\{r\})' \subseteq Q'_0 \subseteq Q_0$-$\{r\}$. Since r is not in Q we have that $Q \subseteq Q_0$-$\{r\}$ and thus Q_0-$\{r\}$ is a chain that contains Q and so $Q_0 \subseteq Q_0$-$\{r\}$ and hence r is not in Q_0. So $Q_0 \subseteq Q + Q'_0$.
  - PM (without reducibility) argue directly that if $r \in Q_0$ then $r \in Q + Q'_0$.
    - Take $r \in Q_0$. Then for every chain α that contains Q, α-$\{r\}$ is not a chain that contains Q for otherwise r would not be in Q_0. Hence either Q is not a subset of α-$\{r\}$ in which case $r \in Q$, or $(\alpha$-$\{r\})'$ is not a subset of α-$\{r\}$ in which case $r \in \alpha'$. But then $r \in \cap\alpha'$ for every chain α that contains Q and so $r \in Q'_0[=\cap\alpha']$.

Clearly the proof of Dedekind and Peano is more elegant because it stays in the language of sets and does not switch to calculations involving elements. It is remarkable that Zermelo and WR did not spot it. This proof, however, could not work for PM without reducibility so Russell's changed proof was necessary. That proof appears to be more elegant than the earlier proofs of Zermelo and PM because

[20] Poincaré does not prove that $Q + Q'_0$ is a chain, only states it and its consequence.

it avoids a logical argument and because it reveals an important aspect of the chains not openly stated before: the image of the intersection is the intersection of the images.

Note that the impredicative proof makes no direct use of the frame structure of chains. So perhaps the pushdown the chain metaphor should be restated as ‘collapse of the chain’.

The gestalt of the inductive proof of PM includes ‘beginnings, generations, residues’ and the metaphoric descriptor is ‘roll’em up, roll’em over, roll’em down’. The proof is different from all previous proofs given to that theorem in several ways:

- The construction of the metaphor of two reflections R|S and S|R. Only in those proofs that shift the proof from the two-set formulation to the single-set formulation, a reflection was previously generated, and then for the target set only. But WR’s proof, though it does not glide into a proof for the single-set formulation, does use reflections and for both sets.[21]
- The gestalt of generations of a reflexive relation, with its origin in the beginning, was not noticed in previous proofs that are focused on the frames. Likewise was ignored the construction of the metaphor of the 1–1 relation between the beginnings from the relations given in the conditions of the theorem and the metaphor of obtaining the 1–1 relations between the generations by rolling down the relation between the beginnings using the two reflections R|S and S|R.
- The proof by an inductive argument (metaphor) that the residue is equal to the intersection of the nesting ranges of the relations that generate the generations. Likewise, the proof by two inductive arguments (metaphors) that the residues are equivalent. Usually these steps are overlooked.[22]
- The interlacing of the ranges metaphor that was used in the proof of *94.5, a metaphor that was not observed by previous proofs.

With regard to the cardinal proof, there was only one proof of CBT in the language of cardinal numbers – the proof of Zermelo 1901. However, while Zermelo’s proof was indeed a CBT proof that was leveraged from his Denumerable Addition Theorem applying the reemergence argument, PM’s cardinal proof rested on CBT and only translated it from the language of classes and relations to the typically ambiguous language of cardinal numbers. Thus PM’s cardinal proof of CBT is a product uniquely devised for the context of PM with no general mathematical interest.

Looking back at this chapter one cannot help but wonder how the plain transparent prose of Cantor, Dedekind and Zermelo was transformed into a formidable prayer-book of forms and formulations. Is this turn in the context of thought really unavoidable?

---

[21] It is perhaps because of its use of reflections that WR state the conclusion of (CBT-I) for the fields of R|S and S|R rather than the domains of R and S.

[22] Schoenflies and Poincaré did take notice of them.

Chapter 27
The Origin of Hausdorff Paradox in BDT

We conclude the third part of our excursion with a presentation of Hausdorff's paradox, conjecturing on its origin in BDT.

Hausdorff's presentation of the theorem known as Hausdorff Paradox runs as follows:[1]

> The insolvability of the problem of content. We demonstrate that it is not possible to assign to every point-set of the surface K of a sphere, a content that fulfills the requirements (α), (γ) from p 401[2] and whereby f(K) turns out to be positive. The proof rests on the remarkable fact that half a sphere and a third of a sphere can be congruent, or that K (apart from a denumerable set) can be partitioned into three sets A, B, C which are congruent with each other as well as with B + C.

The paradoxical appearance of the theorem arises from the use of concept stretching (Lakatos 1976) made in it of the notions 'half' and 'third', which inserts the unwarranted presupposition that the whole (K) and the parts (A, B, C) are measurable by a common scale. Because of its paradoxical appearance the theorem became known as Hausdorff Paradox. A decade later it stimulated the Banach-Tarski Paradox.

The opening sentence of the theorem and its link to §X.1 place it in the context of the problem of content.[3] This problem Hausdorff presented in §X.1 (p 401) as follows:[4]

> To assign each (bounded) set A of the E_n space [Euclidean space of n dimensions] as content a number $f(A) \geq 0$ under the following conditions:

[1] We quote from Hausdorff 1914a. There the theorem appears in the appendix (p 469), in an endnote to a passage (p 402) in §1, titled "the problem of definition of content" (*Inhaltsbestimmung*), of Chap. 10, titled "Contents [*Inahlte*] of point sets". The theorem, with some of its background from section X.1, also appeared in *Mathematische Annalen*, perhaps a couple of months before the book (Hausdorff 1914b).

[2] See below.

[3] We maintain Hausdorff's terminology though currently the problem is referred to as the 'measure problem'.

[4] Hausdorff relates the formulation in this presentation to Lebesgue.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_27, © Springer Basel 2013

(α) Congruent sets have the same content.
(β) The content of the unit cube is 1.
(γ) f(A + B) = f(A) + f(B).[5]
(δ) f(A + B + C…) = f(A) + f(B) + f(C) + … for a bounded sum of denumerably many sets.

In §X.1 Hausdorff demonstrated that the problem is insolvable already for E_1 (and hence for all spaces with higher dimension by taking an appropriate cylinder – p 401 footnote 1) by providing a partitioning of the unit segment into denumerably many congruent partitions. If each partition be assigned the same (according to (α)) positive number, obviously the number assigned to the segment, according to (δ) will have to be infinite, contrary to (β).[6]

Then in a short passage (p 402) Hausdorff remarks that it is noteworthy that the problem is insolvable even if one gives up requirement (δ), at least in E_3 (and hence in all higher dimension spaces). For proof of this assertion Hausdorff refers to the appendix. It thus appears quite possible that Hausdorff had found his paradox when his 1914 book was in advanced stages of setting. This observation is perhaps supported by another remark in the appendix (p 451) where it is explained that the promised (p 14) role of "symmetric sets" in the theory of the content of point sets, "relates to an originally planned presentation which is now settled by another, shorter one".

Notwithstanding the context of the problem of content in which Hausdorff Paradox was presented, it is our opinion that the origin of the theorem lies in a certain observation that Hausdorff made about BDT, which he combined with J. Kőnig's string gestalt from his 1906 CBT proof. It is Hausdorff's remarkable achievement that he was able to turn the natural association of these observations with the problem of content into a rigorous theorem.

27.1 The Metaphor

In BDT (see Chap. 14), we are given a set that is divided in two different ways into a finite number of equivalent partitions; it turns out that the partitions of the different partitionings are also equivalent. In the language of cardinal numbers this means that if $2\mathfrak{m} = 2\mathfrak{n}$ then $\mathfrak{m} = \mathfrak{n}$ or if $3\mathfrak{m} = 3\mathfrak{n}$ then $\mathfrak{m} = \mathfrak{n}$. Now the question naturally rises whether when $3\mathfrak{m} = 2\mathfrak{n}$ also $\mathfrak{m} = \mathfrak{n}$. Of course with the axiom of choice the theorem is trivial, because $3\mathfrak{m} = \mathfrak{m}$ and $2\mathfrak{n} = \mathfrak{n}$, but is it provable without the axiom of choice? If Hausdorff asked himself this question he probably obtained the idiom that a third is equivalent to a half, right from the start, perhaps enhancing his interest in the problem because of its possible link to the problem of content (the bright idea – Pólya 1945).

[5] It is tacitly assumed that A, B are disjoint; this point is noted on p 400. For (δ) it surely must be assumed that the sets are pairwise disjoint.

[6] This example resembles the one given by Vitali 1905 (Schoenflies 1900 p 374f).

In the language of sets and mappings, which is the language used in Bernstein's proof, the conjecture is worded as follows: Let S be an infinite set that is partitioned into three equivalent partitions M_1, M_2, M_3 and also into two equivalent partitions N_1, N_2; then $M_1 \sim N_1$. Let ψ be a 1–1 mapping which takes M_1 onto M_2, M_2 onto M_3 and M_3 onto M_1 such that $\psi^3 = 1$ – the identity mapping. Let φ be a 1–1 mapping that takes N_1 onto N_2 and N_2 onto N_1 such that $\varphi^2 = 1$. In Bernstein's proof of BDT the idea is to follow the images of a certain subset under the mappings composed from the given mappings. In the suggested context of $3m = m$ and $2n = n$, these are the mappings
$\varphi, \psi\varphi, \psi^2\varphi, \varphi\psi\varphi, \varphi\psi^2\varphi\ldots \psi, \psi^2, \varphi\psi, \varphi\psi^2\ \psi\varphi\psi, \psi\varphi\psi^2$, etc. These composite mappings can be regarded as words composed of the letters φ and ψ.[7] Following J. Kőnig's gestalt of strings, the following network (see Sect. 22.2.) emerges for an element x of S:
$x, \varphi(x), \psi(x), \psi^2(x), \psi\varphi(x), \psi^2\varphi(x), \varphi\psi(x), \varphi\psi^2(x)$, and so on. It is a network and not a string because from certain words composed of the letters φ, ψ, ψ^2, more than one word can emerge. The network stretches indefinitely if the images of x by these mappings are all different, a property which Hausdorff described by saying that the mappings are independent. The mappings in Bernstein's proof of BDT and J. Kőnig's CBT proof (except perhaps at the residue, see Sect. 21.3.) were indeed independent in this sense.

Now it is a short step to the observation that all the images of x can be partitioned in three partitions: a partition that contains those images that are generated by mappings that begin on the left with φ, denote it by A_x, another, B_x, that contains those images of x generated by mappings that begin on the left with ψ, and a third, C_x, that contains those images of x generated by mappings that begin on the left with ψ^2. It is easy to see that A_x is equivalent to $B_x + C_x$ by φ, that by ψ A_x is equivalent to B_x, B_x to C_x and C_x to A_x, and that finally by ψ^2 A_x is equivalent to C_x, B_x to A_x and C_x to B_x. Notice that the result was obtained without using the axiom of choice.

The above construction does not lead to a proof of the extended BDT, namely that $3\mathfrak{m} = 2\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$, because in the general case it is not possible to assert that the mappings carrying over the partitionings are indeed independent. However, in the special case where S is the surface of a sphere, Hausdorff was able to demonstrate that there are rotations[8] by half turn and third turn of the sphere, which are independent and so the desired result can be obtained.[9] As he himself stated: "this is the heart of the entire consideration". In the next section we bring a translation of Hausdorff's proof of his paradox.

---

[7] Such 'words' appeared already in Dedekind 1855–8, Dedekind 1930–2 vol 3 p 440. See Pla i Carrera 1993 p 220.

[8] Hausdorff was directed towards rotations as the form of congruence by his use of rotations of a circle in his counterexample to the possibility of ω-additive content. He turned to spatial figure no doubt because he needed independent rotations.

[9] Moore (1982 p 187) remarks that Hausdorff, in his proof of this theorem, "extended earlier arguments yielding a non-measurable set", but he does not give any references.

27.2 Proof of Hausdorff's Paradox

We cite the proof from Hausdorff 1914a p 469:

> Now, a denumerable set Q [on K] must anyway have zero content: because choosing a rotation axis which passes through neither point of Q, and a rotation angle which is equal to none of the differences in geographic longitudes of two points of Q,[10] we obtain a rotation that transposes Q into a subset of P = K−Q, and through iterations of this procedure it is recognized that K has arbitrarily many pairwise disjoint subsets which are congruent with Q, so that $f(\mathrm{Q}) \leq 1/\mathrm{n}\, f(\mathrm{K})$[11] for n = 1, 2, 3, ... and thus $f(\mathrm{Q}) = 0$.[12] By the specified partitioning K = Q + A + B + C must *f*(A) be at the same time 1/2 *f*(K) and 1/3 *f*(K), which contradicts the assumption $f(\mathrm{K}) > 0$.
>
> In order to produce such partitioning we denote by φ a half turn (at π) and by ψ a third turn (at 2/3 π) about one of the axes different from the axis of the first [rotation]. They generate a group G, of rotations, the elements of which, arranged by the number of factors (whereby φ, ψ, ψ^2 are counted as simple factors), we write as follows:
>
> (G) 1 | φ, ψ, ψ^2 | $\varphi\psi, \varphi\psi^2, \psi\varphi, \psi^2\varphi$ | . . . . By suitable choice of the rotation axes there exist no relations between φ and ψ except that $\varphi^2 = \psi^3 = 1$. As this is the heart of the entire consideration, we wish to demonstrate it elaborately. The product of two or more factors is of one of the four forms[13]
>
> $$\alpha = \varphi\psi^{\mathrm{m}_1}\varphi\psi^{\mathrm{m}_2}\ldots\varphi\psi^{\mathrm{m}_\mathrm{n}}$$
> $$\beta = \psi^{\mathrm{m}_1}\varphi\psi^{\mathrm{m}_2}\ldots\psi^{\mathrm{m}_\mathrm{n}}\varphi$$
> $$\gamma = \varphi\psi^{\mathrm{m}_1}\varphi\psi^{\mathrm{m}_2}\ldots\varphi\psi^{\mathrm{m}_\mathrm{n}}\varphi$$
> $$\delta = \psi^{\mathrm{m}_1}\varphi\psi^{\mathrm{m}_2}\ldots\varphi\psi^{\mathrm{m}_\mathrm{n}},$$
>
> where n is a natural number and $m_1, m_2, \ldots, m_n$ are equal to either 1 or 2. A relation $\sigma = \rho$ between two formally different products would have as consequence $\rho\sigma^{-1} = 1$, that is, one product formally different from the identity 1 must actually be =1; this product must have at least two factors, and the relation in question can be assumed in the form $\alpha = 1$ (out of $\beta = 1$ follows $\varphi\beta\varphi = \alpha = 1$, out of $\gamma = 1$ likewise $\varphi\gamma\varphi = \delta = 1$,[14] finally, out of $\delta = 1$ $\psi^{-m_1}\delta\psi^{m_1} = \alpha' = 1$ follows).[15] It is therefore to show that by a suitable choice of both rotation axes all products α are different from 1.

[10] The number of these differences is equal to the number of pairs of elements from Q, which is denumerable (see Sect. 1.2.). As the number of axes passing through the center of the sphere is not denumerable, and likewise is the number of rotation angles, the choice of an axis and rotation angle as required is possible.

[11] Namely, f(K)/n.

[12] In the proof that f(Q) = 0, a chain of Q is generated, proof-processing from *Zahlen*.

[13] These are reduced forms, after omitting φ^2 or ψ^3 that may occur when words of different types are combined.

[14] Here Hausdorff reduces the case of γ to the not yet handled case δ.

[15] Hausdorff does not explain what α' stands for. If for the m_1, m_n of δ, $m_1 + m_n \neq 3$, then α' is α; otherwise, α' is γ. This is transformed by $\varphi\gamma\varphi$ to a δ with n-2 φ and n-1 ψ, if originally δ had n ψ and n-1 φ. Repeating this procedure up to n times, we must encounter an α variant for otherwise we reduce the formula to φ which is clearly not 1.

We put through the middle point of the sphere a right angled axes system, and take the ψ-axis in the z-axis, the φ-axis in the xz-plane, denoting the angle between the two by 1/2 θ and setting

$\lambda = \cos 2/3\ \pi = -1/2,\ \mu = \sin 2/3\ \pi = \surd 3/2,$

our rotations become the corresponding orthogonal transformations:

$$(\psi)\ x' = x\lambda - y\mu$$
$$y' = x\mu + y\lambda$$
$$z' = z$$
$$(\varphi)\ x' = -x\cos\theta + z\sin\theta$$
$$y' = -y$$
$$z' = x\sin\theta + z\cos\theta$$
$$(\varphi\psi)^{16}\ x' = -x\lambda\cos\theta + y\mu + z\lambda\sin\theta$$
$$y' = -x\mu\cos\theta - y\lambda + z\mu\sin\theta$$
$$z' = x\sin\theta + z\cos\theta.$$

By replacement of μ with -μ, ψ^2 comes in place of ψ.[17] Now denote by α a product of n double factors $\varphi\psi$ or $\varphi\psi^2$, and by $\alpha' = \alpha\varphi\psi$ or $\alpha' = \alpha\varphi\psi^2$ [18] a product of $n + 1$ such factors; the point with the coordinates 0, 0, 1 goes over through α to x, y, z, and through α' to x', y', z', so that between these coordinates exist the equations ($\varphi\psi$) or the equations ($\varphi\psi^2$) that emerge from the former by replacement of μ with -μ. We claim that x/sinθ, y/sinθ are polynomials in the (n-1)th degree, z a polynomial in the nth degree, of cosθ,[19] so that

$$x = \sin\theta(a\cos\theta^{n-1} + ...)$$
$$y = \sin\theta(b\cos\theta^{n-1} + ...)$$
$$z = c\cos\theta^{n} + ...$$

This is indeed correct for $n = 1$, as the point 0, 0, 1 goes over through $\varphi\psi$ or $\varphi\psi^2$ to the point λsinθ, $\pm\mu$sinθ, cosθ, and it is carried from n to $n + 1$, because after the equations ($\varphi\psi$) or ($\varphi\psi^2$) we have:[20]

$$x' = \sin\theta(a'\cos\theta^{n} + ...)$$
$$y' = \sin\theta(b'\cos\theta^{n} + ...)$$
$$z' = c'\cos\theta^{n+1} + ...$$

[16] To obtain the transformation matrix for ($\varphi\psi$) we have to multiply the transformation matrix of ψ on the left of the transformation matrix of φ, namely, we first apply φ and then ψ; Hausdorff uses here the same convention as did Bernstein, namely, that $\varphi\psi$ means that first φ is executed then ψ; this convention applies to all members of G.

[17] In both (ψ) and ($\varphi\psi$); this is because $\sin 4/3\ \pi = -\sin 2/3\ \pi$.

[18] This is not the α' mentioned above.

[19] The claim is laid for all x, y, z that are obtained from 0, 0, 1 by α-type members of G.

[20] To obtain these equations put in place of x, y, z of ($\varphi\psi$) (or ($\varphi\psi^2$)) their expansion under the induction hypothesis. For z' we obtain the factor $\sin\theta^2$ which should be replaced by $1-\cos\theta^2$.

For the highest coefficients exist thereby the formulas

$$a' = \lambda(c\text{-}a),\ b' = \pm\mu(c\text{-}a),\ c' = c\text{-}a,\ c'\text{-}a' = (1\text{-}\lambda)(c\text{-}a) = 3/2(c\text{-}a),$$

from which, through repeated application, c-a = $(3/2)^n$ is concluded. The z coordinate of the point 0, 0, 1 will thus be transformed through a product α of n double factors into $z = (3/2)^{n\text{-}1}\cos\theta^n + \ldots$ and anyway is not reducible identically (for all θ) to the value 1. Thus there are only finitely many values of $\cos\theta$, for which $\alpha = 1$ can be, and it is possible, by avoidance of at most denumerably many values, so to choose the angle θ that no product α will equal 1.[21]

The hereafter not only formal, but truly pairwise different rotations (G) we now divide in three classes A, B, C in the manner that of two rotations ρ, $\rho\varphi$, one to A, the other to B + C and from three rotations ρ, $\rho\psi$, $\rho\psi^2$ one belongs to each of A, B, C. This division is possible. When it is already successful for products of at most n factors (to the extent that the rotations in question ρ, $\rho\varphi$ and so on, have at most n factors), it is conductible for products of at most $n + 1$ factors; because a product of $n + 1$ factors is either $= \rho\varphi$, where ρ is a product of n factors and ends with ψ or ψ^2, and would, indeed as ρ belongs to the classes A, B, C, be assigned to the classes B, A, A, or it is $= \sigma\psi$ or $\sigma\psi^2$, where σ is a product of n factors which ends with φ, and then would, indeed as σ would belong to A, B, C, $\sigma\psi$ would be assigned to the classes B, C, A and $\sigma\psi^2$ to the classes C, A, B. A collision in these designations is impossible. The beginning of the processes is obviously of the following composition (when we count the identity 1 to class A):[22]

A	1		$\psi\varphi, \psi^2\varphi, \varphi\psi^2$	$\varphi\psi\varphi$	...
B		φ, ψ		$\varphi\psi^2\varphi, \psi\varphi\psi, \psi^2\varphi\psi$	...
C		ψ^2	$\varphi\psi$	$\psi\varphi\psi^2, \psi^2\varphi\psi^2$	...

Finally let Q be the denumerable set of fixed-points (the rotations poles) of the rotations of our group G which are different from 1[23] and K = P + Q. A point x of P goes over by the rotations (G) in the pairwise different points x, $x\varphi$, $x\psi$, $x\psi^2$, …,[24] denoted by P_x, and two such sets P_x, P_y are either identical or have no point in common. Choose out of each of the sets P_x exactly one point and in this fashion the set $M = \{x, y, \ldots\}$[25] is constructed; then $P = M + M\varphi + M\psi + M\psi^2 + \ldots$ is finitely molded, according to the above division into classes of the rotations, into three sets, which we wish again to denote by A, B, C:

[21] There are only n values for $\cos\theta$ for which the equation $(3/2)^{n\text{-}1}\cos\theta^n + \ldots \text{-}1 = 0$ is fulfilled, and the totality of these values (for every n) is denumerable. It is thus possible to take a θ such that $\cos\theta$ is not in that totality.

[22] The partitioning of G given here is different from the one we used in Sect. 1.1 because here 1 is included and put in A.

[23] Every member ρ of G can be represented by one rotation, the matrix of which is the multiplication of all the matrices of the components of ρ. All these rotations are different. The poles of the axis of this rotation are its fixed points. Since G is denumerable so is Q.

[24] Note Hausdorff's convention $x\varphi$ (and below $M\varphi$) instead of the standard today $\varphi(x)$. It explains why in composite mappings the order of execution is from the left. Bernstein did not use this convention.

[25] In the definition of M the axiom of choice enters. This definition follows Vitali 1905 (Schoenflies 1913 p 374f).

$$\begin{aligned}
P &= A + B + C,\\
A &= M + M\psi\varphi + M\psi^2\varphi + M\varphi\psi^2 + \ldots\\
B &= M\varphi + M\psi + \ldots\\
C &= M\psi^2 + M\varphi\psi + \ldots
\end{aligned}$$

and by the construction $A\varphi = B + C$, $A\psi = B$, $A\psi^2 = C$, so that the sets A, B, C, B + C are congruent,[26] whereby the proof is complete.

27.3 Analytic Geometry Background

The following excerpt from Candy 1904 (p 68) explains the origin of the linear transformations used by Hausdorff (italics in the original).

> 56. To transform from one set of rectangular axes to another, having the same origin
>
> Let (x, y) be the coordinates of any point P referred to the old axes *OX* and *OY*; and (x', y') the coordinates of the *same* point referred to the new axes *OX′* and *OY′*. Let the angle $XOX' = \theta$.
>
> Draw the coordinates *MP* and *NP*, and the lines *QN* and *RN* parallel to *OX′* and *OY* respectively. Then the angle $NPQ = \theta$, $OM = x$, $MP = y$, $ON = x'$, $NP = y'$, $OR = ON\cos\theta = x'\cos\theta$, $RN = ON\sin\theta = x'\sin\theta$, $QN = NP\sin\theta = y'\sin\theta$, $QP = NP\cos\theta = y'\cos\theta$. But $OM = OR\text{-}QN$, and $MP = RN + QP$. Therefore $x = x'\cos\theta\text{-}y'\sin\theta$, and $y = x'\sin\theta + y'\cos\theta$.

For (ψ) we have to reverse the roles of the X−Y and X′−Y′ coordinate systems, namely, to take the rotation clockwise (Fig. 27.1). For θ in the figure we take 2/3 π; the fact that our θ is obtuse does not change the development below. For (φ) we have a case of reflection; that is, of rotation of the X−Z plane around an axis in the plane, at an angle $\theta/2$ from Z. Clearly in this case Y changes its direction hence $y' = \text{-}y$, as stated. Now, in the figure above, let us put Z for X′, X for Y′, Z′ for Y and X′ for X. The rotation axis is the line intersecting the angle between Z and Z′ (Y and X′, in the notation of the figure). For the θ in the figure we take $\pi/2\text{-}\theta$, where $\theta/2$ is the angle between the rotation axis and Z (so θ is the angle between Z and Z′). The same proof as in the figure, with the changed notation, yields the desired transformation equations for x', z'.[27] More directly, we can simply change Z for Y and Z′ for Y′ and add to the drawing the following auxiliary lines: the perpendicular from M on X′, and let T be its point of intersection and the line continuing NR downward

[26] Each P_x is similarly partitioned into three partitions with the same relations. But a P_x is not a counterexample to Lebesgue's problem because its content is 0. Taking a $x' \neq x$ in P_x will generate a different partitioning of P_x and so to obtain the desired partitioning a specific member from each network must be chosen by AC. Changing the choice-set M will generate different partitioning of P.

[27] We can assume that θ is acute for otherwise we get a symmetrical case with the rotation axis on the other side of Z. Thus we can use in the proof the familiar identities $\cos(\pi/2\text{-}\theta) = \sin\theta$, $\sin(\pi/2\text{-}\theta) = \cos\theta$.

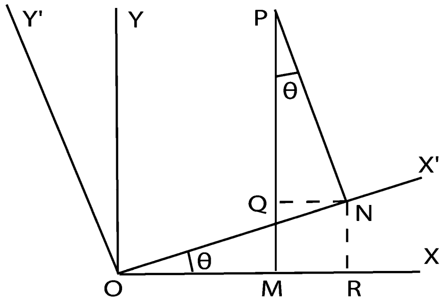


Fig. 27.1 Transformation of axes

meeting the previous line at P′. Then P′N is z and P′M = PN = x'. It is then easy to see that:

z′ = OT + TN = OMcos(π/2-θ) + P′Nsin(π/2-θ) = $x$sinθ + $z$cosθ and

x′ = PN = P′M = P′T-TM = zcos($\pi/2$-θ)-xsin($\pi/2$-θ)= -xcosθ + zsinθ

as given in (φ).

Chapter 28
Sierpiński's Proofs of BDT

We review two proofs of BDT ($2\mathfrak{m} = 2\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$; see Chap. 14) given by Sierpiński[1]: the first from 1922 without use of the axiom of choice and the second from 1947 in the context of Lebesgue measure problem. We suggest that the proofs can be extended to prove the generalized BDT ($\nu\mathfrak{m} = \nu\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$ for any finite ν). We then review the proof of the inequality-BDT for $\nu = 2$ ($2\mathfrak{m} \leq 2\mathfrak{n} \rightarrow \mathfrak{m} \leq \mathfrak{n}$) from another paper of Sierpiński from 1947.

28.1 Sierpiński's First Proof

Sierpiński, a central figure within the Polish school and a devout scholar of cardinal arithmetic, paid attention to BDT on several occasions. His 1922 paper was one of the first papers in the research program for results in cardinal arithmetic without use of the axiom of choice. Within the Polish school, it was the first to employ J. Kőnig's strings gestalt (see Chap. 21). In this, Sierpiński followed the research of D. Kőnig, who also attempted to use the strings gestalt to prove BDT. However, whereas D. Kőnig was led in his research to interpret BDT in the theory of graphs, Sierpiński remained in the context of set theory.

Sierpiński provides two reasons for his motivation to give a new proof of BDT: one is that Bernstein's proof is "rather complicated",[2] and the other is that it does not solve a related problem that Sierpiński sets out to solve:

> Let M, N, P, Q be four given sets, such that M ~ N, P ~ Q, and M + N ~ P + Q and suppose as given the 1–1 mappings φ, ψ and θ respectively between the elements of M and N, of P and Q and of M + N and P + Q: it's about determining a 1–1 mapping between the elements of the sets M and P.

[1] Sierpiński, like D. Kőnig (1916), refers to BDT as Bernstein's theorem.

[2] D. Kőnig agreed with this view, see Sect. 22.2.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_28, © Springer Basel 2013

Sierpiński's stress on the mappings as given, a point ignored by Bernstein, was made because Sierpiński wanted to improve Bernstein's BDT result, just as Zermelo and Peano improved the Bernstein-Borel proof of CBT, by providing a mapping for the result obtained from the given mappings. The stress that the mappings are between members of the sets was made perhaps because already at that time mappings between collections of sets were discussed. By 1927, this will become another research project within the Polish school (see Chap. 31).

Sierpiński's assumptions in the theorem are simply the translation of the assumptions of BDT from the language of cardinal numbers to the language of sets and mappings. Bernstein himself made the same translation as soon as he started his proof. So, with regard to the hypothesis, Sierpiński is adding nothing to the original. But by focusing on finding the 1–1 mapping between M and P, he does step aside from Bernstein. Bernstein transformed the given sets, while maintaining their equivalence, in order to produce an infinite sequence of disjoint images of M in both Q∩N and M∩P. Then he obtained that Q is equivalent to a subset of N and M to a subset of P, so, by CBT and the given equivalences, M ~ P. However, Bernstein did not compile the transformations to suggest how a mapping between M and P is obtained. In fact, the transformations that he uses are too complicated to provide practically such mapping. Sierpiński's proof does not have the transformation stage, for which reason it can be considered less complex; also, his proof does suggest how the desired mapping is obtained.

Two other reasons to justify his new proof, Sierpiński mentions in a footnote with regard to the proof given by D. Kőnig (1916). He claims that D. Kőnig's proof makes use of the axiom of choice and that it "does not give any means to establish effectively the correspondence in question". The first point is correct and Sierpiński's proof avoids the axiom of choice, of which D. Kőnig made unintentional use. For this reason the second point, which is not more than the second point mentioned already above with regard to Sierpiński's problem, is correct too, for D. Kőnig obtained the existence of a mapping assuming the axiom of choice and Sierpiński (1921) saw a definition of an object by way of the axiom of choice not an effective definition.

Sierpiński did not mention the generalized BDT but it seems that his proof for $\nu = 2$ can be extended to any finite ν as we will suggest below.

Sierpiński first notes that it can always be assumed that the given four sets are pairwise disjoint. In this already Sierpiński begins to diverge from Bernstein's proof, for Bernstein assumed that the members of each pair are disjoint but that M + N = P + Q = S. Namely he took the pairs M, N and P, Q to represent two different partitioning of the same set S. Bernstein then proceeded to manipulate with the four "quarters": M∩P, M∩Q, N∩P, N∩Q. Instead of S and its four quarters

Sierpiński defines T = M + N + P + Q, with its four quarters.[3] Following Bernstein, Sierpiński takes φ to be defined over N by φ^{-1}. Sierpiński then extends φ over P by equating it with ψ and over Q with ψ^{-1}. Thus φ is a transformation of T, and $\varphi^2 = 1$ – the identity mapping. Likewise θ is extended over P + Q by θ^{-1} and so $\theta^2 = 1$. It turns out that over the set T two transformations are available, just as in Bernstein's proof over S. Here already the similarity of gestalt with CBT emerges: two 1–1 mappings are given, which through repeated interlaced application generate sequences of images of the given sets. How exactly these sequences provide the desired result, the metaphor, is the crux of the proof of BDT and Sierpiński's proof differs from Bernstein's principally in this regard.

Sierpiński now defines a sequence of functions χ_k, Bernstein's metaphor (without the confusions of the latter): $\chi_0 = 1$, $\chi_1(a) = \theta(a)$, $\chi_3(a) = \varphi\theta(a)$,[4] $\chi_5(a) = \theta\varphi\theta(a)$, $\chi_7(a) = \varphi\theta\varphi\theta(a), \ldots$; $\chi_2(a) = \varphi(a)$, $\chi_4(a) = \theta\varphi(a)$, $\chi_6(a) = \varphi\theta\varphi(a)$, $\chi_8(a) = \theta\varphi\theta\varphi(a), \ldots$. The χ_n form a group[5] under the composition operation since for every n there is n' such that $\chi_n\chi_{n'} = 1$, namely: ${\chi_{4k+1}}^2 = {\chi_{4k+2}}^2 = 1$, $(k \geq 0)$, $\chi_{4k-1}\chi_{4k} = \chi_{4k}\chi_{4k-1} = 1$ $(k > 0)$. Each of the χ_n is 1–1 because if $a \neq b$ and $\chi_n(a) = \chi_n(b)$ then $a = \chi_{n'}\chi_n(a) = \chi_{n'}\chi_n(b) = b$.

From this point on, however, Sierpiński diverges completely from Bernstein. He designates by S(a) the sequence (or string) of type $\omega^* + \omega$: $\ldots, \chi_{2k}(a), \ldots, \chi_4(a), \chi_2(a), a, \chi_1(a), \chi_3(a), \ldots, \chi_{2k+1}(a), \ldots$.[6] His idea is that the relation of belonging to the same S(a) is an equivalence relation that partitions T. Indeed if $g \in$ S(a) then either S(g) is S(a) or the reverse of S(g) is S(a), so that if S(a) and S(b) have an element in common they contain the same elements.[7] The idea of focusing on the strings S(a) as partitioning T, Sierpiński must have obtained from J. Kőnig's 1906 proof of CBT.[8]

---

[3] The difference between S and T is not analogous to the difference between the single-set and the two-set formulations of CBT because we cannot glide from S to T, or vice-versa, as we can glide from the two-set formulation of CBT to its single-set alternative. But the fact that in both theorems there are two arrangements to the given sets is perhaps worth noting as a gestalt of the situation given by the conditions of the proof and the question arises whether similar alternative arrangements show up in other proofs that lead to apply the metaphors of CBT or BDT.

[4] Sierpiński introduces the convention, for any two functions ψ_1, ψ_2, that $\psi_1(\psi_2(\mathrm{a}))$ is written as $\psi_1\psi_2(\mathrm{a})$ and $\psi_1\psi_1(a)$ as ${\psi_1}^2(a)$.

[5] Unlike Bernstein, Sierpiński does not point this out though he proves the 1–1 character of the χ's by using their inverses.

[6] Compare Banach 1924 (see Sect. 29.1) notions (S) and C(a).

[7] We have somewhat shortened Sierpiński's argument here.

[8] Already in Sierpiński's 1914 paper with Mazurkiewitcz, such strings appear. In that paper a result akin to Hausdorff's paradox was presented. The use of alternating mappings in Mazurkiewitcz-Sierpiński 1914 seems to be the reverse of its use in CBT: instead of the shrinkage evident in CBT proofs (e.g., by Schröder or Schoenflies) an expansion occurs. One begins with one point and generates more points through the application of interlaced mappings (with the set of these points having naturally two subsets congruent to it). This paper seems to us to provide an example of proof-processing and it may hide a whole research project.

Two more lemmas regarding the S(a)'s are stated by Sierpiński: One is that the first element from M + P following or preceding an element $m \in$ M ($p \in$ P) is an element of P (M)[9]; the second is that S(a) contains infinitely many elements of M and of P.

The first lemma is established by noting that if m is a member of M in S(a), of the following six consecutive elements of S(a) (in this or the reverse order) $\varphi\theta\varphi(m)$, $\theta\varphi(m)$, $\varphi(m)$, m, $\theta(m)$, $\varphi\theta(m)$, the following must belong to P: either $\theta(m)$ or $\varphi\theta(m)$, and either $\theta\varphi(m)$ or $\varphi\theta\varphi(m)$. In other words, either $\chi_1(m)$ or $\chi_3(m)$, and either $\chi_4(m)$ or $\chi_6(m)$. An analogous argument applies when p and P are used instead of m and M. Now, since for every m there are p's on both its sides and for every such p there is an m in the direction going away from its m, there are infinitely many M's and P's in every S(a), though not necessarily different (as Sierpiński notes on page 5). It is easy to see that if one element is repeated then the string is composed of a repeating finite cycle[10] and the image under any χ_n of a repeating element does not depend on its occurrence place in the string.

Thus, the natural plan seems to be to define for each equivalence class a mapping of m to the next p and then take the union of all these mappings over all equivalence classes. But there is a crucial difference between the situation in J. Kőnig's proof of CBT and the situation devised by Sierpiński: there the application of either mapping moved from left to right along the strings while here the application of θ may move an element either to its right or to its left depending on where the element is located on the string. The reason for the difference is that there we are dealing with bijections while here with transformations. Thus we cannot assign to m the nearest p on its right; the equivalence class of a string can be arranged in two ways and the notion of 'left to right' is not fixed. We can fix it by choosing an arbitrary element from each equivalence class, say K, and fixing the arrangement of the other elements relative to it, but because the number of K's is infinite such choices involve the axiom of choice, which Sierpiński wants to avoid.

Therefore Sierpiński devised a different strategy[11] and in it lies the novelty of his proof. Define by induction subsets K_n of K as follows: K_1 is the set obtained from K after removal of all m's and p's such that $p = \chi_1(m)$. Assuming K_n is defined, define K_{n+1} to be the set obtained from K_n after the removal of all m's and p's such that $p = \chi_{n+1}(m)$. Note that some of the K_n may be empty but we can drop these K_n so we can assume the K_n to be non-empty. Note further that the definition of the K_n does not depend on a particular arrangement of K.[12] Now there are two possibilities: Either $K_\omega = \cap K_n$[13] contains no member of M + P, in which case

[9] This results from Sierpiński's choice of T and would not result for Bernstein's S.

[10] Apparently, Sierpiński regards even strings composed of finite cycles to be of order-type $\omega^* + \omega$.

[11] Note that our discussion here is about the metaphor of the proof.

[12] Strangely, Sierpiński seems to be working with a particular S(g) for each K but his construction by-passes this appearance. Also Sierpiński does not define K_n for every n but only for such n that K_n is not empty. The procedure works without change even when K is composed of only a finite number of different elements.

[13] Sierpiński uses the notation K_ω following perhaps Schoenflies (see Sect. 12.1).

the mapping from the members of M in K and the members of P in K is given by $p = \chi_n(m)$ if m and p were omitted at the stage n. Or, K_ω is not empty, in which case it must contain only one member, say g, of M + P, for otherwise there would be m and p in K_ω and there would be an n such that $p = \chi_n(m)$ but then m and p would have been eliminated at the nth stage and not appear in K_i for $i \geq n$, so not in K_ω. Given this g we can take for K the arrangement of S(g) and assign to every m the nearest p on its right without invoking the axiom of choice.[14]

The final mapping between M and P is obtained by unifying the mappings of the various K's. Can we say that the proof gives this mapping effectively? The answer, according to Sierpiński, seems to be yes. Though the mapping within a K is defined to be either of two possible mappings, so it is an ambiguous definition; but according to Sierpiński 1932, such ambiguous definitions are effective.[15]

Sierpiński, unlike Bernstein and D. Kőnig, did not mention the generalization $\nu\mathfrak{m} = \nu\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$. The generalized theorem should be worded thus: Let M_i, P_i $1 \leq i \leq \nu$, be 2ν given pairwise disjoint sets, such that $M_i \sim M_{i+1}$, $P_i \sim P_{i+1}$, $M_\nu \sim M_1, P_\nu \sim P_1$, all by φ, and $M' = \sum M_i \sim \sum P_i = P'$ by θ, so that φ, θ transform $T = M' + N'$ and $\theta^2 = 1$ and $\varphi^\nu = 1$; determine a 1–1 mapping between M_1 and P_1. It is not clear why Sierpiński did not relate to the general case. It is unlike him to have disregarded the general case whether he had a proof for it or not. To us it seems that his proof can be adapted to prove the general case.

In the construction of the mappings for the general case we have to allow powers of φ, φ^i, $1 \leq i \leq \nu-1$, and at each step we must also allow multiplication by θ to maintain that the generated network, no longer string, contains the same elements no matter with which element we start constructing it. $\chi_0 = 1$ and the other mappings are of the form $\varphi^{i_j}\theta \ldots \varphi^{i_2}\theta\varphi^{i_1}$, where the indices are all between 1 and $\nu-1$. These mappings are still 1–1. Using any procedure for enumerating finite sequences of natural numbers (Troelstra 1969 p 34) we can represent the mappings as χ_n.[16] The networks constructed by the χ_n still form equivalence classes. The two lemmas hold for M_1 and P_1 and Sierpiński's definition of the K_n can be maintained with respect to M_1 and P_1. If K_ω is not empty then it can contain only one element g from $M_1 + P_1$. If $g \in M_1$ then if $m \in M_1$, $m = \chi_n(g)$, we can assign to m the $p = \chi_\mu(g)$ with the smallest index $\mu > n$. If $g \in P_1$ we define in the same way the mapping from P_1 to M_1 and then take its inverse.

[14] D. Kőnig (Kőnig 1926 p 133 footnote 1) conceded that Sierpiński's proof does not make use of AC. He added that this achievement seemed to him impossible when the 1923 paper was written (1914). He seems to be using hindsight here because it seems that in 1914 he was not aware of his use of AC. Unfortunately, D. Kőnig maintained there (the text to the referenced footnote) that only for $\nu = 2$ Bernstein's proof avoided AC, while also Bernstein's general proof made no use of AC.

[15] It seems improbable that Cantor would have accepted an ambiguous definition of a set with his emphasis on 'well-defined' (see Sect. 3.1). Anyway, accepting ambiguous sets as effective seems to be a matter of convention.

[16] Tarski (1949a p 84) said of a similar family of mappings that it can be so represented because it is denumerable. A specific enumeration can be defined by complete induction from the enumeration of the set of all finite sequences of natural numbers as given in Troelstra or from Cantor's enumeration of the rationals.

28.2 Sierpiński's Second Proof

In 1947, after the Polish school was destroyed, Sierpiński (1947a) came back to BDT. More precisely, he proved, for $v = 2$, D. Kőnig's Factoring Theorem[17]:

> If there exists a bi−ν mapping between two sets P and Q, there exists also a 1–1 mapping f such that it makes only such elements correspond that correspond also by the bi−ν mapping (namely, such that f(p) ∈ K(p) for p ∈ P).[18]

The bi−v mappings were introduced by D. Kőnig (see Sect. 22.4). Sierpiński defines these mappings as follows:

> Between two arbitrary sets P and Q a bi−ν mapping [*correspondence bi−ν−voque*] is defined (ν a natural number) if a mapping is defined which corresponds to each element p of P a set K(p) consisting of ν elements of Q, such that for every q of Q there exist exactly ν elements p of P such that q ∈ K(p).

D. Kőnig proved his bi−v mappings Factoring Theorem by way of a Factoring Theorem for graphs, which he had developed between 1914 and 1924, partly by applying observations he made on J. Kőnig's strings gestalt. Sierpiński made no recourse to graph theory but he likewise applied J. Kőnig's strings gestalt. Sierpiński assumed tacitly in the definition of the bi−v mappings that all images are different, whereas D. Kőnig allowed multiplicities and did not use K(p). Sierpiński explicitly structured the proof to apply AC, unlike D. Kőnig who used the axiom unintentionally. In the 1947a paper, Sierpiński connected the result to the measure problem, demonstrating how the proof provides a Lebesgue non-measurable set of real numbers.[19] The elegance of the proof and its application to the measure problem, justified its publication.

Also with regard to his second proof Sierpiński did not mention the generalized BDT; it appears that this proof too can be adapted to provide for the general case. D. Kőnig proved the theorem for any v. Obviously, Sierpiński was interested in the application of the case $v = 2$ not in BDT itself.

Sierpiński's proof proceeds as follows: Sierpiński denotes by S the set of ordered-pairs (p, q) for which $q \in K(p)$ and for every $s = (p, q) \in S$ he denotes by $\varphi(s)$ the other member of S with the same p and by $\psi(s)$ the other member of S with the same q. Clearly $\varphi\varphi(s) = s = \psi\psi(s)$. By E($s$) Sierpiński denotes the set of the following two sequence of type $\omega^* + \omega$:

$$\sigma_s = (\ldots, \psi\varphi\psi(s), \varphi\psi(s), \psi(s),\ s, \varphi(s), \psi\varphi(s), \varphi\psi\varphi(s), \ldots)$$

$$\sigma_s{}^* = (\ldots, \varphi\psi\varphi(s), \psi\varphi(s), \varphi(s),\ s, \psi(s), \varphi\psi(s), \psi\varphi\psi(s), \ldots)$$

---

[17] Sierpiński notes (p 41) that he was not able to prove the equivalence of BDT and the Factoring Theorem without the axiom of choice. As the Factoring Theorem clearly implies BDT, it must be the opposite implication that requires AC.

[18] The bracketed statement was added by Sierpiński; see shortly.

[19] So it makes sense that proving the Factoring Theorem from BDT would require AC.

These two sequences contain the same elements but in reverse order. Sierpiński notes that the elements of the sequences are not necessarily different. But he does not note that in this case the sequences are composed of a finite repeating cycle. It is easy to see that σ_s and $\sigma_{s'}$ are either disjoint or identical and so are the E(s). Let F be the collection of all E(s). By the axiom of choice for a collection of pairs, there exists a set Z that contains one and only one member of each E(s). The order from left to right of Z Sierpiński denotes by $<$.

To every $p \in$ P there is one and only one element in Z, e_p, that contains $s = (p, q)$ for some $q \in$ Q and e_p contains also $s' = \varphi(s) = (p, q')$ where q' is the second element in Q that corresponds to p, namely, that belongs to K(p). If s' is to the right of s Sierpiński defines $f(p) = q$; else, in which case s is to the right of s', $f(p) = q'$. Note that the definition of $f(p)$ does not depend on the choice of s or s'; choosing $s = (p, q'), f(p)$ would still obtain the same element in Q. Note further that $\varphi((p, f(p)))$ is always to the right of $(p, f(p))$. If $p \neq p_1$ and $f(p) = f(p_1)$ then denoting $s = (p, f(p))$ and $s_1 = (p_1, f(p_1))$, we have that $s < \varphi(s)$, $s_1 < \varphi(s_1)$ and $s_1 = \psi(s)$, so that $\psi(s) < \varphi\psi(s)$. But it is not possible to have $s < \varphi(s)$ and $\psi(s) < \varphi\psi(s)$ together in σ_s or σ_s*. So if $p \neq p_1$ then $f(p) \neq f(p_1)$ and f is 1–1.

To prove finally that f is on let $q \in$ Q. Then there are $p, p_1 \in$ P such that $s = (p, q)$ and $(p_1, q) = \psi(s)$ are in S. Let $\varphi(s) = (p, q_1)$ and $\varphi\psi(s) = (p_1, q_1)$. Then in e_p we have either $\varphi\psi(s) < \psi(s) < s < \varphi(s)$ or $\varphi(s) < s < \psi(s) < \varphi\psi(s)$. In the first case we have $q = f(p)$ and in the second case $q = f(p_1)$. Hence f is on Q.

If the theorem could be proved without the aid of the axiom of choice, Sierpiński demonstrated how a linear set could be found that is not Lebesgue measurable. We will not follow his arguments but we note that the context of BDT links, again, with the problem of measure, as in Hausdorff's paradox.

The proof can be generalized to a proof of the general BDT but it loses its charm. The application of AC would amount to choosing a member from each network, which allows the definition of the desired mapping along the lines it was defined in the first proof when K_ω was not empty.

28.3 $2\mathfrak{m} \leq 2\mathfrak{n} \rightarrow \mathfrak{m} \leq \mathfrak{n}$

In yet another paper of 1947 Sierpiński (1947d) proved the inequality version of BDT (inequality-BDT): $v\mathfrak{m} \leq v\mathfrak{n} \rightarrow \mathfrak{m} \leq \mathfrak{n}$ for $v = 2$. This theorem appeared, among many similar results, in Lindenbaum-Tarski 1926 (p 305 and passim). There it was claimed that for $v = 2$ the theorem was proved by Tarski in 1924 and for any finite v by Lindenbaum in 1926. However, Lindenbaum was killed during World War II, on account of him being a Jew, probably in the Ponary massacre site, and so his proof of the theorem was never published. Sierpiński admitted to his inability to prove the general theorem and he noted that it seems to him that Lindenbaum's proof of the generalized theorem would be difficult even for the case $v = 3$. It was perhaps this comment that provoked Tarski to publish his 1949b paper, dedicated to Sierpiński "in celebration of his 40 years as teacher and scholar", where a proof of the general theorem is given (see Chap. 34).

Sierpiński presents the theorem in the same terminology of his first proof of BDT with one difference: θ is not assumed to be on P + Q but *into*. Thus it is required to find a 1–1 mapping from M *into* P. When $x \in$ P + Q-θ(M + N) it is said that $\theta(x)$ does not exist. The S(a) partition T but not all members of S(a) exist. Sierpiński distinguishes four cases for an element x[20]:

1. S(x) is infinite only on its right side. Then the first element on the left of S(x) is $\chi_{8p+4}(x)$ for some finite p.
2. S(x) is infinite only on its left side. Then the first element on the right of S(x) is $\chi_{8q+5}(x)$ for some finite q.
3. S(x) is finite. Then the first element on the left of S(x) is $\chi_{8p+4}(x)$ for some finite p and the first element on the right of S(x) is $\chi_{8q+5}(x)$ for some finite q. In this case the number of elements in S(x) is $4(p + q) + 6 = 2(2(p + q) + 3) = 2s$ where s is odd. S(x) can then be presented by $a_1(x)$, $a_2(x)$,... $a_{2s}(x)$. We have $a_1(x)$, $a_2(x) \in$ P + Q, because otherwise the sequence could be extended to the left. Thus we have that $a_2(x) = \varphi a_1(x)$ and therefore $a_3(x) = \theta a_2(x)$, $a_4(x) = \varphi a_3(x)$, etc., and finally, as s is odd, $a_{s+1}(x) = \varphi a_s(x)$ and $a_s(x) \in$ M + P iff $a_{s+1}(x) \in$ N + Q.
4. S(x) is infinite on both sides (possibly with repetitions), namely, all elements of S(x) exist.

The mapping f from M into P is now defined as follows for $x \in$ M:

If we have for x the case (1) we define $f(x) = \theta(x)$ when $\theta(x) \in$ P or $\varphi\theta(x)$ when $\theta(x) \in$ Q.
If we have for x the case (2) we define $f(x) = \theta\varphi(x)$ when $\theta\varphi(x) \in$ P or $\varphi\theta\varphi(x)$ when $\theta\varphi(x) \in$ Q.

If we have for x the case (3) we define $f(x)$ as follows:

$f(x) = \theta\varphi(x)$ when $a_s(x) \in$ M + P and $\theta\varphi(x) \in$ P,
$f(x) = \varphi\theta\varphi(x)$ when $a_s(x) \in$ M + P and $\theta\varphi(x) \in$ Q,
$f(x) = \theta(x)$ when $a_s(x) \notin$ M + P and $\theta(x) \in$ P,
$f(x) = \varphi\theta(x)$ when $a_s(x) \notin$ M + P and $\theta(x) \in$ Q.

The conditions on $a_s(x)$ are added to qualify the string. Since we do not know the order of the string and we cannot say that we assign to x the first P on its right, we use a_s as a marker to signal if we use the φ direction or the θ direction. The marker will be used below.

Let M_1 be the set of all x's for which case (4) holds. Let N $= \varphi(M_1)$, P_1 be the set of all elements of P that appear in S(x) of some $x \in M_1$ and $Q_1 = \varphi(P_1)$. Sierpiński then proves the lemma that $P_1 + Q_1 = \theta(M_1 + N_1)$ and hence he defines f on M_1 to equal the 1–1 mapping defined in his first proof of BDT to make $P_1 \sim M_1$ (see Sect. 28.1).

To prove the lemma the following is observed: If $x \in M_1$ then as case (4) applies to x, S(x) is infinite and $\theta(x)$ is in S(x). If $\theta(x) \in$ P then $\theta(x) \in P_1$; otherwise,

[20] The four cases appear in J. Kőnig's proof of CBT (see Sect. 21.2).

$\varphi\theta(x) \in P$, hence $\varphi\theta(x) \in P_1$. But then $\varphi(\varphi\theta(x)) = \theta(x) \in Q_1$ and so for every $x \in M_1$ $\theta(x) \in P_1 + Q_1$. If $y \in N_1$ then $y = \varphi(x)$ for some $x \in M_1$ and $\theta(y)$ belongs to $S(x)$. So either $\theta(y) \in P_1$ or $\varphi\theta(y) \in P_1$. In the second case again we have that $\theta(y) \in Q_1$. So $\theta(M_1 + N_1) \subseteq P_1 + Q_1$.

To prove the reverse inclusion, let $y \in P_1 + Q_1$. If $y \in P_1$ there is then $x \in M_1$ such that y appears in $S(x)$. As $S(x)$ is infinite on both sides, also $\theta(y)$ exists and appears in $S(x)$. If $\theta(y) \in M_1$ then $y \in \theta(M_1 + N_1)$; otherwise, $\varphi\theta(y) \in M_1$ and $\theta(y) = \varphi\varphi\theta(y) \in N_1$ and again $y \in \theta(M_1 + N_1)$. If $y \in Q_1$ there is $z \in P_1$ such that $y = \varphi(z)$ and there is $x \in M_1$ such that z is in $S(x)$. So y is also in $S(x)$ and so is $\theta(y)$. If $\theta(y) \in M$ then since $S(x)$ is infinite $\theta(y) \in M_1$ and $y \in \theta(M_1 + N_1)$. If $\theta(y) \in N$ then $\varphi\theta(y) \in M$ and hence $\varphi\theta(y) \in M_1$. So $\theta(y) = \varphi\varphi\theta(y) \in N_1$ and again $y \in \theta(M_1 + N_1)$ and so the lemma is proved.

Hence f is defined for all $x \in M$. Note that in all cases $f(x)$ is a member of $S(x)$. We still have to prove that f is 1–1, namely that if $x \neq x'$ then $f(x) \neq f(x')$, but by the last remark we have to do this only when both x and x' fall under the same case and $S(x) = S(x')$.

If both fall under case (4) then the result follows from the fact that f in this case is 1–1 by the first proof of BDT Sierpiński gave.

If for both x and x' we have case (1) we have to distinguish four cases according to the way f is defined for each of x and x':

(a) If $f(x) = \theta(x)$ and $f(x') = \theta(x')$ then from $f(x) = f(x')$ we obtain $\theta(x) = \theta(x')$ and hence $x = x'$, contrary to the assumption.
(b) If $f(x) = \varphi\theta(x)$ and $f(x') = \varphi\theta(x')$ we again obtain, noticing that $\varphi\theta$ is 1–1, that $x = x'$.
(c) If $f(x) = \theta(x)$ and $f(x') = \varphi\theta(x')$ we have from $\theta(x) = \varphi\theta(x')$ that $x = \theta\varphi\theta(x')$ and also $x' = \theta\varphi\theta(x)$. As we are under case (1) both $S(x)$ and $S(x')$ have the same order and the above two possibilities contradict as x and x' appear to be both on the left and the right of each other.
(d) Like (c) with x and x' reversed. This case is treated analogously to case (c).

If we have for both x and x' the case 2) we have again to distinguish four cases. The four cases are analogous to the four cases of (1). If we have both x and x′ come under case 3) the S(x) and S(x′) are finite and can be ordered similarly or inversely.

In the first case we have $a_s(x) = a_s(x')$ and we distinguish two cases again: If $a_s(x) \in M + P$ then f is defined as in case (2) and likewise it would follow a contradiction to the supposition that $f(x) = f(x')$. If $a_s(x) \notin M + P$ then f is defined as in case (1) and analogously a contradiction arises from assuming that $f(x) = f(x')$.

In the second case the order is reversed. It is in this case that we need the cue of a_s to tell us how f is defined. Then $a_s(x) = a_{s+1}(x')$ and $a_{s+1}(x) = a_s(x')$. We have two symmetric cases according to $a_s(x) \in M + P$ or $a_s(x') \in M + P$. So it is enough to consider the first case. In this case we have four cases to consider:

(a) $\theta\varphi(x) = \theta(x')$, hence $x' = \varphi(x)$ but this is impossible because $x, x' \in M$ but $\varphi(x) \in N$.

(b) $\theta\varphi(x) = \varphi\theta(x')$, hence $x' = \theta\varphi\theta\varphi(x) = \chi_7(x)$ and $x = \varphi\theta\varphi\theta(x') = \chi_8(x')$. But this implies that in both S(x) and S(x') the order of x and x' is the same while we are under the case that their order is reversed.
(c) $\varphi\theta\varphi(x) = \theta(x')$, hence $\theta\varphi(x) = \varphi\theta(x')$ which is again impossible as in case (b).
(d) $\varphi\theta\varphi(x) = \varphi\theta(x')$, hence $\theta\varphi(x) = \theta(x')$ which is impossible as in case (a).

This completes the proof.

The above proof is very calculative and does not seem to be extendible to $v > 2$ because in this case the network does not offer the same analysis into four simple strings. Sierpiński's frustration facing his inability to generalize the proof was understandable. Tarski's route was quite different.

Chapter 29
Banach's Proof of CBT

In 1924, 42 years after Cantor had first mentioned CBT, and after over 15 published proofs, Banach offered a new proof, with new gestalt and new metaphors.

Banach begins his paper by saying that in analyzing the different proofs of CBT,[1] he concluded that they all "make use in an implicit fashion of a general theorem concerning 1–1 mappings (*transformations biunivoques*)":

> Theorem 1: If the function φ transforms in a 1–1 fashion the set A into a subset of B and likewise[2] the function ψ transforms a subset of A into the set B, there exists a decomposition[3] of the sets A and B: $A = A_1 + A_2$, $B = B_1 + B_2$ that satisfies the conditions: $A_1 \cap A_2 = 0 = B_1 \cap B_2$, $\varphi(A_1) = B_1$ and $\psi(A_2) = B_2$.[4]

Among the previous proofs of CBT some were satisfied in demonstrating that there is a 1–1 mapping between A and B (Borel). Others actually showed that the mapping is a certain combination of the given mappings (Peano, J. Kőnig). Still others pointed out the partitioning that Banach speaks about and noted how the given mappings are to be combined to provide the mapping requested (Dedekind, Zermelo 1908b, Peano). Banach's "Partitioning Theorem" thus seems to be a porism of the CBT proofs of the third type mentioned: such a proof simply has to stop short before it reaches the final argument of CBT.

The need for this porism was raised in the context of the Banach-Tarski 1924 paper, containing the Banach-Tarski paradox, which was published in the same Fundamenta Mathematica volume as Banach's paper discussed here. There (Theorem 8) CBT had to be proved for the relation of equivalence by finite decomposition (see below). That proof is an early example in a trend within the structuralist movement in mathematics, of porting CBT to mathematical structures other than sets with equivalences. By applying his Partitioning Theorem to the

[1] Banach refers to CBT as the Schröder-Bernstein theorem or the Equivalence Theorem.

[2] Namely, ψ is also 1–1.

[3] Namely, a partitioning.

[4] Banach uses X for ∩ and 0 is tacitly used as his sign for the empty set.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_29,

relation "equal power", Banach also obtained a proof of CBT in its regular setting (see below).

Banach demonstrated the Partitioning Theorem by combining J. Kőnig's string gestalt (see Sect. 21.2),[5] with the partitioning gestalt of the third group of proofs mentioned above and its impredicative definition of a minimal set of a given property (see Sect. 24.1). However, Banach only mentioned J. Kőnig's paper.

In Banach's paper, features of early structuralistic mathematics are apparent, which had important generalizations (see Chaps. 31 and 34). In particular, CBT was generalized for various relations, besides the relation of 'equal power'.

29.1 Proof of the Partitioning Theorem and Consequences

Proof: If a is an arbitrary element of A, there exist subsets X of A (for example A itself), that fulfill the following conditions:
i. $a \in X$; ii. if $x \in X$ then $\psi^{-1}\varphi(x) \in X$; iii. if $x \in X$ and $\varphi^{-1}\psi(x)$ exists, then $\varphi^{-1}\psi(x) \in X$.
Denote by C(a) the intersection (*la partie commune (le produit)*) of all the subsets X of A, which satisfy the conditions i–iii.[6] It is easy to become convinced[7] that the set C(a) is composed of the following sequence (of type ω or $\omega^* + \omega$)[8]:
(S) $\ldots\varphi^{-1}\psi(a)$, a, $\psi^{-1}\varphi(a)$, $\psi^{-1}\varphi\psi^{-1}\varphi(a)\ldots$
We conclude that the sets C(a) enjoy the following properties:
(1) for all $a \in A$, C(a) fulfills the conditions i–iii;
(2) if $a_1 \in A$ and $a_2 \in A$, then $C(a_1) = C(a_2)$ or $C(a_1) \cap C(a_2) = 0$.[9]

[5] Mańka and Wojciechowska (1984 p 196) noticed that Banach's original proof is based on an idea from the proof of J. Kőnig to CBT. This is an example of an identification of proof-processing relation between proofs that appears often in the literature.

[6] The C(a) are the parts of the strings defined in J. Kőnig 1906 which belong to A, for J. Kőnig's strings included alternatively elements of A and B (using alternatively at each step either φ^γ or ψ^γ, where γ is either 1 or −1). Banach defines the C(a) not by J. Kőnig's inductive procedure but by the impredicative metaphor applied in Zermelo's 1908b CBT proof, which leveraged on Dedekind's definition of the chain of a set. D. Kőnig (1926 p 130f, see Sect. 30.3) pointed out the indebtedness of Banach's result (D. Kőnig mentions Banach's theorem 2 but it seems that he refers to theorem 1) to J. Kőnig's 1906 paper. D. Kőnig's claim that Banach's result is implicit in J. Kőnig's paper is, however, not exact: not only is the use of the impredicative definition of C(a) missing from J. Kőnig, but the focus of Banach is different: he points out the correlated partitionings and this result leads him to structuralistic conclusions – a train of thoughts alien to J. Kőnig.

[7] It seems to us that such conviction can be gathered only using complete induction. If one wants to avoid this procedure then the presentation of C(a) as (S) can only be used as heuristic.

[8] Some of the C(a) may be finite and their (S) cyclic, so apparently, Banach, like Sierpiński, allows the $\omega^* + \omega$ sequences to contain repetitions.

[9] Proof of (2) without complete induction could be as follows: if C(a_1) and C(a_2) are not disjoint, let d be an element of both. Then $C(d) \subseteq C(a_1)$ because C(d) is an intersection of all subsets of A that fulfill i–iii. But a_1 and d are connected by i–iii and this connection is symmetric, so a_1 must be in C(d) and therefore $C(a_1) \subseteq C(d)$ so $C(d) = C(a_1)$. Likewise $C(a_2) = C(d)$, hence (2). It seems, however, that Banach was not interested in having his proof avoid the notion of number.

Denote now by A_2 the set of all elements a of A that verify the formula $C(a)\subset\psi^{-1}(B)$[10]; the set A_2 is thus composed of all the elements a such that the set C(a) can be represented by the sequence (S) of type $\omega^* + \omega$[11] or of type ω, the first element of which belongs to $\psi^{-1}(B)$.
Let (3) $A_1 = A\text{-}A_2$; evidently (4) $A = A_1 + A_2$, (5) $A_1 \cap A_2 = 0$.
Observe several properties of the sets A_1 and A_2:
(6) If $b \in B$ and $\psi^{-1}(b) = a \in A_1$, $\varphi^{-1}(b)$ exists. If, in fact, $\varphi^{-1}(b) = \varphi^{-1}\psi(a)$ does not exist, the sequence (S) would be of type ω and would have as first element the element a, that belongs, by the equality: $a = \psi^{-1}(b)$, to the set $\psi^{-1}(B)$. Hence it could be deduced that $a \in A_2$ contrary to (3).
(7) If $b \in B$ and one of the elements $a_1 = \psi^{-1}(b)$ and $a_2 = \varphi^{-1}(b)$ belongs to A_1, the other one belongs to it as well. As $a_1 = \psi^{-1}\varphi(a_2)$, (7) results immediately from (1), (2) and (3).
Following (1), i and the definition of A_2, it is easily obtained that (8) $A_2\subset\psi^{-1}(B)$.
Pose, in conformity with (8) and the hypothesis of the theorem: (9) $B_1 = \varphi(A_1)$, $B_2 = \psi(A_2)$, whence (10) $B_1 + B_2\subset B$. To establish the inverse inclusion, envisage an arbitrary element b of B. If $\psi^{-1}(b) \in A_2$, b evidently belongs to $\psi(A_2) = B_2$. If, on the contrary, $\psi^{-1}(b) \in A_1$, it is deduced by (6) and (7) that $\varphi^{-1}(b)$ exists and belongs to A_1, from which $b \in \varphi(A_1) = B_1$. It can therefore be concluded that
(11) $B\subset B_1 + B_2$. The inclusions (10) and (11) give immediately the identity
(12) $B = B_1 + B_2$. It results in addition from (7) and (9) that (13) $B_1 \cap B_2 = 0$.[12]
The formulas (4) and (12) provide us a decomposition of the sets A and B; By (5), (9) and (13), this decomposition fulfills all the conditions of the thesis of the theorem.

Banach concludes this proof with the remark that $A_1=\sum_n C_n$ where $C_0 = A\text{-}\psi^{-1}(B)$ and $C_n = \psi^{-1}\varphi(C_{n-1})$.[13]

Next to the Partitioning Theorem Banach gives its following consequences and related definitions:

Definition 1: The relation R possesses the property (α) when:
I. the relation subsists only between sets;
II. the condition ARB implies the existence of a function φ, which transforms in a 1–1 fashion A into B[14] so that we have $XR\varphi(X)$ for every subset X of A.
In conformity with this definition, Theorem 1 implies immediately the following:
Theorem 2. The relation R having the property (α), if $A'\subset A$, $B'\subset B$, ARB' and $A'RB$,[15] each of the sets A and B can be decomposed into two disjoint subsets: $A = A_1 + A_2$, $B = B_1 + B_2$, such that we have: A_1RB_1 and A_2RB_2.

---

[10] Banach uses $\subset$ for not necessarily proper subsets. Note that for all $C(a)$, $C(a)\text{-}\{a\}$ is a subset of $\psi^{-1}(B)$. Banach prefers to define A_2 first and leave for A_1 the $C(a)$ for which the (S) sequence begins with a member of $A\text{-}\psi^{-1}(B)$. The partitioning here defined is not the only partitioning possible: all or part of the (S) which are finite or of type $\omega^* + \omega$ could be moved from A_2 to A_1.

[11] There is a typo in the original where in place of our $\omega^* + \omega$, $\omega + \omega^*$ is printed.

[12] Banach's proof could profit heuristically if the structure induced in B by the structure in A (the $C(a)$'s) would be explicitly described.

[13] Namely, A_1 is composed of the frames $A\text{-}\psi^{-1}(B)$, $\psi^{-1}(\varphi(A)\text{-}\varphi\psi^{-1}(B))$, etc. Banach presents here A_1 as a chain composed of frames (see Sect. 9.2).

[14] Apparently, the meaning here is that φ is onto.

[15] There seems to be a typo in the original: the sets A' and B' are denoted by A_1 and B_1, but clearly from theorem 1 the sets given in the conditions of the theorem ($\varphi(A)$ and $\psi^{-1}(B)$) are not the decomposition sets. When theorem 2 is then invoked for the proof of theorem 3, A_1 and B_1 of theorem 3 correspond to A' and B' of theorem 2.

As examples of relations that have property (α) Banach gives the relation of similar order between sets, the relation of congruence in geometry and the relation of homeomorphism between topological spaces. Banach then continues with his structuralist corollaries:

Definition 2: The relation R possesses the property (β), when:
I. the relation subsists only between sets;
II. the conditions A_1RB_1, A_2RB_2 and $A_1 \cap A_2 = 0 = B_1 \cap B_2$ imply that $(A_1 + A_2)R(B_1 + B_2)$.
Following this definition, the following is immediately deduced from Theorem 2:
Theorem 3: When the relation R possesses the properties (α) and (β), if $A_1 \subset A$, $B_1 \subset B$, ARB_1 and A_1RB, then ARB.
By applying this theorem to the relation of the equality of power of arbitrary sets, the known Schröder-Bernstein theorem is obtained.

As another example of a relation that possesses both properties (α) and (β) Banach cites the relation of equivalence by finite decomposition, which is defined in the Banach-Tarski paper (1924 p 246) as follows:

"The sets of points A and B are equivalent by finite decomposition" AfB, if there exist sets $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots B_n$, that fulfill the following conditions:
I. $A=\sum_{k=1,\ldots,n} A_k$, and $B=\sum_{k=1,\ldots,n} B_k$;
II. $A_k \cap A_l = 0 = B_k \cap B_l$ when $1 \leq k < l \leq n$;
III. $A_k \cong B_k$ for $1 \leq k \leq n$.[16]

29.2 Aspects of the Proof

Theorem 1 has the conditions of CBT in its two-set formulation, expressed in the language of sets and mappings, with a twist: instead of defining ψ from B to a subset of A, symmetrically to the definition of φ, ψ is defined from a subset of A onto B. Banach's reason for this twist was perhaps that he wanted to obtain first the decomposition of A and then the decomposition of B by the mappings φ, ψ. His presentation also leads to the metaphor that shrinking the domain of φ and the range of ψ should lead to the desired result.

Given φ, ψ, two relations between the members of the power-sets of the given sets of Theorem 1, P(A) and P(B), can be defined. XRY when $Y = \varphi(X)$ and XSY when $Y = \psi(X)$. Properties (α) and (β) serve to generalize the results that can be obtained for R, S to arbitrary relations between members of P(A) and P(B). This gestalt switch, new in the history of CBT, was further developed in 1927 by Whittaker and by Tarski and Knaster (see Chap. 31) and it finally led to Tarski's Fixed-Point Theorem (see Chap. 35).

[16] $A \cong B$ signifies congruence, namely, that there is a 1–1 mapping between A and B that preserves the distance between any two points of A. The notion of equivalence by finite decomposition was no doubt borrowed from the Euclidean geometry notion of scissors congruence.

Fig. 29.1 Yin-Yang symbol

From an alternative gestalt, if A and B have a certain structure and φ, ψ are structure preserving mappings (isomorphisms), then Banach's Theorem is about partitioning A and B into isomorphic structures correlated by φ, ψ. This aspect of the proof reveals the structuralistic spirit of the period.

Another aspect of structuralism evident in Banach's paper is the use of the algebraic closure conditions i–iii instead of the gestalt that sees in the structure induced by φ and ψ the geometric structure of domains, ranges and frames.

A third structuralist aspect of Banach's proof is the way proof-processing is applied: when descriptors from a previous proof are identified in a new proof, the original proof is changed so that the new theorem becomes a particular case of the older one. As if particularization is more respectful than correlation by proof-processing. We are talking here about the preference of Banach-Tarski to present their CBT proof for equivalence by finite decomposition as a particular case of the Partitioning Theorem instead of presenting it as proof-processed from the regular CBT. Compared to other proofs of CBT Banach's approach stands out in that it does not provide a direct proof of CBT but requires the proof to march through the gates of the (α), (β) properties.

The image that we attach to Banach's proof is Yin-Yang. The two parts of the circle represent A and B. The parts with the same colors are the partitions that correspond. If the conditions of CBT give us an uneasy feeling that something needs to be corrected, Banach's Partitioning Theorem in its Yin-Yang depiction endows us with catharsis (Fig. 29.1).

Chapter 30
Kuratowski's Proof of BDT

Preceding the 1924 paper of Banach-Tarski that included the paradox that bears their names, was not only Banach's 1924 paper that we reviewed in the previous chapter, but also a paper by C. Kuratowski, on BDT: $2\mathfrak{m} = 2\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$. Kuratowski mentions that Tarski posed to him the problem of generalizing BDT, because he required the result in order to prove an analogous theorem (Theorem 11, $11'$ in Banach-Tarski 1924) for the relations of equivalence by finite or denumerable decomposition. Kuratowski adds that his theorem stands to BDT just as Banach's Partitioning Theorem stands to CBT. What he must have meant is that both papers introduce a partitioning of the sets in the thesis and the mappings that make these partitions correspond. We can add that both proofs make use of J. Kőnig's gestalt of strings (see Sect. 21.2), that was also used in Sierpiński 1922 (see Sect. 28.1). The analogy breaks down, however, along the following line: Banach's proof contained new structuralist insight while the proof of Kuratowski had no similar import.

In addition to referencing Bernstein 1905 (see Chap. 14), Kuratowski references both D. Kőnig 1916 (see Sect. 22.2 and the following sections there) and Sierpiński 1922 (see Chap. 28) on the same theorem, where new proofs were provided. The outstanding feature of Kuratowski's proof is that it gives the mappings that provide the thesis of the theorem, something not done in Bernstein's own proof or that of D. Kőnig. However, the effectiveness of this result is reduced because Kuratowski makes use of the axiom of choice, avoided by Bernstein and Sierpiński. Kuratowski's paper was complemented few years later by his student, S. Ulam, who gave examples to the effect that Kuratowski's result cannot be improved. We will present Ulam's results too.

D. Kőnig's 1926 paper, published also in Fundamenta Mathematica, contains a section devoted to the comparison of his results to the result of Kuratowski. We will review his contentions there in detail below.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_30,

30.1 The Theorem and Proof

The theorem that Kuratowski sets out to prove is the following:

> If a set E is decomposed in two different ways:
> (1) $E = M + N, M \cap N = 0,$[1]
> (2) $[E] = P + Q, P \cap Q = 0,$
> and if there exists a 1–1 mapping[2] $\varphi(x)$ of M on N, as well as a 1–1 mapping $\psi(x)$ of P on Q, then the sets M, Q decompose into four disjoint parts in such fashion that:
> (3) $M = M_1 + M_2 + M_3 + M_4, Q = Q_1 + Q_2 + Q_3 + Q_4,$
> (4) $Q_1 = M_1, Q_2 = \psi(M_2), Q_3 = \varphi(M_3), Q_4 = \psi\varphi(M_4).$ [3]

Kuratowski brings the theorem in its sets and mappings formulation. Bernstein's original formulation, in the language of cardinal numbers, was cited by Kuratowski in the introduction to his paper. But the difference is not essential since Bernstein translated the theorem to the language of sets and mappings upon the beginning of his proof. In the cardinal numbers formulation, the two sets of power $\mathfrak{m}$ and the two sets of power $\mathfrak{n}$ are not assumed related, but Bernstein assumed the setting to be that of one set partitioned in two ways into two equivalent partitions. While Sierpiński took all the sets to be disjoint, Kuratowski returned to Bernstein's setting, still preserving the notation of Sierpiński for the partitions and the mappings.[4]

Like Bernstein and Sierpiński, Kuratowski extends φ over N and ψ over Q so that $\varphi\varphi(x) = x$ and $\psi\psi(x) = x$.[5] Kuratowski then introduces some notation:

> a being an arbitrary element of E, we call the *chain* of a, the smallest set C(a) that contains a and that contains $\varphi(x)$ and $\psi(x)$ once it contains x. In virtue of the identities
> $\varphi\varphi(x) = x = \psi\psi(x)$, the chain C(a) can be represented by a sequence doubly infinite:
> $\ldots \psi\varphi\psi(a), \varphi\psi(a), \psi(a),\ a, \varphi(a), \psi\varphi(a), \varphi\psi\varphi(a), \ldots.$
> (in which the elements may repeat).

Kuratowski's definition of C(*a*), which he calls 'chain' but we will call 'string' because they are J. Kőnig's strings, as the smallest set with a certain property is taken directly from Banach 1924. It is an impredicative definition. The notation C(*a*) is taken from Banach as well; Sierpiński used S(*a*). Kuratowski does not introduce a special notation for the mappings composed of the φ and ψ as did Bernstein who was followed in this by Sierpiński (1922). Like Sierpiński and unlike

---

[1] Evidently Kuratowski denotes the empty set by 0. Intersection he denotes by $\times$ which we replace by $\cap$.

[2] The paper, like most papers of Fundamenta Mathematica until the second world war, is written in French and Kuratowski uses the term "transformation biunivoque".

[3] This decomposition can be ported to N and P by φ and ψ.

[4] Kuratowski uses the same letters as did Sierpiński though he seeks the mapping between M and Q while Sierpiński searched for the mapping between M and P. This remark is, of course, mathematically void.

[5] Kuratowski does not use the notation φ^2 nor 1 for the identity mapping.

Bernstein, Kuratowski does not point out that the collection of all these mappings forms a group, but he clearly assumes that the composite mappings are 1–1. The arrangement of the members of C(a) in a sequence of type $\omega^* + \omega$ appears both in Sierpiński 1922 and Banach 1924, as well as in J. Kőnig (1906) for the non-cyclic strings. Bernstein did not refer to the strings gestalt. Kuratowski notes that C(x) = C(y) or else C(x)∩C(y) = 0 so that the C(x) are equivalence classes for the relation 'pertaining to the same string' (language that Kuratowski did not use) and thus E is partitioned into strings.

This partitioning of E was discovered already by D. Kőnig (1916)[6] who also noted that each partition is either finite or denumerable. Kuratowski's representation of a finite C(a) is also as $\omega^* + \omega$ strings: it lists all elements starting from a, repeatedly to the right of a, and in reverse order to the left of a. It seems that Kuratowski wanted to address the finite, cyclic, strings under a common gestalt with the infinite strings. That repeating elements can occur only when C(a) is finite is not noted explicitly by Kuratowski. It can be proved as follows: Let x be a repeating member of a string and let $x_1, \ldots, x_n$ be the elements between one occurrence of x and the next one on its right. Without loss of generality we can assume that $x_1 = \varphi(x)$. We now consider two cases: $x = \psi(x_n)$ and $x = \varphi(x_n)$. In the first case obviously n is odd and the string is made of sequences $x_1, \ldots, x_n, x$. In the second case n is even and $\varphi(x) = x_n$ so $x_1 = x_n$, $x_2 = x_{n-1}$, and so on until finally $x_{n/2+1} = \varphi(x_{n/2})$ and $x_{n/2} = x_{n/2+1}$, which is a contradiction since φ cannot move an element to itself. So the second case never happens.

Kuratowski next chooses from each string an element belonging to M (such an element must exist). Therewith he consciously made an appeal to the axiom of choice, which was no hindrance for the proof's intended application in the paper of Banach-Tarski because the axiom of choice was anyway assumed in that paper for results concerning equivalence by finite decomposition. Kuratowski adds in a footnote that AC can be avoided if in the theorem instead of decomposition into four sets the decomposition is denumerable. In this case, he says, the theorem will be applicable to equivalence by denumerable decomposition. For proof of the thesis made in the footnote Kuratowski refers the reader to Sierpiński 1922. Sierpiński did not prepare for this application of his proof but his proof can provide the partitioning to which Kuratowski alludes. We have here a clear example of proof-processing. Sierpiński avoided AC but in his proof, when K_ω is not empty, a chosen member from a string is obtained, and for this case Sierpiński's proof is similar to the proof of Kuratowski.

Assume that for an arbitrary string C, x_0 is the chosen M element. So we can write the sequence of C as: (6) $\ldots x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots$ where $x_{2n} = \psi(x_{2n-1}) = \varphi(x_{2n+1})$, $x_{2n+1} = \varphi(x_{2n}) = \psi(x_{2n+2})$.[7] For each x_k in C that belongs to M we define a corresponding element $x_{g(k)}$ of Q as follows: If k is even, $g(k) = k + 1$ if x_{k+1}

[6] J. Kőnig did not point out the partitioning aspect of arranging the members of the set in strings.

[7] There is a typo in the original where it is written $\psi(x_{2n+1})$ and $\varphi(x_{2n+2})$.

belongs to Q, else x_{k+2} belongs to Q and we set $g(k) = k + 2$. If k is odd, $g(k) = k$ if x_k belongs to Q, else x_{k+1} belongs to Q and we set $g(k) = k + 1$.

Since each member of M belongs to one string only, and since with the use of the choice-set each member of M is a certain x_k, this definition covers all the elements of M. But it is not clear that (1) for finite strings, each element of M is assigned in this way only one element of Q, (2) to different elements of M different elements of Q are assigned, and (3) all elements of Q are thereby assigned to members of M.

For (1), all repeating M elements will have the same corresponding element in Q because their index has the same parity by the above proven lemma[8] and is assigned the same $g(k)$ as their following elements are the same so obey the condition of belonging or not belonging to Q in the same way.

For (2), Kuratowski proves that if $x_{g(i)} = x_{g(j)}$ then $x_i = x_j$ when x_i and x_j are in M. If $g(i) \neq g(j)$ the string is finite and by the argument for (1) we have the required result. Otherwise, $g(i) = g(j)$ and we have to prove that $i = j$. We have four cases to consider: that both i, j are even or both are odd or one is even and the other odd. If both are even then when $g(i) = i + 1$ $(i + 2)$, $g(j) = j + 1$ $(j + 2)$ so $j + 1 = i + 1$ $(j + 2 = i + 2)$ and $i = j$. If both i, j are odd, then when $g(i) = i$ $(i + 1)$, $g(j) = j$ $(j + 1)$ so again $i = j$. If i is even and j is odd then if $g(i) = i + 1$ $(i + 2)$, $g(j) = j$ $(j + 1)$. In both cases we have $j = i + 1$. But then $x_j = \varphi(x_i)$ which is impossible because in this case both x_i and x_j cannot be in M. If i is odd and j is even then if $g(i) = i$ $(i + 1)$, $g(j) = j + 1$ $(j + 2)$. In both cases we have $j = i$-1. But then again $x_j = \varphi(x_i)$, and we obtain the same contradiction as before.[9]

For (3), since every member of Q belongs to some string, every member of Q is some x_k. If k is even then $x_{k-1} = \psi(x_k)$ and so is not in Q. If $x_{k\text{-}1}$ is in M then $g(k\text{-}1) = k$, as k-1 is odd; otherwise $x_{k\text{-}2}$ is in M and $g(k\text{-}2) = k$, as k-2 is even. If k is odd, if x_k is in M, $g(k) = k$ since k is odd; otherwise, $x_{k\text{-}1} = \varphi(x_k)$ is in M and $g(k\text{-}1) = k$ since k-1 is even.[10]

Kuratowski now partitions M according to the definition of $g(k)$: if m is a member of M, then m belongs to some string and is some x_k of that string. If k is odd then m is placed in M_1 or M_2 accordingly to x_k being a member of Q or not; if k is pair then m is placed in M_3 or M_4 accordingly to x_{k+1} being a member of Q or not. Q is then partitioned according to the equations in the statement of the theorem. The strength of Kuratowski's result is that the parts of M and Q are related by simple mappings and so the result appears effective when compared to Bernstein's original proof. But this achievement comes at a price – use of AC, which is as far from effective as can be.

---

[8] Kuratowski proves directly that if $x_i = x_j$ then $d = |i\text{-}j|$ is pair by induction on d.

[9] We have changed here Kuratowski's argument to make it clearer (in our view).

[10] We have slightly detailed Kuratowski's argument.

30.2 Examples and Generalizations

As a simple example of a set that is partitioned in two ways into two equivalent partitions, Kuratowski gives an ellipse that is partitioned by its axes. The mappings φ and ψ are congruences. In this example, if M is one of the halves of the ellipse partitioned by the small axis and Q is one of the halves partitioned by the large axis, then the intersection of M and Q is a quarter that corresponds to itself (M_1) and the remaining quarter of M corresponds by ψ to the remaining quarter of Q. So the two halves are congruent by finite decomposition. This example shows that M can be partitioned into two parts.

Another example that Kuratowski brings he attributes to Tarski and describes as interesting. E is the set of real numbers; M the set of real numbers in the interval $(2n, 2n + 1)$; N = E-M; P is the set of real numbers in the interval $(2n\sqrt{2}, (2n + 1)\sqrt{2})$; Q = E-P; n is a variable integer, which probably means that it is the collection of the intervals for every n that make up M and P. Obtaining φ and ψ for this example is simple and Kuratowski does not give it.[11]

Kuratowski invokes Banach's conditions (α), (β) stating two generalizations of his theorem: (a) If R has property (α) and is symmetric, transitive and reflexive,[12] and the partitions of each of the partitioning of the theorem stand in R relation, namely, MRN and PRQ, then $M_iRN_iRP_iRQ_i$, $1 \leq i \leq 4$. (b) If R satisfies (β) too then all four partitions are in the relation R to each other. As examples where (a) can be applied Kuratowski gives the relation of homeomorphism, similarity of order and congruence. As examples for (b) he gives the relation of 'same power' and 'equivalence (geom.)[13] by decomposition'.

In 1929,[14] Ulam answered negatively a question raised by Kuratowski in a seminar whether the number four of the decomposition of M can be reduced. Ulam provides an example of a set E partitioned in two ways into two equivalent partitions M, E-M, Q, E-Q, in which case M and Q cannot be decomposed into two disjoint parts that correspond by mappings from Γ (– the group of all combinations of φ and ψ) [15] and then another example showing that a decomposition into three parts is not possible. The examples are given for finite sets but by taking infinitely many similar sets one gets examples with infinite sets. It is of interest to note that the examples are for graphs, which link the discussion to D. Kőnig's context.

The examples are provided with the following drawings (Fig. 30.1):

[11] The importance of this example escapes us, though we sense its whimsical beauty.

[12] Namely, R is an equivalence relation but Kuratowski does not use this term.

[13] By adding the bracketed expression it seems that Kuratowski wanted to stress that the notion of equivalence by decomposition, as introduced by Banach-Tarski, is in a geometric context and not in the general context of set theory.

[14] Interestingly, Ulam's paper is in English.

[15] Ulam does point out that Γ is a group.

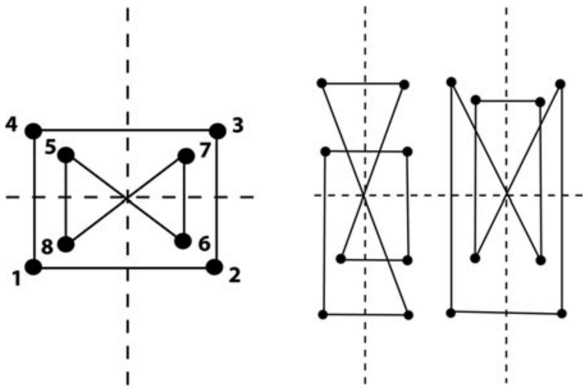


Fig. 30.1 Ulam's counterexamples

In the first example the set E is composed of the numbered points, M is the set of points to the left of the vertical line and the set Q is below the horizontal line. $\varphi(1) = 2$, $\varphi(4) = 3$, $\varphi(5) = 6$, $\varphi(8) = 7$, $\psi(1) = 4$, $\psi(2) = 3$, $\psi(8) = 5$, $\psi(6) = 7$, and inversely. There are only four functions in Γ: 1,[16] φ, ψ, $\varphi\psi$ ($=\psi\varphi$); all other combinations of φ and ψ only produce functions from these four. Assume that α_1 and α_2 are two functions from Γ such that $M = M_1 + M_2$, $Q = Q_1 + Q_2$, $Q_1 = \alpha_1(M_1)$, $Q_2 = \alpha_2(M_2)$. Now, (1) the elements 1, 4 of M must correspond to 1, 2 of Q, and (2) 5, 8 to 5, 6. For (1) to happen either of α_1, α_2 must be the identity 1 and the other must be $\varphi\psi$, or one is φ and the other is ψ. For (2) to happen either of α_1, α_2 must be the identity 1 and the other must be φ, or one is ψ and the other is $\varphi\psi$. So we have a contradiction and the decomposition of M cannot be into two parts. M can be decomposed into three parts $M_1 = \{1, 8\}$, $M_2 = \{5\}$, $M_3 = \{4\}$. Then $Q_1 = M_1$, $Q_2 = \varphi(M_2)$, $Q_3 = \varphi\psi(M_3)$.

The second example can be treated analogously. The horizontal edges stand for φ and the vertical or diagonal edges for ψ. The two sets to the left of the vertical lines compose M and the two sets below the horizontal line compose Q. Γ is reduced to 1, φ, ψ, $\psi\varphi$, $\varphi\psi$, $\varphi\psi\varphi$, $\psi\varphi\psi$, $\varphi\psi\varphi\psi(=\psi\varphi\psi\varphi)$. In this example M cannot be decomposed into three parts that correspond by members of Γ to a composition of Q. Ulam omits the details and so will we.

30.3 D. Kőnig on Kuratowski's Paper

In 1926, D. Kőnig published in Fundamenta Mathematica a paper that summarized his research program on results in graph theory that emerged during his attempt to apply his father's string gestalt, to prove BDT. Towards the end of the 1926 paper, D. Kőnig linked it with Kuratowski's paper discussed here. He showed that

[16] Which Ulam denotes by 0.

Kuratowski's theorem can be deduced from his own, earlier (1916), results, thus establishing his priority in providing a basic lemma to the Banach-Tarski paradox.[17] While Kuratowski's result can be derived from the results of D. Kőnig, it is clear that Kuratowski had different gestalt and metaphor in mind from those of D. Kőnig: the latter wanted to establish the existence of a mapping (metaphor) between M and Q while the former was after the decomposition of M and Q (gestalt) and the indication of which mappings (metaphors) correspond the parts of M and Q. Thus, D. Kőnig's view (1926 p 130f), that Banach's proof of CBT has to his father's CBT proof of 1906 the same relation as Kuratowski's paper has to his own results of 1916, cannot be upheld. Banach's proof certainly relies on the gestalt of J. Kőnig, and Banach did admit the point though Banach sought for different gestalt and metaphor (the partitioning and the equivalence of the partitions), but Kuratowski's proof does not have any similarity to gestalt and metaphors used by D. Kőnig. Indeed, when in 1947 Sierpiński wrote a couple of papers on issues that relate to BDT, he quoted D. Kőnig's result as summarized in 1926 in his 1947a and Kuratowski's 1924 result in his 1947d, depending on the aspects he wanted to emphasize.

D. Kőnig's theorem from which Kuratowski's result can be deduced was presented in 1926 as follows:

> A) If there exists a bi-ν mapping[18] between two arbitrary sets M and N, there exists also a 1–1 mapping such that makes correspond two elements only when they correspond by the given ν−ν mapping. (ν is an arbitrary natural number.)

In 1916 this theorem was worded thus:

> Every bipartite regular graph has a factor of first degree.

If we connect corresponding elements of the sets M, N by edges we obtain a graph. This graph is bipartite because its points are partitioned into two sets (M, N) and only elements that belong to different sets are connected. The graph is regular of degree v because to each element exactly v elements correspond. The factor then gives the 1–1 mapping. So we see that the theorem of 1926 appeared indeed already in D. Kőnig's 1916 paper.

D. Kőnig claimed that Kuratowski's result follows from (A) for $v = 2$. To demonstrate his contention D. Kőnig first added the following terminology (p 129): He introduces a new set E′ disjoint from E and equivalent to it by a 1–1 mapping χ. Thus D. Kőnig adopted the layout used by Sierpiński (1922). Then he extended φ and ψ over E′ so that if $a' = \chi(a)$, $b' = \chi(b)$ and $b = \varphi(a)$ then $b' = \varphi(a')$, and likewise for ψ. Thus $\chi\varphi = \varphi\chi$ and $\chi\psi = \psi\chi$. D. Kőnig now looks at the set of pairs $(e, \varphi(e))$ from E and the set of pairs $(e', \psi(e'))$ from E′. To every pair of E two, not necessarily different, pairs of E′ correspond and the graph of this

[17] Perhaps this was the reason D. Kőnig published this paper in Fundamenta Mathematica and not in Mathematische Annalen, where his 1916 and 1926 (with Valko′) papers appeared.

[18] "*transformation bi-ν-ivoque*" in the original, probably following Sierpiński (see Sect. 28.2).

correspondence is bipartite of second degree. Then by Theorem (A) there is a factor G_1 of this graph that is of first degree. Thus for every $m \in M$, to the couple $(m, \varphi(m))$ either of the following couples is related by G_1: $(m', \psi(m'))$ or $(\varphi(m'), \psi\varphi(m'))$. In each of these couples, one and only one is in Q' and thus has a referent in Q. If $(m, \varphi(m))$ is correlated to $(m', \psi(m'))$ then if $m' \in Q'$ then $m \in Q$ and we assign m to M_1 and correlate m to itself; otherwise, m' is not in Q' and then $\psi(m')$ is in Q' and let its referent be q. Then we relate m to q and put m in M_2. If $(m, \varphi(m))$ is correlated to $(\varphi(m'), \psi\varphi(m'))$ then if $\varphi(m')$ is in Q' let its referent be q; we then correlate m to q and put m in M_3; otherwise, $\varphi(m')$ is not in Q' and then $\psi\varphi(m')$ is in Q' and let its referent be q; we then correlate m to q and put m in M_4.

One can appreciate the virtuosity by which D. Kőnig reduced Kuratowski's proof to his own but it is perhaps too farfetched to suggest that Kuratowski got his cue from D. Kőnig's 1916 paper. Thus, when D. Kőnig says that "not only is the theorem of Mr. Kuratowski an immediate consequence of Theorem (A) for $v = 2$ – as we have come to see – but the reasoning of Mr. Kuratowski, by which he arrived at his decomposition, is – in principle – the reasoning of my demonstration" it is difficult to agree. No decomposition gestalt is evident in D. Kőnig's original discussion and the decomposition gestalt he arrived at was certainly construed after Kuratowski's gestalt.

Kuratowski did not relate to the generalization of BDT, given by Bernstein himself, namely, that $v\mathfrak{m} = v\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$, v finite. D. Kőnig, who proved the general theorem, did suggest an extension of Kuratowski's result to cover the generalized BDT:

> If one decomposes a set E in two different fashions:
> $E = M_1 + M_2 + \ldots + M_\nu = N_1 + N_2 + \ldots + N_\nu$, $M_i \cap M_j = 0$, $N_i \cap N_j = 0$ (i, j $= 1, 2, \ldots, \nu$; $i \neq j$), and if there exists 1–1 mappings $\varphi_i(x)$ from M_1 to M_i and 1–1 mappings $\psi_i(x)$ from N_1 to N_i for $i = 1, 2, \ldots, \nu$, φ_1 and ψ_1 are the identity mapping, then the sets M_1 and N_1 can be decomposed into ν^2 disjoints parts $M_1 = \sum_\alpha M_1^{(\alpha)}$, $N_1 = \sum_\alpha N_1^{(\alpha)}$ $[\alpha = 1, \ldots, \nu^2]$[19] such that
> $N_1^{(\alpha)} = \psi_i\varphi_j(M_1^{(\alpha)})$, $(\alpha = 1, 2, \ldots, \nu^2)$ designating by $\alpha = (i\text{-}1)\nu + j$ the rank that the couple (i, j) occupies in the suite (1,1), (1,2), …, (1,ν), …, (2,ν), …, (ν,1), …, (ν, ν).

D. Kőnig provided no proof for this generalized theorem, so neither will we, but the proof seems straightforward, based on the experience gained in the case $v = 2$. D. Kőnig noted that because Kuratowski's proof was only for $v = 2$, Banach-Tarski (corollaries 12. 12′) needed to limit their division results to $v = 2^n$, where n is finite, just as he, D. Kőnig, limited in 1916 his factoring results for infinite graphs to graphs of degree 2^n. But as by 1926, in collaboration with Stephan Valkó, he was able to extend his result to any v, even for infinite graphs, Kuratowski's result could be extended as well, and thus, as D. Kőnig indicated, a similar extension of the Banach-Tarski results was enabled.

For further extension of Kuratowski's and D. Kőnig's results see Lindenbaum and Tarski 1926 §2, Tarski 1930 p 246, 248 and passim.

---

[19] For typographic reasons we have not attached the range of α to the $\sum$ sign.

Chapter 31
Early Fixed-Point CBT Proofs: Whittaker; Tarski-Knaster

On December 9, 1927, at a session of the Polish Mathematical Society, Tarski presented several results that he obtained in the wake of Banach's Partitioning Theorem (see Chap. 29). Apparently, what guided Tarski was his study of images of sets under relations and their relation to mappings of sets of sets, which continued section *37 of Principia Mathematica (see Sect. 39.6). In that same session, Knaster presented a fixed-point theorem for power-sets, which he had obtained with Tarski. The notes of Tarski and Knaster were published in 1928 (see Tarski 1928; Knaster 1928; cf. Tarski 1955 p 286 footnote 2). From the fixed-point theorem, Banach's Partitioning Theorem can be derived. The fixed-point theorem was published without proof. The Tarski-Knaster notes grew within the Polish school into a research program. Cf. Szpilrajn-Marczewski 1939 and the bibliography there and Chap. 35.

Interestingly, earlier that year, in mid-January 1927, J. M. Whittaker presented in a meeting of the Edinburgh Mathematical Society, a two pages paper that was published in the proceedings of the society. It contained a proof of the same fixed-point theorem, without noticing the fixed-point essence of the result. Whittaker used the theorem as a lemma to derive a partitioning theorem from which he derived CBT[1] directly, not through Banach's (α), (β) properties. Whittaker was 22 at the time, Tarski 26 and Knaster 34.

It was 3 years after Banach's paper (see Chap. 29) was published. Banach's name is not mentioned in Whittaker's paper and so there is no explicit evidence that Whittaker read it; but some textual evidence so suggests. Like Banach, Whittaker first proves a partitioning theorem and then states CBT. In addition, with Banach, Whittaker speaks of relations that have a property (property (M) for Whittaker, properties (α) and (β) for Banach). Other points of contact are mentioned below. On the other hand, Whittaker is using relations in his Partitioning Theorem where Banach uses mappings, and Whittaker's formulation of the Partitioning Theorem is symmetrical in the roles of A and B, while Banach had both mappings

[1] Whittaker calls the theorem the Schröder-Bernstein theorem.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_31, © Springer Basel 2013

going from A to B. Most importantly, however, is that Whittaker's context is different from Banach's: Whittaker shifts the discussion on the Partitioning Theorem and CBT from a context that involves the members of the sets concerned, which is the context in which Banach operates, to that of their subsets.

Dr. Roman Mańka, who was Knaster's last Ph.D. student, told me in correspondence that Knaster and Tarski were not aware at the time of Whittaker's paper.[2] Thus, we have here a case of simultaneous discovery.

31.1 Whittaker's Partitioning Theorem and CBT

Whittaker begins his paper with the following statement:

> Let R be a (1–1) relation between the members of two similar classes A, B_1.

Whittaker is using here Russell's terminology of 'class' and 'similar' for Cantor's 'set' and 'equivalence'. Later in the paper Whittaker uses another of Russell's terms, 'ordinally similar', for Cantor's 'similar'. Whittaker is using B_1, and not B, for a generic sign for a class because of the notation he is using in what follows. We will use Whittaker's terms.

Strangely, Whittaker presents in the quoted passage the similarity of the two classes as a feature independent from the relation R, while it is obvious that given such R the classes are similar. This, we believe, is a slip of pen, because the extra assumption is not used in the paper.

From R Whittaker defines a relation between the subclasses of A and B_1:

> It [R] correlates the members of a subclass X of A to the members of a certain subclass Y of B_1 and thus defines a relation ρ connecting X and Y.

The ρ so defined is a relation between the members of the power-classes of A and B_1. Whittaker notes that ρ is also 1–1 and that it has the following property (M):

> If $X_1 \rho Y_1$, $X_2 \rho Y_2$, then $X_1 \subset X_2$ implies $Y_1 \subset Y_2$.

Whittaker's $\subset$ is $\subseteq$ because of the 1–1 property.[3] The Partitioning Theorem Whittaker states as follows:

> If $A \rho B_1 \subset B$, $B \sigma A_1 \subset A$, there are subclasses A_0, B_0 of A, B such that $A_0 \rho B_0$, $B\text{-}B_0 \sigma A\text{-}A_0$.

Whittaker says nothing to introduce σ so we naturally conclude that there is a 1–1 relation S between the members of B and A_1 from which σ is generated just as ρ was, to be 1–1 with property (M).

---

[2] Apparently, even Fraenkel who was a follower of the history of set theory, learned of Whittaker's proof only in 1943. It was when he met Whittaker who was during World War II a British Army officer stationed in Jerusalem, where Fraenkel lived. The anecdote is told in Fraenkel 1953 (p 103 footnote).

[3] For the same reason, $Y_1 \subset Y_2$ implies $X_1 \subset X_2$.

Whittaker next introduces a metaphor: For every $X \subset A$, if $X\rho Y$, he denotes by X' the subclass of A that corresponds by σ to B-Y. Therewith Whittaker defines a mapping $'$ from P(A) to itself. If X is such that X and X' overlap, Whittaker says that X is a U; otherwise X is an L. Apparently L denotes 'lower' and 'U' upper (and M perhaps monotonous). Whittaker is not using U, L as adjectives, for he does not say that X is L but an L, so he uses them as classes, which is what they actually are: subclasses of the power-class of A. U and L provide, in fact, a new gestalt. In the proof below Whittaker uses L also as a variable ranging over the class L, namely, as a metaphor.

Note that when X is an L, $X \subseteq A\text{-}X'$ and this is an equivalent characterization of the classes in L. Whittaker does not mention this observation but uses it in the proof below.

Whittaker provides next the plan of the proof. The subclass A_0 of the theorem will be a class that fulfills $A_0' = A\text{-}A_0$, namely, a class such that its complement is its image under$'$, which is the image under σ of the complement of its image under ρ. A_0 "is to be an L but as nearly as possible a U", says Whittaker, "a largest L". The division of the power-class of A into U and L classes Whittaker describes, metaphorically for sure, as a 'Dedekind section' [cut]. It seems that Whittaker proof-processed his characterization of A_0 from the proof of Banach where A_0 is closed under the mappings and so the image of the complement of its image is disjoint from it.

As first step for the proof of the Partitioning Theorem, Whittaker states without proof the following lemma as an immediate consequence of (M):

If $X_1 \subset X_2$ then $X_2' \subset X_1'$.

Whittaker does not detail the proof of the lemma, no doubt, because it is simple: By (M) for ρ: $Y_1 \subset Y_2$; therefore $B-Y_2 \subset B-Y_1$; by (M) for σ: $X_2' \subset X_1'$. The proof of the Partitioning Theorem now unfolds as follows:

> Let A_0 = sum of all L's. Then in the first place (1) $A_0 \subset A\text{-}A_0'$. For by the lemma, $A_0' \subset L'$ for every L and so $L \subset A\text{-}L' \subset A\text{-}A_0'$. Thus $A\text{-}A_0'$ contains every L and so it contains A_0. By (1) and the lemma $(A\text{-}A_0')(A\text{-}A_0')' \subset (A\text{-}A_0')A_0' = 0$,[4] i.e., $A\text{-}A_0'$ is an L and so is contained in A_0. But by (1) A_0 is contained in $A\text{-}A_0'$. Thus $A_0 = A\text{-}A_0'$ or $A_0' = A\text{-}A_0$ which is the result stated.

For the proof of (1), note that $L \subset A\text{-}L'$ by the equivalent characterization of the L's mentioned above; by the definition of A_0, $L \subset A_0$ and thus $A_0' \subset L'$ and therefore $A\text{-}L' \subset A\text{-}A_0'$. Now, because of the equivalent characterization of L, $L \subset A\text{-}L'$ so that $L \subset A\text{-}A_0'$ and therefore $A_0 \subset A\text{-}A_0'$ which is (1). Note that the equivalent characterization of the L's cannot provide (1) directly because we don't know yet that A_0 is an L. It is (1) that provides that A_0 is an L! Note further that as A_0 is an L, its definition as the union of all L's is impredicative. This is no surprise, for Whittaker's proof avoids the notions 'number' and 'complete induction', and in such cases, impredicativity is invoked. Whittaker's proof can thus be classified with the proofs of Dedekind, Zermelo and Peano.

[4] Whittaker denotes intersection by juxtaposition and the empty set by 0.

For the proof that A-A_0' is an L, note that since by (1) and the lemma $A_0 \subset$ A-A_0', (A-A_0')$' \subset A_0'$. As clearly (A-A_0')$A_0' = 0$, also (A-A_0')(A-A_0')$' = 0$ so by the definition of L classes, A-A_0' is an L.

Whittaker points out in a footnote that "there may be no L's but this does not matter since the null class is counted as a subclass of A". Indeed, if A_0 is empty and it is taken by ρ to B_0, then (B-B_0)σA. But (B-B_0) $\subset$ B$\sigma A_1 \subset$ A contrary to property (M), unless A_1 = A. So in this case BσA.

Whittaker's use of the sum [union] operation for the definition of A_0 appears as the dual to Zermelo's use of intersection in his 1908b CBT proof. The metaphor of taking the union of a set of sets that have a good property but not good enough, Whittaker may have proof-processed from Zermelo's first proof of the Well-Ordering Theorem.

Following the proof, Whittaker stated CBT as an immediate corollary:

> If A is similar to a part of B and B is similar to a part of A, then A is similar to B.[5]

Indeed the proof is simple: If the similarity of A to B_1 is given by R and the similarity of B and A_1 is given by S, then we have the conditions of the Partitioning Theorem and A_0 is provided. So the similarity of A and B is given by combining R reduced to A_0, which relates A_0 to B_0, and S^{-1} reduced to A-A_0, which relates A-A_0 to B-B_0. Because the correlated classes are disjoint, the combined relation is 1–1. Here it is used that A_0 is both disjoint from A_0' and its complement. The possibility of reducing a similarity to a subset of its domain and the additive nature of similarity, are Banach's properties (α) and (β).

When the relations R, S behind ρ, σ of the Partitioning Theorem, have more properties than just being 1–1, the combined relation defined as the similarity of A and B in CBT may not have the additional properties. Even though the Partitioning Theorem can still be applied. Whittaker gives the following example with relations of ordinal similarity. Let A be the class [0,1] and the rational numbers of [1,2] and B the class [0,2]. Then A is a subclass of B and so is ordinally similar to a subclass of B by the identity relation. B is also ordinally similar to a subclass of A, actually [0,1], by the relation $y = x/2$.[6] A partitioning of A, B as warranted by the Partitioning Theorem, is given by setting A_0, B_0 to be both equal to the set of all rational numbers of A (namely, the rational numbers of [0,2]).[7] Then A_0 is ordinally similar to B_0 by the identity and B-B_0 is ordinally similar to A-A_0 by the relation $y = x/2$. Combining these two relation as per the proof of CBT does not provide an ordinal similarity of A and B, which indeed are not

[5] In a footnote Whittaker also gives the version in the language of cardinal numbers using the $\leq$ relation. He is not aware of the problem in this formulation (see Sect. 17.2).

[6] Whittaker wrote $y = 2x$ which expresses the relation from [0,1] to B.

[7] A subclass X of A that contains an irrational number ζ will never be disjoint from X$'$. The reason is that there is a maximal n such that $n\zeta$ belongs to X and if Y corresponds to X by ρ, then $2n\zeta$ belongs to B-Y and thus ζ belongs to X$'$. Thus the suggested A_0, B_0 are indeed the classes generated by the proof of the Partitioning Theorem, a point that Whittaker had omitted.

ordinally similar, but only "ordinal similarity by finite decomposition", according to the notion introduced by Banach-Tarski in 1924 in the context of their "paradox". Banach mentioned that the Partitioning Theorem is applicable to the relation of ordinal similarity between two ordered-sets because it possesses property (α) so the partitions are ordinally similar, but he did not mention this relation as possessing property (*β*) and Whittaker showed that it does not. This relatedness of Whittaker to Banach is further evidence in favor of the claim that Whittaker read Banach's paper.

Whittaker, in a footnote to the proof, points out that the proof of the Partitioning Theorem does not assume R, S but only the two relations *ρ*, *σ*, which have to be 1–1 and have property (M). And he adds:

> It is not necessary that the members of X should be in (1-1) relation with those of Y, nor that those subclasses of B to which the subclasses of A are correlated by ρ should be all the subclasses of a certain part B_1 of B.

Thus *ρ* has only to be a relation with domain the power-class of A, P(A), and range in P(B), with property (M); *σ* likewise a relation from P(B) into P(A) with property (M). In Banach's proof, 1–1 mappings are used instead of the relations R, S, and even when the discussion was switched, in a structuralist fashion, to relations between sets, the relations were supposed to imply the existence of a 1–1 mapping; therefore property (M) was not stated by Banach because it is an innate property for relations obtained from 1–1 mapping.

31.2 The Tarski-Knaster Fixed-Point Theorem

Some of the results of which Tarski reported in 1927 translated Banach's theorem from mappings to relations which, Tarski showed, need not be 1–1. Knaster then retranslated one of Tarski's reformulations of Banach's theorem back to the language of mappings but required monotonicity instead of 1–1 (we cite the theorem in Sect. 35.4). He then stated the following fixed-point theorem:

> h(x) being a monotone function of sets and A a set such that $h(A) \subseteq A$, there exists a subset D of A such that $D = h(D)$.

From the fixed-point theorem Knaster derived Banach's Partitioning Theorem by setting $h(X) = A\text{-}g(B\text{-}f(X))$, namely, to X corresponds the complement of the image by *g* of the complement of the image of X by *f*.[8] *h* fulfills the condition of the theorem and so there exists a fixed-point for it which provides one of the partitions of Banach's theorem from which the others can be defined. Knaster did not detail in his note how the fixed-point theorem is proved. Clearly, Whittaker's A_0 is D. If one

[8] Mańka-Wojciechowska (1984 p 196f) note that this is how the theorem is proved in Sierpiński 1951 and they further note that the Tarski-Knaster Fixed-Point Theorem can also serve to prove the Cantor-Bendixson Theorem.

does not care to avoid the natural numbers then the union of the set $\{h^n(\varnothing)\}$ for $n \geq 0$ is also a fixed-point (see Sect. 35.5).

Whittaker probably missed the significance of the lemma to which Tarski-Knaster were conscious: that the lemma provides a topological fixed-point theorem for power-classes that could be generalized (cf. Drake 1974 p 49).

Chapter 32
CBT and BDT for Order-Types

Not all results of cardinal arithmetic can be transported to order-type arithmetic, or even to ordinal arithmetic. For example, Sierpiński (1947b p 74) noted that the Mean-Value Theorem (see Sect. 25.2) cannot be extended to ordinals instead of cardinals: If M is a set of type $\omega + 1$ and P a subset of M containing only the last element of M, so obviously P is of type 1, there is no subset of M that has type ω and contains P. Another example is a set A of order-type $\omega + \eta + \omega^*$ and a set B of order-type η; each is ordinally similar to a subset of the other (the conditions of CBT) but the two are not ordinally similar.

Still, with certain restrictions, results analogous to CBT and BDT can be expressed for order-types. This was first noted in the Lindenbaum-Tarski paper of 1926,[1] several theorems from which (§3) we will review here. Since the paper contains only theorems without proofs,[2] we rely, for proofs of the theorems discussed, on papers by Sierpiński, mainly his paper from 1948.[3]

32.1 Comparability Theorems

Among the results of Lindenbaum-Tarski 1926 §3 that are analogous with CBT, first is the following (small Greek letters signify order-types):

Theorem 3. *If* $\alpha = \sigma + \beta$ *and* $\beta = \alpha + \rho$ *then* $\alpha = \beta$.

In the language of sets and mappings, instead the language of order-types, Theorem 3 is stated thus:

[1] Where CBT is still called 'the Schröder-Bernstein theorem' – p 302.

[2] The theorems are attributed to either Tarski or Lindenbaum. The theorems described in this chapter are attributed to Lindenbaum, unless otherwise stated.

[3] In the late 1940s, after Lindenbaum was murdered in Ponary-Tarski obtained shelter in the USA, Sierpiński attempted to supply systematically proofs of theorems in Lindenbaum-Tarski 1926. Lindenbaum was Sierpiński's student.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_32,

Theorem 4. When an ordered-set A is similar to a segment of an ordered-set B, and B is similar to a residue of A, then the sets A and B are similar.

Lindenbaum-Tarski reference Hausdorff 1914a p 88 for definition of 'segment' (*Anfangsstück*) and 'residue' (*Endstück*): A segment (residue) contains with every element also all smaller (greater) elements.

Lindenbaum-Tarski indicate that Theorem 3 is an immediate corollary of:

Theorem 2. $\alpha = \sigma + \alpha + \rho$ *iff* $\alpha = \sigma + \alpha$ *and* $\alpha = \alpha + \rho$.

Proof of Theorem 3: From the hypothesis of Theorem 3 it follows by substitution that $\alpha = \sigma + \alpha + \rho$ Hence, by Theorem 2, $\alpha = \alpha + \rho = \beta$.

Proof of Theorem 2: The right to left implication is immediate by substitution For the other side of the proof the following is necessary (as indicated by Lindenbaum-Tarski):

Theorem 1. $\alpha = \sigma + \alpha + \rho$ *iff there is an order-type* ξ *such that* $\alpha = \sigma\omega + \xi + \rho\omega^*$.

Proof of Theorem 2, continued: If $\alpha = \sigma + \alpha + \rho$, then by Theorem 1 there is an order-type ξ such that $\alpha = \sigma\omega + \xi + \rho\omega^*$; but then, since (*) $\omega = 1 + \omega$ and $\omega^* = \omega^* + 1$, $\alpha = \sigma(1 + \omega) + \xi + \rho\omega^* = \sigma + \sigma\omega + \xi + \rho\omega^* = \sigma + \alpha$ and $\alpha = \sigma\omega + \xi + \rho\omega^* = \sigma\omega + \xi + \rho(\omega^* + 1) = \sigma\omega + \xi + \rho\omega^* + \rho = \alpha + \rho$

This proof of Theorem 2 from Theorem 1 is given in Sierpiński 1948 (p 5 Corollary 1).

Proof of Theorem 1: The right to left implication in Theorem 1 is obtained by the argument used in the second part of the proof of Theorem 2:

$\alpha = \sigma\omega + \xi + \rho\omega^*$ entails

$\alpha = \sigma(1 + \omega) + \xi + \rho(\omega^* + 1) = \sigma + \sigma\omega + \xi = \rho\omega^* + \rho = \sigma + \alpha + \rho$. The left to right implication of Theorem 1 is argued thus: If $\alpha = \sigma + \alpha + \rho$, let A be a set of type α. A then has a segment S_1 of type σ and a residue R_1 of type ρ with the middle section A_1 of type α again. By complete induction we can define the sequences S_n, A_n, R_n. Let $S = \cup S_n$, $R = \cup R_n$. Then S is a segment of A of order-type $\sigma\omega$ and R is a residue of A of type $\rho\omega^*$. Let A′ be the section between S and R. A′ maybe empty. Let ξ be the order-type of A′. Then clearly $\alpha = \sigma\omega + \xi + \rho\omega^*$. This proof is given in Sierpiński 1948 p 4 Lemma 1. Theorem 1 is analogous to Zermelo's 1901 Denumerable Addition Theorem (see Chap. 13) from which CBT can be derived in the language of cardinal numbers. Similar gestalt and metaphor are used in both proofs: A set is partitioned (gestalt) with one partition being equivalent (similar) to the set so that the partitioning can be repeated (metaphor). The reemergence metaphor is modestly applied when using the equalities which we denoted by (*). It seems that we have here an example of proof-processing.

It is noted in Lindenbaum-Tarski that in general $\alpha = \sigma + \beta$ and $\beta = \rho + \alpha$ do not entail $\alpha = \beta$. For example, if $\alpha = \omega^*\omega$, $\beta = 1 + \omega^*\omega$, $\sigma = \omega^*$ and $\rho = 1$, then $\alpha = \sigma + \beta$ and $\beta = \rho + \alpha$ but $\alpha \neq \beta$. Yet the following, another analog of CBT, holds:

Theorem 5. *If $\alpha = \sigma + \beta$ and $\beta = \rho + \alpha$, then $\alpha = \beta$ iff $(\sigma + \rho)\omega = (\rho + \sigma)\omega$. This theorem is not mentioned in Sierpiński* 1948.

Proof of Theorem 5: From $\alpha = \sigma + \beta$ and $\beta = \rho + \alpha$ we get $\alpha = (\sigma + \rho) + \alpha$ and $\beta = (\rho + \sigma) + \beta$ Thus, by the procedure employed in the left to right part of the proof of Theorem 1, $\alpha = (\sigma + \rho)\omega + \xi$ and $\beta = (\rho + \sigma)\omega + \xi'$. Note further that $\sigma + (\rho + \sigma)\omega = (\sigma + \rho)\omega$ and $\rho + (\sigma + \rho)\omega = (\rho + \sigma)\omega$.[4]

For the right to left implication we have:

$$\alpha = \sigma + \beta = \sigma + (\rho + \sigma)\omega + \xi' = (\sigma + \rho)\omega + \xi' = (\rho + \sigma)\omega + \xi' = \beta.$$

For the left to right implication: If $\alpha = \beta$ then we have $(\sigma + \rho)\omega + \xi = (\rho + \sigma)\omega + \xi'$. Let M be an ordered-set of type α, M′ an ordered-set of type β, φ a similarity mapping from M onto M′. Let M^* be the segment of M of order-type $(\sigma + \rho)\omega$ and M'^* be the segment of M′ of order-type $(\rho + \sigma)\omega$. Now there are three possibilities: either $\varphi(M^*) = M'^*$ or $\varphi(M^*)$ is a segment of M'^* and the residue M'-M'^* is of order-type ζ, or $\varphi^{-1}(M'^*)$ is a segment of M_* and the residue M-M^* has order-type ζ'. In the first case M^* and M'^* are directly similar and the result is obtained. In the second case we have $(\rho + \sigma)\omega = (\sigma + \rho)\omega + \zeta = \sigma + (\rho + \sigma)\omega + \zeta$ and by Theorem 2 $(\rho + \sigma)\omega = \sigma + (s\rho + \sigma)\omega = (\sigma + \rho)\omega$, and the required result follows. The third case is symmetrical to the second so we can skip a direct argument.

When in Theorem 5 we have $\rho = \sigma$ the condition $(\sigma + \rho)\omega = (\rho + \sigma)\omega$ is fulfilled and so we get that if $\alpha = \sigma + \beta$ and $\beta = \sigma + \alpha$, then $\alpha = \beta$, which is Theorem 6 in Lindenbaum-Tarski. Lindenbaum-Tarski note that dual versions of Theorems 5, 6 can be obtained, replacing $\sigma + \beta$, $\rho + \alpha$ and ω, respectively, by $\beta + \sigma$, $\alpha + \rho$, ω^*.

As a consequence of Theorem 3 (order-types CBT), Lindenbaum-Tarski present the following Theorem 7: If $\alpha = \sigma + \alpha^*$ or $\alpha = \alpha^* + \rho$, then $\alpha = \alpha^*$. Indeed, in the first case clearly $\alpha^* = \alpha + \sigma^*$ and in the second case $\alpha^* = \rho^* + \alpha$, so in both cases the conditions of Theorem 3 hold and we get $\alpha = \alpha^*$.

Paradigm examples of order-types equal to their inverse, are: $\omega^* + \omega$, $\omega + \omega^*$, $\omega + 1 + \omega^*$. By symmetry considerations, this is the general rule, as is stated in Theorem 8: $\alpha = \alpha^*$ iff there is an order-type ξ such that $\alpha = \xi^* + \xi$ or $\alpha = \xi^* + 1 + \xi$. An "analytic" proof of this theorem, which we will not attempt to provide, seems to require the use of AC.

32.2 Division Theorems

Also a theorem analogous to BDT ($k\mathfrak{m} = k\mathfrak{n} \rightarrow \mathfrak{m} = \mathfrak{n}$, $\mathfrak{m}$, $\mathfrak{n}$ cardinal numbers, k natural number) was given in Lindenbaum-Tarski §3:

[4] Something similar was used in Schröder 1898, in his proof of CBT p 339 Eq. 54.

Theorem 15. If μ or μ^* is an ordinal number $\neq 0$, then $\mu\alpha = \mu\beta$ entails $\alpha = \beta$.

Since $\mu\alpha = \mu\beta$ iff $\mu^*\alpha^* = \mu^*\beta^*$ and $\alpha = \beta$ iff $\alpha^* = \beta^*$, it is enough to prove the theorem when μ is an ordinal number. Under this supposition, the theorem is proved in Sierpiński 1948 (Theorem 1 p 1) as follows:

Let M be the set of all ordinal numbers ξ, $0 \leq \xi < \mu$; M is well-ordered and the ordinal number of M is μ. Let A and B be two disjoint ordered-sets with order-types α, β. Let P be the set of all ordered-pairs (m, a) where $m \in$ M and $a \in$ A, ordered by reverse lexicographic order, and Q the set of ordered-pairs (m, b), $m \in$ M, $b \in$ B, ordered likewise in reverse lexicographic order. Let us call the set of all pairs (m, a) for some a, the slice of a and denote it by P_a; likewise call the set of all pairs (m, b) for some b, the slice of b, denoted by Q_b. The order-types of P and Q are $\mu\alpha$ and $\mu\beta$, and as $\mu\alpha = \mu\beta$, there is a similarity mapping f from P on Q. Define a mapping φ from A into P by $\varphi(a) = (0, a)$ and a mapping ψ from Q onto B by $\psi((m, b)) = b$. It will be proved that $g(a) = \psi f \varphi(a)$ is a similarity mapping of A onto B, thus proving that $\alpha = \beta$.

Obviously, g is a mapping from A into B. We need to prove that g preserves order (and therefore it is 1–1) and that it is on B. Let then a, a' be two members of A such that $a < a'$. Then $(0, a) < (0, a')$.

Let $f((0, a)) = (\xi, b)$, $f((0, a')) = (\eta, b')$, where ξ, $\eta \in$ M and b, $b' \in$ B. We have $(\xi, b) < (\eta, b')$ and hence $b < b'$ or $b = b'$ and $\xi < \eta$. In the first case g preserves the order. In the second case we would have that a copy of M, there is at least one such copy between $(0, a)$ and $(0, a')$, is similar to a subset of a segment of M, the segment is determined by η, and the subset is the section between ξ and η, which is a contradiction as M is well-ordered. Therefore, the second case cannot occur and g preserves order and is 1–1. We now prove that it is on.

Assuming that there is a $b_0 \in$ B such that for no $a \in$ A, $g(a) = b_0$, we will reach a contradiction by constructing an infinite descending sequence of ordinals.

Let $f^1((0, b_0)) = (\alpha_0, a_0)$. Necessarily, $0 < \alpha_0$, for otherwise $g(a_0) = b_0$, contrary to the assumption on b_0. The set $\{(\xi, a_0) \mid \alpha_0 < \xi < \mu\}$, the α_0 residue of the slice P_{a_0}, is similar to the set $\{(\eta, b_0) \mid \eta < \mu\}$, the slice Q_{b_0}, for otherwise one of the following two cases will happen: (1) There is a ξ' such that for all $\eta < \mu, f^1((\eta, b_0)) = (\xi, a_0)$, with $\alpha_0 < \xi < \xi'$. In this case M will be similar to a subset of its segment defined by ξ', a contradiction to the well-ordering of M. (2) There is $\eta < \mu$ such that $f^1((\eta, b_0)) = (0, a)$, with $a > a_0$. But then $g(a) = b_0$ contrary to the assumption on b_0. As the slice of a_0 is of ordinal μ and it is composed of two parts the first of ordinal α_0 and the second of ordinal μ, we obtain $\mu = \alpha_0 + \mu$.

Let $f((0, a_0)) = (\beta_1, b_1)$. Since $(0, a_0) < (\alpha_0, a_0)$ we have $(\beta_1, b_1) < (0, b_0)$. But then we must have $b_1 < b_0$. If there is $b_1 < b < b_0$ then the set of all (η, b), which is similar to M, would be similar to a subset of the set of all (ξ, a_0), $\xi < \alpha_0$, which is similar to a segment of M, a contradiction to the assumption that M is well-ordered. Therefore, b_1 must be the predecessor of b_0 in B. A similar contradiction is obtained if we assume that $\beta_1 = 0$. So now we have that the residue of Q_{b_1}, defined by β_1, corresponds to the segment of P_{a_0} defined by α_0. So the slice Q_{b_1}, which has ordinal μ, is composed of two parts the first of ordinal β_1 and the second of ordinal α_0, so that $\mu = \beta_1 + \alpha_0$.

Let $f^1((0, b_1)) = (\alpha_1, a_1)$. Since $(0, b_1) < (\beta_1, b_1)$ we have $(\alpha_1, a_1) < (0, a_0)$. Then necessarily $a_1 < a_0$. If there is an a such that $a_1 < a < a_0$, then the slice of a would be similar to a subset of the β_1 segment of the slice of b_1, and we would get the same contradiction as before. Therefore, a_1 must be the predecessor of a_0. Again, α_1 cannot be 0 because otherwise we would get that the slice of a_1, which is of ordinal μ, is similar to the β_1 segment of the slice of b_1, the ordinal of which is necessarily smaller than μ. So the slice of a_1 is composed of two parts, the first of ordinal α_1 and the second of ordinal β_1, so $\mu = \alpha_1 + \beta_1$. As $\beta_1 < \mu$, $\alpha_0 + \mu$ has a segment of ordinal $\alpha_0 + \beta_1$ so $\alpha_0 + \beta_1 < \mu$ and we must have $\alpha_0 < \alpha_1$.

Let $f((0, a_1)) = (\beta_2, b_2)$. Since $(0, a_1) < (\alpha_1, a_1)$ we have $(\beta_2, b_2) < (0, b_1)$. Then necessarily $b_2 < b_1$. Again, we find that b_2 must be the predecessor of b_1 and that β_2 cannot be 0. So we get $\mu = \beta_2 + \alpha_1$ and as $\mu = \beta_1 + \alpha_0$ and $\alpha_1 > \alpha_0$, we have $\beta_2 < \beta_1$.

Continuing in this fashion to define α_n, β_n, a_n, b_n. The α_n and β_n are never 0 so the definition can be continued, and each a_{n+1} (b_{n+1}) is the immediate predecessor of a_n (b_n). As the β_n form a descending sequence of ordinals, we reach a contradiction, which implies that g is on.

The metaphor that won the proof lies in the definition of the mappings φ, ψ, g. For the part of the proof that concerns that g is on, the gestalt is that if two slices do not match then the mismatch must begin earlier, down along a ω^* part of α and β. At each step downward, the new slice is divided (on the B side) at a smaller ordinal. The argument is of the back-and-forth type, with careful calculations (the metaphor) regarding the sequences of ordinals encountered. The proof does not resemble the proofs of BDT for sets. It does associate with the generation of frames in CBT proofs. This is interesting because we see that gestalt and metaphor, which were devised for a positive proof, can be used also in a negative proof.

Lindenbaum-Tarski also brought a right division theorem (p 321):

Theorem 13. *If $\alpha n = \beta n$, n natural number not 0, then $\alpha = \beta$.*

Sierpiński (1948 p 3 Theorem 2) gave a proof of the theorem. After establishing Theorems 1, 2 of Lindenbaum-Tarski mentioned above, Sierpiński proved the following lemma (p 5 Lemma 2):

Lemma. *If U is an ordered-set, n a natural number [not 0], and $U = A_1 + \ldots + A_n = B_1 + \ldots + B_n$, where each element of A_k (respectively B_k) precedes in U the elements of each A_l (respectively B_l) for $1 \leq k < l \leq n$,[5] then there exists an index p, $1 \leq p \leq n$, such that $B_p \subseteq A_p$.*

Proof of the lemma: If B_1 is not contained in A_1 then $B_2 \subseteq A_2 + \ldots + A_n$ for otherwise there would be a $b \in B_2$ in A_1 and all the members of B_1, necessarily smaller than b, will belong to A_1, contrary to the assumption Continuing in this way we get that $B_n \subseteq A_n$, contrary to the accumulated assumptions.

[5] Tacitly it is assumed here that the A_k (B_k) are pairwise disjoint.

As a consequence of the lemma Sierpiński proved the following corollary (p 6 Corollary 2):

Corollary. *If* $\alpha n = \beta n$ *then there are two order-types* σ, ρ *such that* $\alpha = \sigma + \beta + \rho$.

Proof of corollary: From the hypothesis we can devise a set U partitioned into n partitions of order-type α and n partitions of order-type β By the lemma there is an index $1 \leq p \leq n$ such that $B_p \subseteq A_p$. Then A_p, which is of order-type α, is partitioned to three partitions: S of all elements that are smaller than all elements of B_p, R that consists of all elements that are larger than all elements of B_p, and the elements of B_p. If σ is the order-type of S and ρ that of R, the corollary is obtained.

Note that in the lemma and corollary, each A_k (B_k) contains all elements of U that are between any two of its elements. Note further that the lemma and corollary are symmetric so that we also have $\beta = \sigma' + \alpha + \rho'$.

Proof of Theorem 13: Let U be a set as in the lemma Then U has both segments and residues of both α and β.[6] If either the segments or residues coincide, then $\alpha = \beta$. Otherwise, we have $\alpha = \beta + \rho$ or $\beta = \alpha + \rho$ and $\alpha = \sigma + \beta$ or $\beta = \sigma + \alpha$. If the crossed cases occur then we have by Lindenbaum-Tarski Theorem 3 that $\alpha = \beta$. Otherwise, we have that $\alpha = \beta + \rho$ and $\alpha = \sigma + \beta$ or $\beta = \alpha + \rho$ and $\beta = \sigma + \alpha$. Also by the corollary we have order-types σ', σ'' and ρ', ρ'' such that $\beta = \sigma' + \alpha + \rho'$ and $\alpha = \sigma'' + \beta + \rho''$. So in the first case we have:
$\alpha = \sigma + \beta = \sigma + \sigma' + \alpha + \rho'$ and by Lindenbaum-Tarski Theorem 2, $\alpha = \alpha + \rho'$. On the other hand we have in the first case that $\beta = \sigma' + \alpha + \rho' = \sigma' + \beta + \rho + \rho'$ and by Theorem 2 again $\beta = \beta + \rho + \rho'$. But $\alpha = \beta + \rho$ so that we have
$\beta = \alpha + \rho' = \alpha$. The remaining case is symmetrically proved.

Sierpiński continued to demonstrate that Theorem 13 remains valid when n is replaced by a successor ordinal but fails for limit ordinals. He then brought over 10 theorems on the divisibility of order-types in various special cases. Most of the proofs, like that of Theorem 13, are of arithmetic nature.

For other results on order-types arithmetic, see Morel 1959, Ginsburg 1955 and the bibliography there.

[6] This is the gestalt of the proof. Each of the cases that emerge is handled by arithmetic metaphors obtained on formulas (gestalt) derived by substitution, leveraging on Theorems 3, 2 of Lindenbaum-Tarski and the above corollary. This remark is intended to stress that the gestalt and metaphor in a proof need not be limited to the context of the main entities (here partitioned sets) and can extend to linguistic entities (formulas).

Chapter 33
Sikorski's Proof of CBT for Boolean Algebras

In a paper of 1948 (its main result was presented in 1946 at a session of the Polish mathematical society), Roman Sikorski ported Banach's (1924) Partitioning Theorem and CBT to Boolean algebras. This was a developing field within structuralist mathematics since the 1930s, and Sikorski was one of its main propagators. Banach's Partitioning Theorem was already dressed for structuralist application (see Chap. 29) and the move towards Boolean algebras was perhaps influenced by the work, in 1927, of Whittaker, Tarski and Knaster (see Chap. 31) who discussed CBT within power-sets. After Sikorski, as the movement towards mathematical structuralism boomed, it became standard to try CBT for the various mathematical structures discussed. The property sought was whether, if the given mappings of CBT preserve the structure's properties, so does the emerging mapping.

For Sikorski a Boolean algebra is a non-empty set **A**, closed under two operations: addition (+) and complementation (′), satisfying Huntington's (1933) axioms: Addition is commutative and associative and the following is satisfied for any two members A, B of **A**: $(A' + B)' + (A' + B')' = A$. Sikorski defines multiplication (intersection), denoted by juxtaposition or ·, by $AB = (A' + B')'$, difference, denoted by -, by $A\text{-}B = AB'$ and a partial order, denoted $\subset$ and called inclusion, by $A \subset B$ when $A + B = B$. It is easy to see that by the commutativity of addition, $A \subset B$ and $B \subset A$ entail $A = B$. The paradigm for Boolean algebras is the power-set of a set; Sikorski denotes by **S**(X) the Boolean algebra of the power-set of X. A Boolean algebra is σ-complete if every denumerable sequence of its members has a union – a minimal member that contains (i.e., that is greater than) all members of the sequence.

Two Boolean algebras are isomorphic if there exists a 1–1 mapping that correlates their members and preserves inclusion.[1] It is a standard exercise in Boolean algebra to prove that isomorphisms preserve all operations, including denumerable union if the algebras are σ-complete. A σ-homomorphism between σ-complete algebras is a

[1] Compare Whittaker 1927 property (M).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_33,

mapping which is not necessarily 1–1 or on and which preserves complementation and countable (finite or denumerable) unions. Since all other operations and the inclusion relation are defined using the two basic ones, a σ-homomorphism preserves all operations. An isomorphism is thus a σ-homomorphism.

If **A** is a Boolean algebra and E one of its elements, then the set of all A ∈ **A** such that A ⊂ E, is a Boolean algebra, which Sikorski denotes by E**A**. In it, complementation is made relative to E and is denoted by A'_E. Thus E**A** is composed of all EA for A ∈ **A**. E**A** is σ-complete if so is **A**. If X ⊂ $\mathfrak{X}$ then XS($\mathfrak{X}$) = S(X).

33.1 Theorems and Proofs

Sikorski begins with a statement of Banach's theorem (Theorem I) and of CBT (Theorem II) claiming that these theorems "are particular cases of two more general theorems from the theory of σ-complete Boolean Algebras" that he will demonstrate. We will return to this contention below. He then gives three "obvious" lemmas:

> (A) Let **A**, **B**, **C** be σ-complete Boolean algebras, B ∈ **B** and C ∈ **C**. If f and g are σ-homomorphisms of **A** in B**B** and of **B**[2] in C**C** respectively, then the mapping gf is a σ-homomorphism of **A** in g(B)**C**.[3]
> (B) Let **A** and **B** be two Boolean algebras, C ∈ **A** and D ∈ **B**. If C**A** is an isomorph of D′**B**, and D**B** is one of C′**A**, then **A** and **B** are isomorphic.
> Namely, if f is an isomorphism of C**A** on D′**B** and g is one of C′**A** on D**B**, then
> h(A) = f(AC) + g(AC′) (for A ∈ **A**) is an isomorphism of **A** on **B**.[4]
> (C) If **A** and **B** are two Boolean algebras, and if f is an isomorphism of **A** on **B**, then A**A** is isomorphic to f(A)**B** for any A ∈ **A**.
> In fact, the mapping f(E) restricted to E ⊂ A[5] is an isomorphism of A**A** on f(A)**B**.

Next Sikorski provides the Partitioning Theorem for Boolean algebras:

> Theorem 1. Let **A** and **B** be two σ-complete Boolean algebras, A ∈ **A** and B ∈ **B**, and let f and g be two σ-homomorphisms of **A** in B**B**, and of **B**[6] in A**A** respectively. Then there exist two elements C ∈ **A** and D ∈ **B** such that f(C) = D′ and g(D) = C′.
> [*Proof*:] Let A_1 = gf(A′) and by induction A_{n+1} = gf(A_n) (n = 1, 2, 3, ...), and let
> C = A′ + A_1 + A_2 + A_3 + ... and D = (f(C))′. Obviously A_n ⊂ A, C ∈ **A**, D ∈ **B** and f(C) = D′. On account of (A), gf is a σ-homomorphism; therefore
> gf(C) = gf(A′) + gf(A_1) + gf(A_2) + ... = A_1 + A_2 + A_3 + ... = AC and consequently
> g(D) = g[(f(C))′] =[7] A-gf(C) = A-AC[8] = AC′ = C′, since C′ ⊂ A by definition of C.

[2] In the original B is printed here instead of **B**.

[3] For every member A of **A**, *f*(A) ⊂ B and so *gf*(A) ⊂ *g*(B) so *gf*(A) ∈ C**C**.

[4] D′ can be denoted by D and D by D′. Why the "crossed" notation is preferred we don't know. That *h* is 1–1 and on follows from the 1–1 nature of the mapping A → {AC,AC′}.

[5] The wording here is perhaps not so clear. It is f that is restricted to A.

[6] In the original B is printed here instead of **B**.

[7] Actually, *g*[(*f*(C))′] = *g*[B-*f*(C)] = *g*(B)-*gf*(C) = A-*gf*(C).

[8] A-AC=A(AC)′ = A(A′ + C′) = AA′ + AC′ = AC′.

Following Theorem 1, Sikorski proves the original Banach's Partitioning Theorem by moving from its given situation, which relates each of two sets by a 1–1 mapping to a subset of the other, to the Boolean algebra of their power-sets.[9] The given assumptions then provide the assumptions of Theorem 1, and by its thesis, Banach's Partitioning Theorem is proved. Sikorski says of his proof of Banach's theorem that it is obtained by "suitable modification" of Banach's proof. It seems more accurate to say that what Sikorski did was to translate Banach's proof first to the context of the power-sets of the sets given in the Partitioning Theorem; then he switched Banach's gestalt from the J. Kőnig strings of the elements to Dedekind's chains of frames, the chain of A′ is imaged by *f* to a chain in B. Sikorski's C corresponds to Banach's A_2. Finally, he replaced Banach's language of sets and mappings with the Boolean algebra language, transforming chains to elements correlated by *gf*. For no obvious reason, Sikorski regarded Theorem I' as more general than Theorem I.

Sikorski then demonstrated that σ-completeness is a necessary condition for Theorem 1 by giving the following counterexample: Let **A** and **B** be both the Boolean algebra composed of all the finite subsets of the set N of natural numbers and their complements. Let A and B be equal to $N-\{1\}$. Let φ be the 1–1 function $x \to x + 1$, $x \in N$, and let *f*, *g* be defined by $f(X) = \varphi(X) = g(X)$, X in the mentioned Boolean algebra. *f* is clearly an isomorphism of **A** on B**B** and of **B** on A**A**. It is easy to verify, contends Sikorski, that there are no C, D that fulfill the theorem. Indeed, if C is finite and c is its maximal number then $c + 2$ is in C′ and $c + 1$ is in D′; but then $c + 1$ is not in D so $c + 2$ cannot be in C′. If C is infinite then C′ is finite and let c be its largest member. Then $c + 1$ is in C and it is in D and thus not in D′; but then $c + 2$ cannot be in C′ contrary to the assumption on c.[10] The conditions of Theorem 1 are fulfilled except the σ-completeness requirement, which is not fulfilled since the sequence of the singletons of even numbers has no union in the discussed algebras. Incidentally, this example shows that not all sub-Boolean algebras of a Boolean algebra of a power-set are power-sets. Thus it shows that not all sub-Boolean algebras of a Boolean algebra **A** are of the form A**A** for some A ∈ **A**.

Now Sikorski proves what he considers a generalization of CBT:

> Theorem 2. Let **A** and **B** be σ-complete Boolean algebras, A ∈ **A** and B ∈ **B**. If **A** is isomorphic to B**B** and **B** to A**A**, then **A** and **B** are isomorphic.
> [*Proof*:] Suppose that f and g are isomorphisms of **A** on B**B** and of **B** on A**A** respectively. Since f and g are σ-homomorphisms, then there exist by Theorem 1 two elements C ∈ **A** and D ∈ **B** such that f(C) = D′ and g(D) = C′. By Lemma (C), C**A** is isomorphic to D′**B** and D**B** is isomorphic to C′**A**. By Lemma (B), the algebras **A** and **B** are isomorphic.

---

[9] There is a trifling difference between Banach's theorem cited at the beginning of the paper as theorem I, and its restatement as theorem I′ proved via theorem 1. In the statement of theorem I′ there is a typo regarding the place of the numberings (1) and (2); the proof is not affected by this typo.

[10] The argument is not detailed in Sikorski's paper.

From Theorem 2 Sikorski derives CBT as a "particular case", by quoting the following theorem, which he attributes to Szpilrajn-Marczewski (1939 p 137): X and Y have the same power if and only if **S**(X) and **S**(Y) are isomorphic. The proof of CBT then runs as follows: If $\mathfrak{X}$, $\mathfrak{Y}$ are two sets, X $\subset$ $\mathfrak{X}$ and Y $\subset$ $\mathfrak{Y}$, and the power of $\mathfrak{X}$ is equal to the power of Y and the power of $\mathfrak{Y}$ is equal to the power of X, then by the Szpilrajn-Marczewski Theorem, **S**($\mathfrak{X}$) is isomorphic to **S**(Y) = YS($\mathfrak{Y}$) and **S**($\mathfrak{Y}$) is isomorphic to **S**(X) = XS($\mathfrak{X}$). But then, by Theorem 2, **S**($\mathfrak{X}$) is isomorphic to **S**($\mathfrak{Y}$) and by the Szpilrajn-Marczewski Theorem again $\mathfrak{X}$, $\mathfrak{Y}$ are of the same power.

Again it is not clear why Sikorski regards a proof by way of a theorem about Boolean algebras as more general than a proof about sets, when it is set theory that provides the domain of Boolean algebras and thus appears as more general. It seems that Sikorski regards Boolean algebra as a new dominant background theory (Lakatos 1976), a generalization of set theory, a view that is hardly tenable. Note that Sikorski's derivation of CBT from the Partitioning Theorem differs from the derivation of Banach not only because of its use of the Szpilrajn-Marczewski's Theorem but also because he disregarded the fact that Banach derived CBT by way of his properties (α) and (β) and not directly from the Partitioning Theorem.

As for the Szpilrajn-Marczewski Theorem referenced by Sikorski, it is in fact the following: "Let **K** and **L** be classes of subsets of non-void spaces X and Y. If **K** and **L** contain all the one-element subsets of X and Y respectively, then each weak isomorphism between **K** and **L** is an equivalence between these classes."

Some explanation regarding terminology is necessary here. Spaces are just sets, for all that concerns us. The classes **K** and **L** are partially ordered by inclusion. A weak isomorphism is a 1–1 mapping between **K**, **L** that preserves inclusion. When **K** and **L** are Boolean algebras, a weak isomorphism between them is what Sikorski called isomorphism. Equivalence, F, is a 1–1 mapping between **K**, **L** for which there is another 1–1 mapping φ between X and Y such that for every K $\in$ **K**, F(K) = φ(K). Equivalence is always a weak isomorphism. The Szpilrajn-Marczewski Theorem stipulates the conditions for the opposite. Its proof is trivial: every weak isomorphism must map the singletons and hence the 1–1 mapping between the X, Y can be defined.

The same proof justifies the right to left direction of the lemma used by Sikorski, while the left to right direction is straightforward. So it appears that invoking the Szpilrajn-Marczewski Theorem by Sikorski was unnecessary.

33.2 Sikorski's Open Problems

At the closing of his paper Sikorski notes that also BDT can be expressed in the terms of Boolean algebras in the following way: If A and B are two elements of a Boolean algebra **A** such that A**A** is an isomorph of A′**A** and B**A** is an isomorph of B′**A**, then A**A** and B**A** are isomorphic. He added that the result can be obtained using BDT for complete and atomic Boolean algebras and presented the question whether this result could be extended to more general Boolean algebras. We do not know if this question was answered.

In the same issue of the journal where Sikorski's paper was published, on p 242, Sikorski raised another question: "Two compact metric spaces of dimension 0, each homeomorphic to an open subset of the other, are they necessarily homeomorphic?" Sikorski added that a negative answer to this question would imply that the requirement on σ-completeness is necessary for the Boolean algebra CBT. A step towards a negative answer to the question was made in Kuratowski 1959, for metric spaces of dimension 1. The result was generalized to metric spaces of dimension 0 in Kinoshita 1953. Then, in Hanf 1957, the result was provided directly in the language of Boolean algebras. Apparently, the different results demonstrate between them proof-processing links. Unfortunately, it is beyond the scope of this work to present this research project in any detail.

Chapter 34
Tarski's Proofs of BDT and the Inequality-BDT

We review Tarski's 1949b proof of the inequality-BDT: If $k \neq 0$ and $k\mathfrak{m} \leq k\mathfrak{n}$ then $\mathfrak{m} \leq \mathfrak{n}$, where k is a natural number and $\mathfrak{m}$, $\mathfrak{n}$ are cardinal numbers, from which Tarski easily deduced, using CBT, BDT: If $k \neq 0$ and $k\mathfrak{m} = k\mathfrak{n}$ then $\mathfrak{m} = \mathfrak{n}$.

In Lindenbaum-Tarski 1926, the general inequality version of BDT (inequality-BDT) was announced perhaps for the first time (§1 Theorem 39). It was stated there that Tarski proved the theorem for $k = 2$ and that in its full generality the theorem was proved by Lindenbaum. From this theorem, BDT was presented (Theorem 40) as a corollary, no doubt obtained using CBT.

Tarski published a proof following Sierpiński's proof of the theorem for $k = 2$ (see Sect. 28.3) and his declared failure to prove the theorem in its full generality, Tarski noted (p 78) that "it is in a sense an extension of the original proof" which he obtained for $k = 2$ in the 1920s. Concerning Lindenbaum's proof of the general theorem, Tarski could only recall that it was based on a lemma, which he had proved (Theorem 6 of the next section). This testimony is perhaps dubious because Theorem 6 is not mentioned in Lindenbaum-Tarski 1926, which is quite detailed in terms of the lemmas its theorems apply.

We also doubt Tarski's statement with regard to Lindenbaum's proof that "the idea used in an important part of the proof was rather similar to the one used by Sierpiński (1922)" proof of BDT. The reason for our doubt is that Tarski further noted that the same similarity to Sierpiński's proof applies to his own proof, yet Tarski himself noted, with regard to his proof of Theorem 7 (p 88 footnote 13), essential differences between his and Sierpiński's proof: Sierpiński defined his functions differently and he considered individual elements (J. Kőnig's strings gestalt) while Tarski "made an extensive use of the algebra of sets and set-images", i.e., used Dedekind's frames gestalt. So what similarity with Sierpiński Tarski had in mind is not clear.

We will give Tarski's proof in detail except that its main lemma (Theorem 7) we will bring only for a special case, which is enough to entail BDT and the inequality-BDT. Hence, perhaps, we come closer to the original proof of Lindenbaum. We finish this chapter with a few remarks on a proof of BDT for $k = 3$, given in 1994 by Doyle-Conway.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_34,

In his 1949b paper Tarski notes (p 78) that he had published a different proof of the inequality-BDT in his 1949a book "Cardinal algebras" (p 30ff). Actually, the reference is to the following theorem (2.31): If $k\mathfrak{m} + \mathfrak{p} \leq k\mathfrak{n} + \mathfrak{p}$ then $\mathfrak{m} + \mathfrak{p} \leq \mathfrak{n} + \mathfrak{p}$ (compare corollary 4 below). From this theorem it is easy to obtain the inequality-BDT (2.33) and BDT (2.34). Tarski further notes (*l.c.* footnote 6) that in the context of 1949a the proof of 2.34 involves the axiom of choice but that when transported to the context of cardinal numbers this use can be avoided. Unfortunately, presentation of the proof in Tarski 1949a is outside the scope of this book.

34.1 Tarski's 1949b Proof

Tarski proceeded as follows:

Theorem 1. *If* $\mathfrak{m} \leq \mathfrak{n}$ *and* $\mathfrak{m} \geq \mathfrak{n}$, *then* $\mathfrak{m} = \mathfrak{n}$.

This is CBT, which Tarski assumed without proof. Tarski defined $\mathfrak{m} \geq \mathfrak{n}$ to hold when there is a cardinal number $\mathfrak{r}$ such that $\mathfrak{m} = \mathfrak{n} + \mathfrak{r}$. Thus he avoided the circularity noted by Jourdain (see Sect. 17.2).

Theorem 2. *If* $\mathfrak{m} + \mathfrak{p} = \mathfrak{m} + \mathfrak{q}$ *then there are cardinals* $\mathfrak{n}$, $\mathfrak{p}'$, $\mathfrak{q}'$ *such that* $\mathfrak{m} = \mathfrak{m} + \mathfrak{p}' = \mathfrak{m} + \mathfrak{q}'$, $\mathfrak{p} = \mathfrak{n} + \mathfrak{p}'$, $\mathfrak{q} = \mathfrak{n} + \mathfrak{q}'$.

For proof of the theorem Tarski referenced his 1948 paper (Lemma A p 85). The proof there we summarize as follows:

By the conditions of the theorem there are sets M, P, M′, Q such that M + P ~ M′ + Q with P disjoint from M and Q from M′. Let U = M + P; we can assume that M′ and Q are subsets of U. Let φ be a 1–1 mapping between M and M′ that exists because M and M′ have the same power. U has the following partitions: M∩M′, P∩Q that is disjoint from both M and M′, P-Q = M′-M, Q-P = M-M′. Define:

$$N_n = (M\text{-}M') \cap \varphi^{-n}(M'\text{-}M),\ N'_n = (M'\text{-}M) \cap \varphi^{n}(M\text{-}M'),$$ [1]

$$C_1 = \sum_{n<\infty} N_n, D_1 = \sum_{n<\infty} N'_n, C_2 = M\text{-}M\cap M'\text{-}C_1, D_2 = M'\text{-}M\cap M'\text{-}D_1.$$

It is easy to become persuaded that $N'_n = \varphi^n(N_n)$ and that the N_n and N'_n are mutually disjoint, as are C_2 and D_2 from the pack. As φ is 1–1 we have that $C_1 \sim D_1$. Let *n* be the power of $(P\cap Q)\cup C_1$, which is the same as the power of $(P\cap Q)\cup D_1$. Since C_2 and D_2 have denumerably many pairwise disjoint copies in M∩M′, or else their members would be in some N_n or N'_n, we get that
$M\cap M' + C_2 \sim M\cap M' \sim M\cap M' + D_2$. We denote by $\mathfrak{p}'$ ($\mathfrak{q}'$) the power of C_2 (D_2). The theses of the theorem are now easy to verify.

The proof clearly leans on the frames gestalt but the frames, the N_n and N'_n, are not frames of a chain and so frames within either sequence are not necessarily

[1] For some n, the N_n, N'_n may be empty.

equivalent to each other. Nevertheless, the frames are obtained by a metaphor similar to the one used to obtain the frames in, say, Borel's CBT proof. The two mappings in this case are φ and φ^{-1} and the construction can be carried between any two disjoint sets and a 1–1 mapping φ. The particular setting of the theorem provides that φ (φ^{-1}) generate in M∩M′ a chain to the residues C_2 (D_2), obtained after removal of the aforementioned frames. Tarski's proof of Theorem 2 is a clear example of proof-processing. In his 1947c (p 113 Lemma 1) Sierpiński gave a proof of the same theorem utilizing instead the string gestalt of J. Kőnig, as he did in his proofs of BDT (1922, 1947a) and the inequality-BDT (1947d). Sierpiński defines the set C_2 (D_2) as the set of all members of M-M′ (M′-M) that generate by φ (φ^{-1}) an infinite string in M∩M′. Then C_1 = M-M′-C_2 and D_1 = M′-M-D_2. The rest of Sierpiński's proof is similar to that of Tarski.[2]

The importance of Theorem 2 is when we have two sets that we cannot assume to be disjoint, as in the situation of BDT where we consider different partitioning of the same set.[3] Theorem 2 explains how in this case the sets can be partitioned into subsets that have certain power relationships among them.

Theorem 2 was mentioned in Lindenbaum-Tarski 1926 (Theorems 5, 6 p 301).[4] The following consequences of Theorem 2 are also mentioned by Lindenbaum-Tarski:

- $\mathfrak{m} = \mathfrak{m} + \mathfrak{q}$ is equivalent to $\aleph_0\mathfrak{q} \leq \mathfrak{m}$ which is equivalent to $\mathfrak{m} = \mathfrak{m} + \aleph_0\mathfrak{q}$.[5] This is Zermelo's 1901 Denumerable Addition Theorem: $\mathfrak{m} = \mathfrak{m} + \mathfrak{q}$ entails $\mathfrak{m} = \mathfrak{m} + \aleph_0\mathfrak{q}$.
- Bernstein's 1905 inequality-BDT result: If $\mathfrak{m} + \mathfrak{m} = \mathfrak{m} + \mathfrak{m}'$ then $\mathfrak{m}' \leq \mathfrak{m}$.[6]

Theorem 3. *If* $k\mathfrak{m} + \mathfrak{p} \leq (k + 1)\mathfrak{m} + \mathfrak{q}$ *then* $\mathfrak{p} \leq \mathfrak{m} + \mathfrak{q}$.

The theorem generalizes Bernstein's inequality result which can be obtained from it by setting $k = 1$, $\mathfrak{p} = \mathfrak{m}'$ and $\mathfrak{q} = 0$ (Tarski 1949b p 81 footnote 10). The theorem is stated in Lindenbaum-Tarski 1926 as Theorem 20 proved by Tarski.

Theorem 3 is proved by induction on k. For $k = 0$ it is trivial. Let us assume it for k and prove it for $k + 1$.

[2] Sierpiński's motivation to prove theorem 2 in 1947c (there lemma 1) was for a proof of theorem 56 of Lindenbaum-Tarski 1926: $2^{\mathfrak{m}} - \mathfrak{m} = 2^{\mathfrak{m}}$. Sierpiński's proof may have provoked Tarski, who referenced Sierpiński's proof, to produce his own version in his 1948 paper.

[3] Still, let us recall that Sierpiński (see Chap. 28) and Kuratowski (see Chap. 30) did provide proofs of BDT under the assumption that the partitions are disjoint.

[4] Theorem 5 of Lindenbaum-Tarski 1926 is in the language of sets and mappings in which the above proof is given; theorem 6 there is in the language of cardinal numbers as is theorem 2 above.

[5] If $\aleph_0\mathfrak{q} \leq \mathfrak{m}$ then $\mathfrak{m} = \mathfrak{m}' + \aleph_0\mathfrak{q} = \mathfrak{m}' + 2\aleph_0\mathfrak{q} = \mathfrak{m}' + \aleph_0\mathfrak{q} + \aleph_0\mathfrak{q} = \mathfrak{m} + \aleph_0\mathfrak{q} = \mathfrak{m} + (\aleph_0 + 1)\mathfrak{q} = \mathfrak{m} + \aleph_0\mathfrak{q} + \mathfrak{q} = \mathfrak{m} + \mathfrak{q}$. However, $\mathfrak{q} \leq \mathfrak{m}$ does not entail $\mathfrak{m} = \mathfrak{m} + \mathfrak{q}$.

[6] For there are three cardinal numbers $\mathfrak{n}$, $\mathfrak{p}_1$, $\mathfrak{q}_1$ such that $\mathfrak{m} = \mathfrak{n} + \mathfrak{p}_1$ and $\mathfrak{m}' = \mathfrak{n} + \mathfrak{q}_1$ and $\mathfrak{m} + \mathfrak{p}_1 = \mathfrak{m} = \mathfrak{m} + \mathfrak{q}_1$. So $\mathfrak{m}' + \mathfrak{p}_1 = \mathfrak{n} + \mathfrak{q}_1 + \mathfrak{p}_1 = \mathfrak{n} + \mathfrak{p}_1 + \mathfrak{q}_1 = \mathfrak{m} + \mathfrak{p}_1 = \mathfrak{m}$ hence $\mathfrak{m} \geq \mathfrak{m}'$.

We begin with $(k+1)\mathfrak{m}+\mathfrak{p} \leq (k+2)\mathfrak{m}+\mathfrak{q}$. From it there is some cardinal $\mathfrak{r}$ such that $(k+1)\mathfrak{m}+\mathfrak{p}+\mathfrak{r} = (k+2)\mathfrak{m}+\mathfrak{q}$. Hence $\mathfrak{m}+(k\mathfrak{m}+\mathfrak{p}+\mathfrak{r}) = \mathfrak{m}+((k+1)\mathfrak{m}+\mathfrak{q})$. Setting $\mathfrak{p}^* = k\mathfrak{m}+\mathfrak{p}+\mathfrak{r}$, $\mathfrak{q}^* = (k+1)\mathfrak{m}+\mathfrak{q}$ we get $\mathfrak{m}+\mathfrak{p}^* = \mathfrak{m}+\mathfrak{q}^*$.[7] By Theorem 2, there are cardinals $\mathfrak{n}, \mathfrak{p}', \mathfrak{q}'$ such that $\mathfrak{m} = \mathfrak{m}+\mathfrak{p}' = \mathfrak{m}+\mathfrak{q}'$, $\mathfrak{p}^* = \mathfrak{n}+\mathfrak{p}'$, $\mathfrak{q}^* = \mathfrak{n}+\mathfrak{q}'$. Then, on the one hand, $\mathfrak{q}^*+\mathfrak{p}' = (k+1)\mathfrak{m}+\mathfrak{q}+\mathfrak{p}' = \mathfrak{n}+\mathfrak{q}'+\mathfrak{p}'$, and on the other, $\mathfrak{p}^*+\mathfrak{q}' = k\mathfrak{m}+\mathfrak{p}+\mathfrak{r}+\mathfrak{q}' = \mathfrak{n}+\mathfrak{p}'+\mathfrak{q}'$, so that $(k+1)\mathfrak{m}+\mathfrak{q}+\mathfrak{p}' = k\mathfrak{m}+\mathfrak{p}+\mathfrak{r}+\mathfrak{q}'$. But $(k+1)\mathfrak{m}+\mathfrak{q}+\mathfrak{p}' = k\mathfrak{m}+\mathfrak{q}+\mathfrak{m}+\mathfrak{p}' = k\mathfrak{m}+\mathfrak{q}+\mathfrak{m} = (k+1)\mathfrak{m}+\mathfrak{q} = k\mathfrak{m}+\mathfrak{p}+\mathfrak{r}+\mathfrak{q}' \geq k\mathfrak{m}+\mathfrak{p}$ so that by the induction hypothesis $\mathfrak{p} \leq \mathfrak{m}+\mathfrak{q}$, as required.

Corollary 4. $k\mathfrak{m}+\mathfrak{p} \leq k\mathfrak{m}+\mathfrak{q}$ *then* $\mathfrak{m}+\mathfrak{p} \leq \mathfrak{m}+\mathfrak{q}$.

For $k = 0$ the corollary is obvious. For $k > 0$, the hypothesis can be stated thus: $(k-1)\mathfrak{m}+(\mathfrak{m}+\mathfrak{p}) \leq k\mathfrak{m}+\mathfrak{q}$. By Theorem 3 then the corollary follows. Corollary 4 does not appear in Lindenbaum-Tarski 1926.

Corollary 5. *If* $k\mathfrak{m}+\mathfrak{p} = k\mathfrak{m}+\mathfrak{q}$ *then* $\mathfrak{m}+\mathfrak{p} = \mathfrak{m}+\mathfrak{q}$.

The corollary follows from Corollary 4 and CBT. Strangely, Corollary 5 appears in Lindenbaum-Tarski 1926 as Theorem 8, before Theorem 20 there, which is Tarski's 1949b Theorem 3, which serves in 1949b for its proof. The reason may have been that in Lindenbaum-Tarski 1926 Corollary 5 (Theorem 8 there) was proved directly from Theorem 6 there (Theorem 2 here) instead of Theorem 3. The direct proof, by induction, runs as follows:

The corollary is trivial for $k = 1$. Let us prove it for $k = 2$ because we need for the induction step to assume that $k > 2$. The hypothesis is $2\mathfrak{m}+\mathfrak{p} = 2\mathfrak{m}+\mathfrak{q}$ and we have to prove that $\mathfrak{m}+\mathfrak{p} = \mathfrak{m}+\mathfrak{q}$.[8] We can present the hypothesis as $\mathfrak{m}+(\mathfrak{m}+\mathfrak{p}) = \mathfrak{m}+(\mathfrak{m}+\mathfrak{q})$ and define $\mathfrak{p}^* = \mathfrak{m}+\mathfrak{p}$ and $\mathfrak{q}^* = \mathfrak{m}+\mathfrak{q}$. The hypothesis is then $\mathfrak{m}+\mathfrak{p}^* = \mathfrak{m}+\mathfrak{q}^*$. By Theorem 2 there are $\mathfrak{n}$, $\mathfrak{p}'$, $\mathfrak{q}'$ such that $\mathfrak{m} = \mathfrak{m}+\mathfrak{p}' = \mathfrak{m}+\mathfrak{q}'$, $\mathfrak{p}^* = \mathfrak{n}+\mathfrak{p}'$, $\mathfrak{q}^* = \mathfrak{n}+\mathfrak{q}'$. Then, on the one hand we have: $\mathfrak{q}^*+\mathfrak{p}' = \mathfrak{m}+\mathfrak{q}+\mathfrak{p}' = \mathfrak{m}+\mathfrak{p}'+\mathfrak{q} = \mathfrak{m}+\mathfrak{q}$, while on the other: $\mathfrak{q}^*+\mathfrak{p}' = \mathfrak{n}+\mathfrak{q}'+\mathfrak{p}' = \mathfrak{n}+\mathfrak{p}'+\mathfrak{q}' = \mathfrak{p}^*+\mathfrak{q}' = \mathfrak{m}+\mathfrak{p}+\mathfrak{q}' = \mathfrak{m}+\mathfrak{q}'+\mathfrak{p} = \mathfrak{m}+\mathfrak{p}$, so finally $\mathfrak{m}+\mathfrak{p} = \mathfrak{m}+\mathfrak{q}$, as needed.

Now we assume the theorem for $k > 2$ and prove it for $k+1$. The hypothesis is $(k+1)\mathfrak{m}+\mathfrak{p} = (k+1)\mathfrak{m}+\mathfrak{q}$. We have to prove that $\mathfrak{m}+\mathfrak{p} = \mathfrak{m}+\mathfrak{q}$. We can present the hypothesis as follows: $\mathfrak{m}+k\mathfrak{m}+\mathfrak{p} = \mathfrak{m}+k\mathfrak{m}+\mathfrak{q}$. Define $\mathfrak{p}^* = k\mathfrak{m}+\mathfrak{p}$ and $\mathfrak{q}^* = k\mathfrak{m}+\mathfrak{q}$. Then we have the hypothesis $\mathfrak{m}+\mathfrak{p}^* = \mathfrak{m}+\mathfrak{q}^*$. By Theorem 2 there are $\mathfrak{n}$, $\mathfrak{p}'$, $\mathfrak{q}'$ such that $\mathfrak{m} = \mathfrak{m}+\mathfrak{p}' = \mathfrak{m}+\mathfrak{q}'$, $\mathfrak{p}^* = \mathfrak{n}+\mathfrak{p}'$, $\mathfrak{q}^* = \mathfrak{n}+\mathfrak{q}'$. Then, on the one hand:

[7] Our notation here slightly differs from that of Tarski and as a result so does our chain of equalities. In Tarski there seems to be a typo on the third equality from the bottom of the proof: $\mathfrak{p}'$ is missing on the left side.

[8] Theorem 30 of Lindenbaum-Tarski 1926 says that $\mathfrak{m}+\mathfrak{p} = \mathfrak{m}+\mathfrak{q}$ iff $2\mathfrak{m}+\mathfrak{p} = 2\mathfrak{m}+\mathfrak{q}$.

$\mathfrak{q}^* + \mathfrak{p}' = k\mathfrak{m} + \mathfrak{q} + \mathfrak{p}' = k\mathfrak{m} + \mathfrak{p}' + \mathfrak{q} = (k\text{-}1)\mathfrak{m} + \mathfrak{m} + \mathfrak{p}' + \mathfrak{q} = k\mathfrak{m} + \mathfrak{m} + \mathfrak{q} = k\mathfrak{m} + \mathfrak{q}$, while on the other hand: $\mathfrak{q}^* + \mathfrak{p}' = \mathfrak{n} + \mathfrak{q}' + \mathfrak{p}' = \mathfrak{n} + \mathfrak{p}' + \mathfrak{q}' = \mathfrak{p}^* + \mathfrak{q}' = k\mathfrak{m} + \mathfrak{p} + \mathfrak{q}' = k\mathfrak{m} + \mathfrak{q}' + \mathfrak{m} = (k\text{-}1)\mathfrak{m} + \mathfrak{m} + \mathfrak{q}' + \mathfrak{p} = (k\text{-}1)\mathfrak{m} + \mathfrak{m} + \mathfrak{p} = k\mathfrak{m} + \mathfrak{p}$, so finally $k\mathfrak{m} + \mathfrak{p} = k\mathfrak{m} + \mathfrak{q}$. By the induction hypothesis $\mathfrak{m} + \mathfrak{p} = \mathfrak{m} + \mathfrak{q}$, and Corollary 5 is proved.

Theorem 6. *If $k \neq 0$, $k\mathfrak{m} \leq k\mathfrak{n}$, and $\mathfrak{m} \geq \mathfrak{n}$, then $\mathfrak{m} = \mathfrak{n}$.*

Since $\mathfrak{n} \leq \mathfrak{m}$ there is a cardinal $\mathfrak{r}$ such that $\mathfrak{m} = \mathfrak{n} + \mathfrak{r}$. On the other hand there is a cardinal $\mathfrak{s}$ such that $k\mathfrak{n} = k\mathfrak{m} + \mathfrak{s} = k\mathfrak{n} + k\mathfrak{r} + \mathfrak{s}$. By Corollary 5 with $\mathfrak{p} = 0$ and $\mathfrak{q} = k\mathfrak{r} + \mathfrak{s}$ we have that
$\mathfrak{n} = \mathfrak{n} + k\mathfrak{r} + \mathfrak{s} = \mathfrak{n} + \mathfrak{r} + (k\text{-}1)\mathfrak{r} + \mathfrak{s} = \mathfrak{m} + (k\text{-}1)\mathfrak{r} + \mathfrak{s} \geq \mathfrak{m}$. Since it is assumed that $\mathfrak{m} \geq \mathfrak{n}$, we have by CBT that $\mathfrak{n} = \mathfrak{m}$, as required.

Theorem 6 clearly entails BDT, under the additional assumption that $\mathfrak{m} \geq \mathfrak{n}$. Thus one is tempted to argue that if $k \neq 0$ and $k\mathfrak{m} \leq k\mathfrak{n}$ and it is not true that $\mathfrak{m} \leq \mathfrak{n}$, then $\mathfrak{m} \geq \mathfrak{n}$, so that $\mathfrak{m} = \mathfrak{n}$, a contradiction. But, as Tarski warned (p 83), without the axiom of choice, the additional assumption $\mathfrak{m} \geq \mathfrak{n}$ cannot always be made, for the Comparability Theorem for cardinal numbers is not given.

In Tarski 1930, BDT is given as Theorem 13 and around it are given several theorems similar to the ones given here, but in different order. Thus, there Corollary 5 is Theorem 13[b] which appears as a corollary to BDT not its lemma. It appears therefore that Theorem 6 can be obtained directly and not through Corollary 5. Indeed, we can prove that $k\mathfrak{n} = k\mathfrak{n} + k\mathfrak{r} + \mathfrak{s} \rightarrow \mathfrak{n} = \mathfrak{n} + \mathfrak{r} + \mathfrak{s}$[9] (compare Theorem 11 of Tarski 1930) from Zermelo's Denumerable Addition Theorem (which is derivable from Theorem 2, as mentioned):

$k\mathfrak{n} = k\mathfrak{n} + k\mathfrak{r} + \mathfrak{s}$ entails by Zermelo's theorem that $k\mathfrak{n} = k\mathfrak{n} + \aleph_0(k\mathfrak{r} + \mathfrak{s})$. From properties of $\aleph_0$ we obtain the following sequence of equations:

$$\begin{aligned} k\mathfrak{n} &= k\mathfrak{n} + \aleph_0(k\mathfrak{r} + \mathfrak{s}) = k\mathfrak{n} + \aleph_0(k\text{-}1)\mathfrak{r} + \aleph_0(\mathfrak{r} + \mathfrak{s}) \\ &= k\mathfrak{n} + \aleph_0(k\text{-}1)\mathfrak{r} + (\aleph_0 + 1)(\mathfrak{r} + \mathfrak{s}) \\ &= k\mathfrak{n} + \aleph_0(k\mathfrak{r} + \mathfrak{s}) + \mathfrak{r} + \mathfrak{s} = k\mathfrak{n} + \mathfrak{r} + \mathfrak{s}. \end{aligned}$$

Let $N_1, \ldots, N_k$ be k disjoint sets each of power $\mathfrak{n}$ and R a disjoint set from the N_i of power $\mathfrak{r} + \mathfrak{s}$. Then, by $k\mathfrak{n} = k\mathfrak{n} + \mathfrak{r} + \mathfrak{s}$ and Zermelo's Denumerable Addition Theorem, R has denumerably many copies in the union of the N_i. Let R_1 be the set of all members of R that have infinite copies in N_1. Let R_2 be the set of all members of $R - R_1$ that have infinite copies in N_2. Similarly we define R_i for all $i = 1, \ldots, k$. Some of the R_i may be empty. We clearly have $R = \Sigma_{i=1,\ldots,k} R_i$. By the reemergence argument (see Sect. 13.2) we have $N_i \sim N_i + R_i$, and therefore $N_1 \sim N_1 + R_i$. By repeated substitution we get $N_1 \sim N_1 + R$, so that $\mathfrak{n} = \mathfrak{n} + \mathfrak{r} + \mathfrak{s}$.

The indirect proofs that Tarski presented have the advantage of using cardinal arithmetic to obtain results about cardinals. The direct proof has the advantage that it maintains the link to the sets and mappings language that is anyhow necessary for the proof of Theorem 2.

[9] From which it follows that $\mathfrak{n} = \mathfrak{m} + \mathfrak{s}$ so that $\mathfrak{n} \geq \mathfrak{m}$ and by the given $\mathfrak{m} \geq \mathfrak{n}$ and CBT, $\mathfrak{m} = \mathfrak{n}$.

Theorem 7. *Let $k \geq 2$. Let $m_1, \ldots, m_k$ and $n_1, \ldots, n_k$ be two sequences of natural numbers, and $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ be a finite sequence of cardinals, such that $\sum_{j=1,\ldots,k}(m_i\mathfrak{p}_i) = \sum_{j=1,\ldots,k}(n_i\mathfrak{p}_i)$. Then there are two sequences of cardinals $\mathfrak{r}_1, \ldots, \mathfrak{r}_k$ and $\mathfrak{s}_1, \ldots, \mathfrak{s}_k$, satisfying the following conditions:*

(i) $\mathfrak{p}_i = \mathfrak{r}_i + \mathfrak{s}_i$ for $\mathfrak{i} = 1, \ldots, \mathfrak{k}$;
(ii) $\sum_{j=1,\ldots,k}(m_i\mathfrak{r}_i) = \sum_{j=1,\ldots,k}(n_i\mathfrak{r}_i)$ and $\mathfrak{r}_1 \geq \mathfrak{r}_2$;
(iii) $\sum_{j=1,\ldots,k}(m_i\mathfrak{s}_i) = \sum_{j=1,\ldots,k}(n_i\mathfrak{s}_i)$ and $\mathfrak{s}_1 \leq \mathfrak{s}_2$.

This theorem is the crux of the paper. It is an amazing example of Tarski's taste for the most general formulation. We wonder if the Theorem was ever used in its full generality. We will give in the next section a proof of the theorem for $k = 2$, $\mathfrak{p}_1 = \mathfrak{m}$, $m_1 = k$, $\mathfrak{p}_2 = \mathfrak{n}$, $n_2 = k$, $m_2 = n_1 = 0$. Under these conditions Theorem 7 becomes:

Simplified Theorem 7. *If $k\mathfrak{m} = k\mathfrak{n}$ then there are $\mathfrak{r}_1$, $\mathfrak{r}_2$, and $\mathfrak{s}_1$, $\mathfrak{s}_2$, such that* (i) $\mathfrak{m} = \mathfrak{r}_1 + \mathfrak{s}_1$, $\mathfrak{n} = \mathfrak{r}_2 + \mathfrak{s}_2$, (ii) $k\mathfrak{r}_1 = k\mathfrak{r}_2$, $\mathfrak{r}_1 \geq \mathfrak{r}_2$, $k\mathfrak{s}_1 = k\mathfrak{s}_2$, *and* $\mathfrak{s}_1 \leq \mathfrak{s}_2$.

The proof of the simplified Theorem 7 runs along the lines of Tarski's proof of Theorem 7. The simplified theorem is sufficient to obtain BDT and the inequality-BDT.

Theorem 8. *If $k \neq 0$ and $k\mathfrak{m} \leq k\mathfrak{n}$ then $\mathfrak{m} \leq \mathfrak{n}$.*

This is the inequality-BDT. Tarski proved the theorem by setting in Theorem 7: $k = 3$, $m_1 = n_2 = \mathrm{k}$, $m_2 = n_1 = n_3 = 0$, $m_3 = 1$, $\mathfrak{p}_1 = \mathfrak{m}$, $\mathfrak{p}_2 = \mathfrak{n}$, $\mathfrak{p}_3 = \mathfrak{t}$, where $\mathfrak{t}$ fulfills $k\mathfrak{m} + \mathfrak{t} = k\mathfrak{n}$. He then obtained (i) $\mathfrak{m} = \mathfrak{r}_1 + \mathfrak{s}_1$, $\mathfrak{n} = \mathfrak{r}_2 + \mathfrak{s}_2$, $\mathfrak{t} = \mathfrak{r}_3 + \mathfrak{s}_3$ (ii) $k\mathfrak{r}_1 + \mathfrak{r}_3 = k\mathfrak{r}_2$ and $\mathfrak{r}_1 \geq \mathfrak{r}_2$ (iii) $k\mathfrak{s}_1 + \mathfrak{s}_3 = k\mathfrak{s}_2$ and $\mathfrak{s}_1 \leq \mathfrak{s}_2$. From (ii), by Theorem 6, $\mathfrak{r}_1 = \mathfrak{r}_2 = \mathfrak{r}$ and hence $\mathfrak{m} = \mathfrak{r} + \mathfrak{s}_1 \leq \mathfrak{r} + \mathfrak{s}_2 = \mathfrak{n}$. We will prove the theorem through the simplified Theorem 7 in the next section.

Theorem 9. *If $k \neq 0$ and $k\mathfrak{m} = k\mathfrak{n}$ then $\mathfrak{m} = \mathfrak{n}$.*

This is finally BDT. The proof is by theorem 8 and CBT.

Following Theorem 9 Tarski brought few other similar results of which we will quote only one: If $k\mathfrak{m} < k\mathfrak{n}$ (then obviously $\mathrm{k} \neq 0$) then $\mathfrak{m} < \mathfrak{n}$. The proof is immediate from Theorem 8 because if $k\mathfrak{m} < k\mathfrak{n}$ then clearly $k\mathfrak{m} \leq k\mathfrak{n}$ so that $\mathfrak{m} \leq \mathfrak{n}$. If $\mathfrak{m} = \mathfrak{n}$ then $k\mathfrak{m} = k\mathfrak{n}$ contrary to the assumption. So we must have $\mathfrak{m} < \mathfrak{n}$.

34.2 The Proof of the Simplified Theorem

The plan of the proof runs as follows: When $k \neq 0$ and $k\mathfrak{m} = k\mathfrak{n}$, or even when $k\mathfrak{m} \leq k\mathfrak{n}$, there is a set U that is partitioned in two ways to k equivalent partitions, M_j of power $\mathfrak{m}$ and N_j of power $\mathfrak{n}$, $1 \leq j \leq k$. Thus there are 1–1 mappings that carry the equivalence of the equivalent partitions, say φ_j from M_1 onto M_j and ψ_j from N_1 onto N_j.[10]

[10] Note that φ_j and ψ_j are chosen from the many equivalences that exist between the partitions but since the number of choices is finite the axiom of choice is not invoked.

Following on Bernstein's metaphor, as implemented by Tarski in the proof of Theorem 7, consider all the compositions of φ_j, ψ_j and their inverses.[11] This set of compositions can be lexicographically ordered and so it can be presented in a sequence (p 84) which we denote by χ_n.[12] The χ_n form a group (a point noted by Bernstein but not by Tarski) and are 1–1.

For any subset X of U, the set $R = \sum\chi_n(X)$, the union of the images of X under all the χ_n, intersects with any two subsets of U, equivalent by some $\chi_{n'}$, in subsets that are equivalent by the same $\chi_{n'}$. The reason is that $\chi_n(R) \subseteq R$ for every n. As χ_n^{-1} is one of the χ_n, also $\chi_n^{-1}(R) \subseteq R$. R is a chain for every χ_n. Also $S = U\text{-}R$ has these properties of R.

The partitioning of U, in particular, casts partitioning of R and S. Let $\mathfrak{r}$ be the cardinal number of $R_j = R\cap M_j$ and $\mathfrak{r}'$ be the cardinal number of $R'_j = R\cap N_j$.[13] Similarly let $\mathfrak{s}$, $\mathfrak{s}'$ be the cardinal numbers of the corresponding partitions of S. We have (i) $\mathfrak{m} = \mathfrak{r} + \mathfrak{s}$, $\mathfrak{n} = \mathfrak{r}' + \mathfrak{s}'$ (ii) If $k\mathfrak{m} = k\mathfrak{n}$ then we clearly have $k\mathfrak{r} = k\mathfrak{r}'$; if $k\mathfrak{m} \leq k\mathfrak{n}$ we have $k\mathfrak{r} \leq k\mathfrak{r}'$ (iii) If $k\mathfrak{m} = k\mathfrak{n}$ then $k\mathfrak{s} = k\mathfrak{s}'$ and if $k\mathfrak{m} \leq k\mathfrak{n}$ then $k\mathfrak{s} \leq k\mathfrak{s}'$.

By a proper choice of X, it is possible to have in addition that $\mathfrak{r} \geq \mathfrak{r}'$ and $\mathfrak{s} \leq \mathfrak{s}'$. In this case, by Theorem 6 for $k\mathfrak{r} \leq k\mathfrak{r}'$, we have that $\mathfrak{r} = \mathfrak{r}'$. Therefore, under the conditions of the inequality-BDT, $\mathfrak{m} = \mathfrak{r} + \mathfrak{s} \leq \mathfrak{r} + \mathfrak{s}' = \mathfrak{n}$, so Theorem 8 is obtained.

For the definition of the set X we proceed to define[14] by induction an infinite sequence of sets G_n as follows: $G_1 = M_1\cap\chi_1^{-1}(N_1)$, namely, G_1 is the set of all members of M_1 that are mapped by χ_1 into N_1; assuming G_m is defined, we define $G_{m+1} = (M_1\text{-}\sum_{j=1,\ldots,m}G_m)\cap\chi_{m+1}^{-1}(N_1\text{-}\sum_{j=1,\ldots,m}\chi_j(G_j))$, namely, G_{m+1} consists of all the members of M_1, not included in previously defined G_js, that are mapped by χ_{m+1} into such members of N_1 that are not in the image of previously defined G_js.[15] The G_js are disjoint and so are their χ_j images, which we denote by G'_j, because the χ_j are 1–1. If we write $G = \sum_{j=1,\ldots,\infty}G_j$ and $G' = \sum_{j=1,\ldots,\infty} G'_j$, we have $G \sim G'$.[16] The X that we were looking for is M_1-G, and thus we define:

[11] Bernstein extended φ and ψ to include their inverse and he was followed in this by Sierpiński 1922 while Tarski used inverses, as we have already remarked with regard to his proof of Theorem 2. The advantage of Bernstein's notation is that φ and ψ, and thus the χ_n, are always defined while the advantage of Tarski's notation will become apparent in the definition of the G_n below. Tarski stressed this difference (p 88 footnote 13) and claimed that the Bernstein-Sierpiński method cannot be generalized, an observation, with regard to BDT, from which we differed (see Sects. 14.3 and 28.1).

[12] The necessity in having the χ enumerated will become clear in the proof.

[13] The R_j (R'_j) are equivalent by φ_j (ψ_j).

[14] Our notation deviates from Tarski's mainly in the signs used for entities.

[15] Compare this definition with the definition of the sets N_n in the proof of Theorem 2 above. This definition is the main metaphor of Tarski's paper discussed here. A similar definition appeared already in Bernstein's proof of BDT (see Sect. 14.2) for the sets to be interchanged. We have here perhaps another case of proof-processing.

[16] The equivalence mapping is the union of the χ_j reduced to G_j.

$$R = \sum_{j=1,\ldots,\infty} \chi_j(M_1\text{-}G),\ S = U\text{-}R, R_j = R \cap M_j,\ R'_j = R \cap N_j,$$
$$S_j = S \cap M_j,\ S'_j = S \cap N_j.$$

We have: $U = R + S$, $M_j = R_j + S_j$, $N_j = R'_j + S'_j$,
$R = \sum_{j=1,\ldots,k} R_j = \sum_{j=1,\ldots,k} R'_j$, $S = \sum_{j=1,\ldots,k} S_j = \sum_{j=1,\ldots,k} S'_j$.

If $x \in M_1$-G then $\chi_j(x)$ is not in N_1-G′ otherwise x would have been in G_{j+1}. So no member of R belongs to N_1-G′. Because G ~ G′ we have that G∩R ~ G′∩R but as R may extend to M_1-G, we have that $\mathfrak{r}' \leq \mathfrak{r}$. As for S_1, we have $S_1 \subseteq M_1$-R, but as one of the χ_j is the identity, we have: $M_1\text{-}R \subseteq M_1\text{-}(M_1\text{-}G) = G$. Hence $S_1 \subseteq G$ so obviously $S_1 = G \cap S \sim G' \cap S \subseteq S'_1$ so that $\mathfrak{s} \leq \mathfrak{s}'$. Thus the proof is complete.

We find Tarski's proof of Theorem 7 to be awesome. We have noted its possible origin in Bernstein's BDT proof but there could be other sources. A map of the proof-processing origins of Tarski's work would be a wonderful piece of art and heuristics.

34.3 The Doyle-Conway Proof

In Doyle-Conway (1994 p 27), Theorem 6, which is named there 'Tarski's lemma', is used to prove that from the inequality-BDT (for $k = 3$), BDT follows.[17] Doyle-Conway thus ignored Tarski's warning that Theorem 6 alone cannot prove BDT because the comparability of sets requires AC. They assumed that one of the sets is greater than the other (p 31). Doyle-Conway used AC in two other places: First, a proof that the connected part of the graph they use is denumerable, if not finite, requires the denumerable AC, as it requires the Denumerable Union Theorem. Second, Doyle-Conway presented their proof (for $k = 3$) assuming two sets of k-tuples with the further assumption that the members of each tuple are numbered (p 28, 31). However, such a numbering is possible only when the sets are finite. In fact, D. Kőnig used a similar idea in his 1916 proof of BDT. In the general case the axiom of choice is needed to provide such numbering.

Incidentally, Doyle-Conway use Sierpiński's 1922 algorithm from his proof of BDT for $k = 2$, which is, we believe, indeed extendible to any k. However, they use it in the case of the inequality-BDT, for which Sierpiński used, in his 1947d proof of the theorem for $k = 2$, another algorithm that requires the consideration of many cases. Doyle-Conway disregarded these subtleties. It thus seems to us that Doyle-Conway's claim that their proof is probably the proof that Lindenbaum announced in Lindenbaum-Tarski 1926, is not warranted.

[17] Incidentally, in their proof Doyle-Conway ignore the lemma $k\mathfrak{n} = k\mathfrak{n} + k\mathfrak{r} \rightarrow k\mathfrak{n} = k\mathfrak{n} + \mathfrak{r}$ but use the lemma $k\mathfrak{n} = k\mathfrak{n} + \mathfrak{r} \rightarrow \mathfrak{n} = \mathfrak{n} + \mathfrak{r}$ (lemma 3 p 26).

Chapter 35
Tarski's Fixed-Point Theorem and CBT

In Tarski's 1955 paper on the lattice theoretical Fixed-Point Theorem, several proofs of CBT are included, for sets, for homogeneous elements of a Boolean algebra and for Boolean algebras. We present these proofs in detail.

It was from Banach's 1924 Partitioning Theorem (see Chap. 29), that the Fixed-Point[1] Theorem for power-sets was extracted: by Whittaker implicitly, and by Knaster and Tarski explicitly (see Chap. 31). In 1955 Tarski ported the result to complete lattices and therewith derived the mentioned versions of CBT which were partially obtained earlier by Sikorski (1948; see Chap. 33).

In the following, after some introduction on lattices and Boolean algebras, we will begin with a description of a proof of CBT for sets and work our way backwards, down to the Fixed-Point Theorem. Along the way we encounter the Mean-Value Theorem and Banach's Partitioning Theorem, both ported to the context of Boolean algebras.

According to Tarski's testimony "the [fixed-point] theorem in its present form and its various applications and extensions were found by the author in 1939 and discussed by him in a few public lectures in 1939–1942 [reference is made to a note on lectures given in 1942, in the American Mathematical Monthly 49 (1942) p 402].[2] An essential part of Theorem 1 [the fixed-point theorem for lattices] was included in [reference is made to the second edition of Birkhoff 1967, dated 1948, p 54]; however, the author was informed by Professor Garrett Birkhoff that a proper historical reference to this result was omitted by mistake [It was added to the third edition of 1967, p 115.]" No leeway regarding priority.

[1] Tarski writes 'fixpoint' which we replaced throughout by 'fixed-point'.

[2] In footnote 6 p 304 of the 1955 paper Tarski notes that theorems 6–13 of that paper "are stated explicitly or implicitly" in his 1949a book.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_35,

35.1 Lattices and Boolean Algebras

In Tarski's paper, lattices and Boolean algebras are introduced thus (p 285, 296): A lattice is a system $\mathfrak{A} = \langle A, \leq \rangle$ formed by a non-empty set A, the domain of the lattice, and a binary relation $\leq$; it is assumed that $\leq$ establishes a partial order in A and that for any two elements $a, b \in A$ there is a least upper bound (join) $a \cup b$ and a greatest lower bound (meet) $a \cap b$. The relations $\geq, <, >$ are defined in terms of $\leq$. The join and the meet are idempotent operations ($a \cup a = a$, $a \cap a = a$).

The lattice is called complete if every subset B of A has a least upper bound $\cup B$ and a greatest lower bound $\cap B$. A complete lattice has in particular two elements $0 = \cap A$ and $1 = \cup A$.[3] Given $a, b \in A$ such that $a \leq b$, we define the interval $[a, b] = \{x \in A \mid a \leq x \leq b\}$.[4] $[a, b]$ is a lattice and is complete if A is. A function f from A to A is called increasing if for any two elements x, y of A, $x \leq y$ implies $f(x) \leq f(y)$. A fixed-point of f is an element x such that $f(x) = x$.

A Boolean algebra is a distributive lattice[5] $A = \langle A, \leq \rangle$ with 0, 1 and an additional operation: for every $a \in A$ there is a unique $\bar{a}$ in A, called the complement of a, such that $a \cup \bar{a} = 1$ and $a \cap \bar{a} = 0$. The uniqueness implies that the complement of the complement is the original element. Using the complement operation, a new operation is defined: for any two elements $a, b \in A$, the difference of a and b, a-b, is defined by $a\text{-}b = a \cap \bar{b}$. Clearly $a\text{-}b \leq a$. If A is a Boolean algebra then for every $a \in A$, $\langle [0, a], \leq \rangle$ is a Boolean algebra with a in the role of 1 and for every $b \in [0, a]$, $\bar{b}$ is $\bar{b} \cap a$. We call $[0, a]$ the interval of a. A complete Boolean algebra is a Boolean algebra that is complete as a lattice.

This introduction of lattices and Boolean algebras is clearly only sketchy; especially it misses to mention that $x = y$ iff $x \geq y$ and $y \geq x$, the analog of equality between sets by extensionality. In addition, the dual lattice is not explicitly mentioned in Tarski's 1955, though it is used. Several basic lemmas necessary for the proofs below, we will state and prove when they are called.

35.2 A Proof of CBT

Towards the end of his paper (p 306 end of Sect. 3), Tarski says that his Theorem 10 "yields the well-known Cantor-Bernstein Theorem". Theorem 10 is the following:

---

[3] Tarski notes (p 305) that most results of the 1955 paper still hold if instead of assuming that the relevant lattice or Boolean algebra is complete, it is assumed to be σ-complete, namely, that $\cup B$ and $\cap B$ exist for countable B. This is so in particular with regard to the theorems required to obtain CBT, which is the focal point of our interest.

[4] Instead of the notation { .. | . . . } Tarski uses the notation E_x[. . .]. It seems to be a variant on Peano's 'x∋', perhaps through Russell (see Kennedy 2002, p 43, Grattan-Guinness 1977 'class of' in the index). We replaced Tarski's notation throughout.

[5] Namely, $(a \cap b) \cup c = (a \cup c) \cap (b \cup c)$ and $(a \cup b) \cap c = (a \cap c) \cup (b \cap c)$. Tarski did not mention the distributivity requirement, or the commutativity and associativity of the meet and join operations, though all these features are necessary to obtain the lemmas that Tarski is assuming tacitly.

$\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, if a, b, c∈A, a ≤ b ≤ c, and a ≈ c, then a ≈ b ≈ c.

Even before the sign ≈ for the relation between homogeneous elements in a Boolean algebra is explained, the formal similarity of Theorem 10 and the single-set CBT is obvious. This is surely why Tarski called Theorem 10 the Equivalence Theorem.

The relation ≈ Tarski introduced by the following:

two elements a, b ∈ A are called homogeneous, in symbols a ≈ b, if the Boolean algebras ⟨[0, a], ≤⟩ and ⟨[0, b], ≤⟩ are isomorphic. In other words, a ≈ b if and only if there is a function f satisfying the following conditions: the domain of f is [0, a]; the range of f is [0, b]; the formulas x ≤ y and f(x) ≤ f(y) are equivalent for any x, y ∈ [0, a].[6]

The main properties of ≈ were summarized by Tarski in Theorem 6 (p 298):

(i) a ≈ a for every a ∈ A; [reflexivity]
(ii) If a ≈ b then b ≈ a; [commutativity]
(iii) If a ≈ b and b ≈ c then a ≈ c; [transitivity]
[Hence ≈ is an equivalence relation] [7]
(iv) If $a_1, a_2, b_1, b_2 \in A$, $a_1 \cap a_2 = 0 = b_1 \cap b_2$, $a_1 \approx b_1$, and $a_2 \approx b_2$, then $a_1 \cup a_2 \approx b_1 \cup b_2$;
(v) If $a, b_1, b_2 \in A$, $b_1 \cap b_2 = 0$, and $a \approx b_1 \cup b_2$, then there are elements $a_1, a_2 \in A$ such that $a_1 \cup a_2 = a$, $a_1 \cap a_2 = 0$, $a_1 \approx b_1$, and $a_2 \approx b_2$.

(i)–(iii) are easy to obtain from the definition of ≈. Re (iv): when two elements of a Boolean algebra have the meet 0, their intervals have only 0 in their intersection and therefore isomorphisms, which necessarily take 0 to itself, with domain on each interval and disjoint ranges, can be combined, as in (iv). Re (v): let *f* be the isomorphism between the intervals of a and of $b_1 \cup b_2$; b_1 and b_2 are in the interval of $b_1 \cup b_2$; define $a_1 = f^{-1}(b_1)$, $a_2 = f^{-1}(b_2)$; it is easily seen the required properties are obtained.

Conditions (iv) and (v) remind us of Banach's properties (α), (β). Indeed, if $a \approx b$ and we denote by R≈ the relation induced by *f* on the subsets of the intervals of *a* and *b*, then R≈ has properties (α), (β). We will come back to this point later.

Obviously, by the extensionality of ≤, the isomorphism mentioned in the definition of ≈ is 1–1. Thus a ≈ b entails [0, *a*] ~ [0, *b*], where ~ is the set theoretic equivalence. When the Boolean algebra is a power-set of a set, ≤ is the relation of subsets ⊆, and the interval [0, A] for a set A in this algebra, is the power-set of A, P(A). An isomorphism in this Boolean algebra preserves the subset relation. If A ~ B, any mapping that provides this equivalence can be extended to a mapping between P(A) and P(B), which preserves the subset relation, so that A ~ B entails A ≈ B. If A, B are sets in some power-set, then if A ≈ B, the power-sets of A and B are isomorphic, namely, they are equivalent, with the equivalence

[6] The relation of homogeneity is central to the theorems of the 1955 paper, but Tarski notes that these theorems are not new and references his 1949a book (§§ 11, 12, 15–17) for an exposition of these results in a different setting, with historical references to earlier publications.

[7] These terms in [ ] are not used by Tarski.

mapping preserving the subset relation. In this case every singleton of A is mapped necessarily to a singleton of B, and every singleton of B is the image of a singleton of A. Thus we can easily define a 1–1 mapping between A and B from the given mapping between P(A) and P(B). Hence A $\approx$ B entails that A ~ B. Thus, in a power-set, A ~ B iff A $\approx$ B. Cf. Chap. 33 regarding the Szpilrajn-Marczewski Theorem.

Prior to his remark on CBT Tarski makes the following remark: "Any given sets A, B, C, . . . can be regarded as elements of a complete Boolean algebra; in fact, of the algebra formed by all subsets of A∪B∪C∪ . . ., with set-theoretic inclusion as the fundamental relation". Thus if we have three sets satisfying the conditions of the single-set formulation of CBT, A $\subseteq$ B $\subseteq$ C ~ A, we look at P(C); there we have A $\approx$ C and by Theorem 10 we get B $\approx$ C, which entails as explained above that B ~ A, namely, the thesis of CBT. This derivation of CBT through the mapping underlying the relation $\approx$, shares a metaphor with the proof of CBT given by Sikorski because of its use of the mapping associated with the Szpilrajn-Marczewski Theorem (see Sect. 33.1).

35.3 The Mean-Value Theorem

For proof of Theorem 10, the Equivalence Theorem, Tarski invokes Theorem 8, which he calls the 'Mean-Value Theorem' (see Sect. 25.2):

> $\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, if a, b, c, a′, c′ $\in$ A, a $\leq$ b $\leq$ c, a′ $\leq$ c′, a $\approx$ a′ and c $\approx$ c′, then there is an element b′ $\in$ A such that a′ $\leq$ b′ $\leq$ c′ and b $\approx$ b′.

If in Theorem 8 we take $a' = c' = c$, then the conditions of Theorem 8 become those of Theorem 10 and necessarily $a' = b' = c'$ so that $a' \approx b' \approx c'$ and as by Theorem 8 $b \approx b'$, by Theorem 6(iii) also $a \approx b \approx c$, as required by Theorem 10.

Theorem 8 generalizes the Schröder-Korselt Theorem (see Sect. 25.2), which was proved by Korselt, proof-processing his proof of CBT. To use the theorem here, however, Tarski needed a proof that avoids CBT. Such a direct proof was mentioned already in Lindenbaum-Tarski 1926 (see Sect. 25.2) but no details were given. In 1955 Tarski appealed for the proof of Theorem 8 to Theorem 7:

> $\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, if a, b_1, b_2, c, d $\in$ A, $b_1 \cap b_2 = 0$, c $\approx$ d, and a∪c $\approx$ $b_1 \cup b_2 \cup d$, then there are a_1, a_2 $\in$ A such that $a_1 \cup a_2 = a$, $a_1 \cap a_2 = 0$, $a_1 \cup c \approx b_1 \cup d$, and $a_2 \cup c \approx b_2 \cup d$.

Theorem 8 follows indeed from Theorem 7, for if in Theorem 7, we substitute c'-a', b-a, c-b, a', a, for a, b_1, b_2, c, d, we obtain that c'-$a' = a_1 \cup a_2$ and $(b\text{-}a) \cup a \approx a_1 \cup a'$. It is easy to see that $(b\text{-}a) \cup a = b$.[8] Defining $b' = a_1 \cup a'$ we have $b \approx b'$ and clearly

[8] $(b\text{-}a) \cup a = (b \cap \bar{a}) \cup a$. By definition of the meet $b \cap \bar{a} \leq b$ and it is given that $a \leq b$ so by the definition of the join $(b\text{-}a) \cup a \leq b$. Let e be an element such that $(b\text{-}a) \cup a \leq e$. By distributivity $(b\text{-}a) \cup a = (b \cup a) \cap (\bar{a} \cup a) = (b \cup a) \cap 1 = b \cup a$, as the meet of every element and 1 is the element itself. Thus $(b\text{-}a) \cup a \leq e$ entails $b \cup a \leq e$ and therefore $b \leq e$ and so b is the least upper bound of b-a and a and $(b\text{-}a) \cup a = b$.

$a' \leq b'$. Now $b' \leq b' \cup a_2 = a_1 \cup a' \cup a_2$ and because of the commutativity and associativity of the join, $a_1 \cup a' \cup a_2 = a_1 \cup a_2 \cup a' = (c'\text{-}a') \cup a'$. Because $a' \leq c'$, $(c'\text{-}a') \cup a' = c'$,[9] so that $b' \leq c'$. Thus Theorem 8 is obtained.

The relation between Theorem 8 and Theorem 7 may appear arbitrary, but shifting to the context of set theory clarifies the situation: If we have $A \subseteq B \subseteq C$ then $C = (C\text{-}B) \cup (B\text{-}A) \cup A$ and if we have $A_1 \subseteq C_1$, then $C_1 = (C_1\text{-}A_1) \cup A_1$. To apprehend the situation in Theorem 7 when coming from Theorem 8 we need a gestalt switch: Instead of apprehending A, B, C as a pile of increasing plates, the reverse of Hanoi tower, we have to see them form a three-floors Roman tower: A, B-A, C-B. A similar notation for the sets in CBT appeared already in Poincaré's second inductive proof. There, however, we did not notice the point because the notation appeared natural for the inductive process that followed.

Theorem 7 expands Theorem 8 because it provides, in the notation of Theorem 8, not only b' which is $\approx$ to b, but also $b'' = a' \cup (c'\text{-}b')$ which is $\approx$ to $a \cup (c\text{-}b)$. The symmetry between b' and b'', which is usually concealed in formulations of CBT, was first mentioned in Zermelo's 1901 paper.

> *Proof [of Theorem 7]*: By the definition of homogeneity, the formula $c \approx d$ implies the existence of a function f which maps isomorphically the Boolean algebra $\langle [0, c], \leq \rangle$ onto the Boolean algebra $\langle [0, d], \leq \rangle$; we have in particular
> (1) $f(c) = d$. Similarly, the formula $a \cup c \approx b_1 \cup b_2 \cup d$ implies the existence of a function g which maps isomorphically $\langle [0, b_1 \cup b_2 \cup d], \leq \rangle$ onto $\langle [0, a \cup c], \leq \rangle$,[10] and we have
> (2) $g(b_1 \cup b_2 \cup d) = a \cup c$.
> We can assume for a while that the domain of g has been restricted to the interval $[0, b_1 \cup d]$. Thus f is an increasing function on $[0, c]$ to A, g is an increasing function on $[0, b_1 \cup d]$ to A,[11] and by applying Theorem 5 we obtain two elements c', d' such that
> (3) $f(c\text{-}c') = d'$ and $g((b_1 \cup d)\text{-}d') = c'$.

At this stage the plan of the proof becomes clear: f is an isomorphism between the intervals of $c\text{-}c'$ and d', and g is an isomorphism between the intervals of $(b_1 \cup d)\text{-}d'$ and c'. If we can show that the respective intervals are disjoint we could apply Theorem 6(iv) to obtain the homogeneity of their union. First, however, we have to define a_1, a_2 and to show that they partition a, as stated in the theorem. Tarski obtains these results as follows:

> The functions f and g being increasing, formulas (1)–(3) imply
> (4) $d' \leq d$ and $c' \leq a \cup c$.[12]

[9] $(c'\text{-}a') \cup a' = (c' \cap \bar{a}') \cup a' = (c' \cup a') \cap (\bar{a}' \cup a') = (c' \cup a') \cap 1 = c' \cup a'$ and because $a' \leq c'$, it is easy to verify that the join of c' and a' is c'.

[10] f and g are taken with crossed domains and ranges to make them ready for the application of Theorem 5 which we will discuss below.

[11] Theorem 5 is about increasing functions from intervals into A so Tarski stresses that f, g are such functions, though the conditions of Theorem 7 provide more properties to f, g.

[12] Clearly, $c\text{-}c' = c \cap \bar{c}' \leq c$ because the meet stands by definition in the relation $\leq$ to its components. Because f is increasing $d' = f(c\text{-}c') \leq f(c) = d$. Similarly, $(b_1 \cup d)\text{-}d' \leq (b_1 \cup d)$ and because the join stands in the relation $\leq$ to its components, $(b_1 \cup d) \leq b_1 \cup b_2 \cup d$. Because g is increasing $c' = g((b_1 \cup d)\text{-}d') \leq g(b_1 \cup b_2 \cup d) = a \cup c$. If we were in a context of power-set, d' would be a subset of d and c' a set with elements in c and in a.

We now let (5) $a_1 = c'$-c and $a_2 = a$-a_1. By (4) we have c'-c $\leq$ a,[13] and hence, by (5), (6) $a_1 \cup a_2 = a$[14] and $a_1 \cap a_2 = 0$.

Now Tarski moves to tackle the heart of the matter:

From (4) and (5) we also obtain
(7) (c-c')$\cup c' = a_1 \cup c$ and (c-c')$\cap c' = 0$,[15]
(8) $d' \cup ((b_1 \cup d)$-$d') = b_1 \cup d$ and $d' \cap ((b_1 \cup d)$-$d') = 0$.[16]
Since f maps isomorphically $\langle [0, c], \leq \rangle$ onto $\langle [0, d], \leq \rangle$, we conclude from (3) that it also maps isomorphically $\langle [0, c$-$c'], \leq \rangle$ onto $\langle [0, d'], \leq \rangle$ and that consequently (9) c-$c' \approx d'$. Analogously, by (3),
(10) $c' \approx (b_1 \cup d)$-d'. By Theorem 6(iv), formulas (7)–(10) imply
(11) $a_1 \cup c \approx b_1 \cup d$.

By a similar train of reasoning Tarski intends to prove that $a_2 \cup c \approx b_2 \cup d$. Since in the previous part we started from $f(c$-$c')$ in this part the starting point is $c \cap c'$. The proof proceeds as follows:

Furthermore, from (4) and (5) we derive
(12) $(c \cap c') \cup ((a \cup c)$-$c') = a_2 \cup c$ and $(c \cap c') \cap ((a \cup c)$-$c') = 0$,[17]
(13) (d-d')$\cup ((b_2$-d$) \cup d') = b_2 \cup d$ and (d-d')$\cap ((b_2$-d$) \cup d') = 0$.[18]
The function f being an isomorphic transformation,[19] we obtain, with the help of (1) and (3), $f(c \cap c') = f(c$-$(c$-$c')) = f(c)$-$f(c$-$c') = d$-d',[20] and hence, by arguing as above in the proof of (9),

[13] c'-$c = c' \cap \bar{c} \leq c', \bar{c}$. Since $c' \leq a \cup c$ by (4), we have that c'-$c \leq a \cup c, \bar{c}$, so that c'-c is a lower bound of $a \cup c, \bar{c}$. Because the meet is the greatest lower bound, c'-$c \leq (a \cup c) \cap \bar{c}$. By distributivity, $(a \cup c) \cap \bar{c} = (a \cap \bar{c}) \cup (c \cap \bar{c}) = (a \cap \bar{c}) \cup 0 = (a \cap \bar{c}) \leq a$ because the join of any element and 0 is that element and the meet stands in the relation $\leq$ to its components.

[14] $a_1 \cup a_2 = a_1 \cup (a$-$a_1) = a_1 \cup (a \cap \bar{a}_1) = (a_1 \cup a) \cap (a_1 \cup \bar{a}_1) = (a_1 \cup a) \cap 1 = a_1 \cup a = a$ because $a_1 \leq a$ and their join is therefore a. A similar argument provides $a_1 \cap a_2 = 0$.

[15] We are looking to partition $a_1 \cup c$ into c-c', because of (3), and its complement in $a_1 \cup c$, which is c'. For the first equality,
$(c$-$c') \cup c' = (c \cap \bar{c}') \cup c' = (c \cup c') \cap (\bar{c}' \cup c') = (c \cup c') \cap 1 = c \cup c'$ while also
$a_1 \cup c = (c'$-$c) \cup c = (c' \cap \bar{c}) \cup c = (c' \cup c) \cap (\bar{c} \cup c) = (c' \cup c) \cap 1 = c' \cup c = c \cup c'$, hence the result. For the second equality $(c$-$c') \cap c' = (c \cap \bar{c}') \cap c' = c \cap (\bar{c}' \cap c') = c \cap 0 = 0$.

[16] We are looking to partition $b_1 \cup d$ into $(b_1 \cup d)$-d', because of (3), and its complement in $b_1 \cup d$, which is d'. For the first equality
$d' \cup ((b_1 \cup d)$-$d') = d' \cup ((b_1 \cup d) \cap \bar{d}') = (d' \cup b_1 \cup d) \cap (d' \cup \bar{d}') = b_1 \cup d \cup d' = b_1 \cup d$ because $d' \leq d$. For the second equality $d' \cap ((b_1 \cup d)$-$d') = d' \cap ((b_1 \cup d) \cap \bar{d}') = (b_1 \cup d) \cap d' \cap \bar{d}' = 0$.

[17] For the first equality,
$a_2 \cup c = (a$-$a_1) \cup c = (a$-$(c'$-$c)) \cup c = (a \cap (\bar{c}' \cup c)) \cup c = (a \cup c) \cap (\bar{c}' \cup c \cup c) = (a \cup c) \cap (c \cup \bar{c}')$ and also $(c \cap c') \cup ((a \cup c)$-$c') = (c \cap c') \cup ((a \cup c) \cap \bar{c}') = ((c \cap c') \cup (a \cup c)) \cap ((c \cap c') \cup \bar{c}') = (a \cup c) \cap (c \cup \bar{c}')$, hence the result. For the second equality, $(c \cap c') \cap ((a \cup c)$-$c') = (c \cap c') \cap ((a \cup c) \cap \bar{c}') = (a \cup c) \cap c \cap c' \cap \bar{c}' = 0$.

[18] For the first equality, $(d$-$d') \cup ((b_2$-$d) \cup d') = (d \cap \bar{d}') \cup d' \cup (b_2 \cap \bar{d}) = ((d \cup d') \cap (\bar{d}' \cup d')) \cup (b_2 \cap \bar{d}) = d \cup (b_2 \cap \bar{d}) = (d \cup b_2) \cap (d \cup \bar{d}) = d \cup b_2$. For the second equality, $(d$-$d') \cap ((b_2$-$d) \cup d') = ((d \cap \bar{d}') \cap ((b_2 \cap \bar{d}) \cup d')) = ((d \cap \bar{d}') \cap d') \cup ((d \cap \bar{d}') \cap (b_2 \cap \bar{d})) = 0$.

[19] Clearly Tarski means 'isomorphism'.

[20] Because f is order preserving it preserves meets, joins, complements and difference.

(14) $c \cap c' \approx d\text{-}d'$.
Since by (4) and the hypothesis, $(b_2\text{-}d) \cup d' = (b_1 \cup b_2 \cup d)\text{-}((b_1 \cup d)\text{-}d')$,[21] we conclude analogously, with the help of (2) and (3), that
$g((b_2\text{-}d) \cup d') = g(b_1 \cup b_2 \cup d)\text{-}g((b_1 \cup d)\text{-}d') = (a \cup c)\text{-}c'$ and therefore
(15) $(a \cup c)\text{-}c' \approx (b_2\text{-}d) \cup d'$.
From (12)–(15), by applying Theorem 6(iv) again, we get
(16) $a_2 \cup c \approx b_2 \cup d$. By (6), (11), and (16), the proof is complete.

The conditions of Theorem 7 are similar, though not identical, to the conditions of Theorem 2 of Sect. 34.1. However, the proofs seem to differ and the question arises whether this is an essential difference. Namely, can Theorem 2 also be proved by way of Banach's Partitioning Theorem rather than its direct application of the proof of CBT. Another question is to explicate how come a single application of the Partitioning Theorem (Theorem 5) was enough to provide for the results for both a_1 and a_2. Finally we wonder whether using the relation R≈ could shorten the proof by reliance on Theorem 3 of Banach (see Sect. 29.1). Answering such questions, which are proof-processing questions and which may provide insight to the proof-processing origin of Tarski's proof, must be left, however, outside the scope of this book.

35.4 The Partitioning Theorem for Boolean Algebras

Theorem 5 is the following:

Let (i) $\mathfrak{A} = \langle A, \leq \rangle$ be a complete Boolean algebra, (ii) a, b be elements of A, f be an increasing function on [0, a] to A, and g an increasing function on [0, b] to A. Then there are elements a′, b′ in A such that $f(a\text{-}a') = b'$ and $g(b\text{-}b') = a'$.

Theorem 5 is a variant of Banach's Partitioning Theorem of 1924 (see Chap. 29), so we call it the Partitioning Theorem for Boolean algebras. Its proof runs as follows:

Consider the function h defined by the formula
(1) $h(x) = f(a\text{-}g(b\text{-}x))$ for every x in A. Let x and y be any two elements of A such that $x \leq y$. We have then $b\text{-}x \geq b\text{-}y$[22] and since b-x and b-y are in [0, b], and g is an increasing function on [0, b] to A, we conclude that $g(b\text{-}y) \leq g(b\text{-}x)$ and
$a\text{-}g(b\text{-}x) \leq a\text{-}g(b\text{-}y)$.[23] Hence the elements a-g(b-x), a-g(b-y) being in [0, a] and f being increasing on [0, a], we obtain $f(a\text{-}g(b\text{-}x)) \leq f(a\text{-}g(b\text{-}y))$, that is, by (1), $h(x) \leq h(y)$. Thus h is an increasing function on A to A and consequently, by Theorem 1 [see below], it has a

[21] $(b_1 \cup b_2 \cup d)\text{-}((b_1 \cup d)\text{-}d') = (b_1 \cup b_2 \cup d) \cap ((\bar{b}_1 \cap d) \cup d') = ((b_1 \cup b_2 \cup d) \cap (\bar{b}_1 \cap d)) \cup ((b_1 \cup b_2 \cup d) \cap d') = (b_1 \cap (\bar{b}_1 \cap d)) \cup (b_2 \cap (\bar{b}_1 \cap d)) \cup (d \cap (\bar{b}_1 \cap d)) \cup ((b_1 \cup b_2 \cup d) \cap d')$. The first term of this expression is obviously 0 so it can be eliminated. The third term is reduced to d' because by (4) $d' \leq d$. The second term is reduced to $b_2 \cap d$, namely to $b_2\text{-}d$ because $b_2 \leq \bar{b}_1$ because:
$b_2 \leq b_2 \leq b_2 \cap 1 = b_2 \cap (b_1 \cup \bar{b}_1) = (b_2 \cap b_1) \cup (b_2 \cap \bar{b}_1) = b_2 \cap \bar{b}_1 \leq \bar{b}_1$. So the desired equality is established.

[22] $\bar{x} \cap y \leq \bar{x}$, x so $\bar{x} \cap y \leq \bar{x} \cap x = 0$ so $\bar{x} \cap y = 0$. Now $\bar{x} = \bar{x} \cap (y \cup \bar{y}) = (\bar{x} \cap y) \cup (\bar{x} \cap \bar{y}) = \bar{x} \cap \bar{y} \leq \bar{y}$ so $x \leq y$ entails $\bar{x} \geq \bar{y}$. $b\text{-}\bar{x} = b \cap x \leq b$, y so $b \cap x \leq b \cap y = b\text{-}\bar{y}$.

[23] By the same lemma as above.

fixed-point b′. Hence, by (1), (2) f(a-g(b-b′)) = b′. We put (3) g(b-b′) = a′. From (2) and (3) we see at once that the elements a′ and b′ satisfy the conclusion of the theorem.

Following the proof of Theorem 5 Tarski noted (p 297) that if we assume in addition that $f(a) \leq b$ and $g(b) \leq a$, Theorem 5 can be improved. In this case $a' \leq a$, $b' \leq b$, and we can define $a'' = a\text{-}a'$, $b'' = b\text{-}b'$; we then have that $f(a'') = b'$, $g(b'') = a'$. Tarski further remarked (footnote 5) that in this form Theorem 5 is a generalization of a set-theoretic theorem obtained by him and Knaster in 1927 (Knaster 1928; see Sect. 31.2). The theorem referenced is the following:

f and g being monotone functions[24] of sets[25] such that $B_1 = f(A)$ and $A_1 = g(B)$ where $A_1 \subseteq A$, $B_1 \subseteq B$, there exist sets D, E, F and G such that A = D∪E, B = F∪G, D∩E = 0 = F∩G and F = f(D) and E = g(G).

Banach proved his theorem using J. Kőnig's string gestalt; Tarski-Knaster proved their Partitioning Theorem by a fixed-point theorem, the predecessor of Theorem 1 referenced above, to the proof of which we now turn.

35.5 The Fixed-Point Theorem for Lattices

Tarski's 'Fixed-Point Theorem for lattices' is the following, which we quote with its proof:

Let (i) $\mathfrak{A} = \langle A, \leq \rangle$ be a complete lattice, (ii) f be an increasing function from A to A, (iii) P the set of all fixed-points of f. Then the set P is not empty and the system ⟨P, ≤⟩ is a complete lattice with $\cup P = \cup\{x \mid f(x) \geq x\}$ and $\cap P = \cap \{x \mid f(x) \leq x\}$.[26]
Proof: Let (1) $u = \cup\{x | f(x) \geq x\}$. We clearly have $x \leq u$ for every element x with $f(x) \geq x$; hence, the function f being increasing, $f(x) \leq f(u)$ and $x \leq f(u)$. By (1) we conclude that (2) $u \leq f(u)$.[27] Therefore $f(u) \leq f(f(u))$, so that f(u) belongs to the set $\{x \mid f(x) \geq x\}$; consequently, by (1), (3) $f(u) \leq u$. Formulas (2) and (3) imply that u is a fixed-point of f; hence we conclude by (1) that u is the join of all fixed-points of f,[28] so that (4) $\cup P = \cup\{x \mid f(x) \geq x\} \in P$.
Consider the dual lattice $\mathfrak{A}' = \langle A, \leq \rangle$. $\mathfrak{A}'$, like A, is complete, and f is again an increasing function in $\mathfrak{A}'$. The join of any elements in $\mathfrak{A}'$ obviously coincides with the meet of these elements in A. Hence, by applying to A′ the result established for $\mathfrak{A}$ in (4), we conclude that (5) $\cap P = \{x \mid f(x) \leq x\} \in P$.

[24] Namely, such that preserve the ⊆ relation, or isomorphic in the 1955 terminology.

[25] Namely, defined in some power-set.

[26] Tarski notes (p 305) that if the lattice is only σ-complete, has 0 and f is "distributive under [over?] countable joins", then f has at least one fixed-point, namely, $0 \cup f(0) \cup f(f(0)) \cup \ldots$. Compare with Zermelo's definition of an infinite set – Sect. 24.2.1, and Dedekind's definition of simple chain – Sect. 9.1.

[27] So $u \in \{x \mid f(x) \geq x\}$ and the definition of u is impredicative.

[28] For every fixed-point u', $f(u') = u'$ so $f(u') \geq u'$ so $u' \in \{x \mid f(x) \geq x\}$. Since u is ≥ every fixed-point of f and u is a fixed-point, the join of all fixed points is u, and (4) follows.

Now[29] let Y be any subset of P. The system $\mathfrak{B} = \langle[\cup Y, 1], \leq\rangle$ is a complete lattice.[30] For any $x \in Y$ we have $x \leq \cup Y$ and hence
$x = f(x) \leq f(\cup Y)$; therefore $\cup Y \leq f(\cup Y)$. Consequently, $\cup Y \leq z$ implies
$\cup Y \leq f(\cup Y) \leq f(z)$.[31] Thus, by restricting the domain of f to the interval $[\cup Y, 1]$, we obtain an increasing function f' on $[\cup Y, 1]$ to $[\cup Y, 1]$.[32] By applying formula (5) established above to the lattice $\mathfrak{B}$ and to the function f', we conclude that the greatest lower bound v of all fixed-points of f' is itself a fixed-point of f'. Obviously, v is a fixed-point of f, and in fact the least fixed-point of f which is an upper bound of all elements of Y; in other words, v is the least upper bound of Y in the system $\langle P, \leq\rangle$.[33] Hence [rather, similarly], by passing to the dual lattices $\mathfrak{A}'$ and $\mathfrak{B}'$, we see that there exists also a greatest lower bound of Y in $\langle P, \leq\rangle$. Since Y is an arbitrary subset of P, we finally conclude that (6) the system $\langle P, \leq\rangle$ is a complete lattice. In view of (4)–(6), the proof has been completed.

The inverse of the Fixed-Point Theorem, namely, that a lattice in which every increasing function has a fixed-point is a complete lattice, was proved, using the axiom of choice, by Tarski's student Anne C. Davis in a paper published next to Tarski's.[34]

Tarski noted that the Fixed-Point Theorem for lattices is a generalization of the set-theoretic Fixed-Point Theorem published in Knaster 1928.[35] He further noted (p 286 footnote 2) that in Knaster 1928, "some applications of this result in set theory (a generalization of the Cantor-Bernstein Theorem) and topology are also mentioned." Actually CBT is not mentioned in Knaster 1928. Perhaps it was mentioned in the session when the paper was presented. Other applications mentioned in Knaster 1928 are to derived sets and the closure of a set in topology. Applications of the Fixed-Point Theorem to the theory of simply ordered-sets and real functions appear in Tarski 1949b §2.

35.6 The Relation $\preceq$

The Schröder-Korselt Theorem mentioned in Sect. 35.3, was ported by Tarski to Boolean algebras in Theorem 9:

[29] Tarski turns to prove that P is a lattice and in fact a complete lattice. He does it by a uniform argument as the following set Y can be finite or infinite.

[30] This was already noted in Sect. 35.1.

[31] Thus $f(z)$ is also in $[\cup Y, 1]$.

[32] Since f is an increasing function in A, f' will be increasing function in $[\cup Y, 1]$ if for every z in $[\cup Y, 1]$, $f(z)$ will also be in $[\cup Y, 1]$. This is provided by the above footnote.

[33] A similar proof is used to obtain that a lattice has a least upper bound to every set iff it has a greatest lower bound to every set.

[34] Anne C. Davis is Anne C. Morel.

[35] We regard the relationship between the two theorems as an example of porting a theorem not of generalization (see Chap. 39). Also we do not see here an example of proof-processing because it is not that certain gestalt and metaphors of one context are used in another context but the entire proof is translated to the new context with necessary changes.

$\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, for any elements a, b $\in$ A the following two conditions are equivalent:
(i) there is an element $a_1 \in A$ such that $a \approx a_1 \leq b$;
(ii) there is an element $b_1 \in A$ such that $a \leq b_1 \approx b$.

Clearly Tarski interpreted Schröder's lemmas (19), (20) as occurring in the context of a specific power-set and thus saw how to generalize these results to Boolean algebras. Proof that (ii) entails (i) follows from Theorem 6(v): b_1 is split into a and b_1-a, and therefore b is split into b' and b'' that are disjoint (have the meet 0) and $b' \approx a$, $b'' \approx b_1$-a. This part of the proof holds for any Boolean algebra. The proof of (i) $\Rightarrow$ (ii), follows from the Mean-Value Theorem (Theorem 8) for complete Boolean algebras: "We consider an arbitrary element a_1 satisfying (i), and we apply Theorem 8 with a, c, a′, c′ respectively replaced by a_1, 1, a, 1." Korselt's reliance on the proof of CBT turns here to reliance on the Mean-Value Theorem, Theorem 8.

For the case when (i) of Theorem 9 holds, Tarski introduced the relation $a \preceq b$; Theorem 9(ii) gives an equivalent formulation when $\mathfrak{A}$ is complete. Using the relation $\preceq$ Tarski reformulated Theorems 8 and 10 as follows:

Mean-Value Theorem. $\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, if a, b, c $\in$ A, $a \leq c$, and $a \preceq b \preceq c$, then there is an element $b' \in A$ such that $a \leq b' \leq c$ and $b \approx b'$.
Equivalence Theorem. $\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, if a, b$\in$A $a \preceq b$, and $b \preceq a$, then $a \approx b$.

The reformulated Equivalence Theorem clearly resembles CBT in its two-set formulation.

The relation $\preceq$ provides a simple way to represent relations of equivalence between Boolean algebras as relations between elements of a certain algebra. Thus Tarski notes (p 305) that every system of Boolean algebras $\langle \mathfrak{A}_i \rangle$ can be represented by means of a system of disjoint elements $\langle a_i \rangle$ of a single Boolean algebra $\mathfrak{A}$, the direct product[36] of all algebras $\mathfrak{A}_i$, in such a way that (i) each algebra $\mathfrak{A}_i$ is isomorphic to the sub-algebra $\langle [0, a_i], \leq \rangle$ of $\mathfrak{A}$; hence (ii) two algebras $\mathfrak{A}_i$, and $\mathfrak{A}_j$, are isomorphic ($\mathfrak{A}_i \cong \mathfrak{A}_j$) iff the elements a_i, and a_j, are homogeneous ($a_i \approx a_j$); (iii) for $i \neq j$, we have $\mathfrak{A}_i \times \mathfrak{A}_j = \mathfrak{A}_k$ iff $a_i \cup a_j \approx a_k$; (iv) $\mathfrak{A}_i$ is isomorphic to a factor[37] of $\mathfrak{A}_k$ if and only if $a_i \preceq a_k$.

As an example, Tarski transported his Theorem 11, described shortly, from elements of a Boolean algebra to Boolean algebras: $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{B}$ being three complete Boolean algebras, we have $\mathfrak{A}_1 \times \mathfrak{B} \cong \mathfrak{A}_2 \times \mathfrak{B} \cong \mathfrak{B}$ if and only if

[36] Which Tarski also call 'cardinal product'. A direct product is a Cartesian product with the Boolean operations defined by coordinates. The image of $\mathfrak{A}_i$, is the sub-algebra [0, 1_i] where 1_i is the element of the product algebra which has 1 at the i coordinate and 0 elsewhere.

[37] This notion does not seem to have been defined in 1949a. It probably signifies here simply a sub-algebra.

$\mathfrak{A}_1 \times \mathfrak{A}_2 \times \mathfrak{B} \cong \mathfrak{B}$.[38] Theorem 11, which is reminiscent of Zermelo's formulation of CBT in 1901 (see Chap. 13), is the following:

> $\mathfrak{A} = \langle A, \leq \rangle$ being a complete Boolean algebra, for any elements $a_1, a_2, b \in A$ the formulas (i) $a_1 \cup b \approx a_2 \cup b \approx b$ and (ii) $a_1 \cup a_2 \cup b \approx b$ are equivalent.
> *Proof*: Obviously, $b \leq a_1 \cup b \leq a_1 \cup a_2 \cup b$ and $b \leq a_2 \cup b \leq a_1 \cup a_2 \cup b$. Hence (ii) implies (i) by Theorem 10. Assume now, conversely, that (i) holds. We clearly have
> $[a_2\text{-}(a_1 \cup b)] \cap (a_1 \cup b) = [a_2\text{-}(a_1 \cup b)] \cap b = 0$ and $a_2\text{-}(a_1 \cup b) \approx a_2\text{-}(a_1 \cup b)$. By Theorem 6 (iv), these two formulas together with (i) imply
> (1) $a_1 \cup a_2 \cup b = [a_2\text{-}(a_1 \cup b)] \cup (a_1 \cup b) \approx [a_2\text{-}(a_1 \cup b)] \cup b$. Since $[a_2\text{-}(a_1 \cup b)] \cup b \leq a_2 \cup b \leq a_1 \cup a_2 \cup b$, we derive from (1), by applying Theorem 10, (2) $a_2 \cup b \approx a_1 \cup a_2 \cup b$. Formulas (i) and (2) obviously imply (ii), and the proof is complete.

The Equivalence Theorem, Theorem 10, the counterpart in Boolean algebra of CBT, is used to prove both directions of Theorem 11. Its use to obtain (ii) ⇒ (i) reverses Zermelo's derivation of CBT in 1901. Tarski needed Theorem 10 also to prove that (i) ⇒ (ii) because, unlike Zermelo, he could not take a_1, a_2, b to be disjoint.

Tarski's way of interpreting relations between Boolean algebras by relations between elements of a Boolean algebra, takes us to Sikorski 1948 (see Chap. 33).[39] In Tarski's terminology, the conditions of Sikorski's Boolean algebra CBT (Theorem 2 in Sect. 33.1), when passing to the direct product of the given Boolean algebras **A**, **B,** imply that $A \preceq B$ and $B \preceq A$, and by his Equivalence Theorem for $\preceq$, we have that $A \approx B$. Passing to the direct product version of Theorem 11 we get that $\mathbf{A} \cong \mathbf{B}$. One can say that Tarski's 1955 trivializes Sikorski's 1948.

35.7 Classifying CBT Proofs by Their Fixed-Points

Banaschewski and Brümmer (BB), in a 1986 paper, took fixed-points to be an implicit basic gestalt in CBT proofs. They thus set out to name the fixed-points in several of the early proofs of CBT (p 3). As a template CBT proof that makes use of Tarski's Fixed-Point Theorem BB suggest the following: Let f:A → B and g:B → A be two injections; define Φ:P(A) → P(A) by $\Phi(X) = A\text{-}g(B\text{-}f(X))$.[40] Φ preserves inclusion, so by Tarski's Fixed-Point Theorem, applied to P(A), which is clearly a complete lattice, there is a $U \in P(A)$ with $\Phi(U) = U$. Then

---

[38] Tarski (1948) posed the question whether this and similar theorems, hold for various algebraic systems. Cf. Sect. 33.1 for a similar question posed by Sikorski and Hanf 1957 for a discussion of the solutions to the problems.

[39] Strangely, Tarski (1955) did not mention Sikorski or his 1948 paper.

[40] We slightly change BB's notations and detail some of their presentation that seems sketchy. With regard to Φ see h in the proof of the Partitioning Theorem above. The definition of Φ appears also in Whittaker 1927 which is not mentioned by BB, who mention Knaster 1928 along with Tarski 1955 and a number of text books.

A-U = A-(A-g(B-f(U))) = g(B-f(U)) and h_U, defined to be equal to f on U and to g^{-1} on A-U, is properly defined and is a bijection from A onto B.

The first proof analyzed by BB is the one due to Dedekind (p 5f, see Sect. 9.2). In this proof B is taken a subset of A and g is the identity mapping. So Φ(X) = (A-B)∪f(X). The minimal fixed-point by Tarski's theorem for Φ is U = ∩{X | (A-B)∪f(X) ⊆ X} = ∩{X | (A-B) ⊆ X and f(X) ⊆ X}. Dedekind indeed defined h as f on U and the identity on A-U.

The second proof analyzed is the one provided by Zermelo in his 1908b paper (p 6, see Sect. 24.1). In this proof B takes the place of A, f(A) the place of B, f reduced to B takes the place of f and g is the identity on f(A). Φ is now defined in P(B) by Φ(X) = (B-f(A))∪f(X) and U = ∩{X ∈ P(B) | B-f(A) ⊆ X and f(X) ⊆ X}. Zermelo indeed defined h as f on U and the identity on B-U.

The third proof scanned by BB is Borel's proof (p 6f, see Chap. 11). Borel also shifted the proof to the case where B ⊆ A but he did not define a set U as did Dedekind and Zermelo; instead, Borel constructed two sequences of frames in each of A and B, with a common residue. He proved that the frames are pairwise equivalent. Hence he concluded that also their union is equivalent (without any argument). Borel did not construct the mapping between A and B! BB's attempt to dress Borel's proof as if he had constructed a set U by taking one of the frames-sequences and defining Φ accordingly has nothing to rely on. In fact, Borel's proof is a counterexample to BB's thesis that in all CBT proofs a mapping combining f and g^{-1} was defined, around a fixed-point of Φ.[41]

BB then (p 7f) discuss the proofs established through the string gestalt that emerged with J. Kőnig's 1906 proof of CBT (see Chap. 21).[42] The string gestalt gives a way to understand the nature of the fixed-points encountered in CBT proofs. In A (or B) an equivalence relation R is defined such that xRy iff for some positive or negative n, $y = (gf)^n(x)$. Alternatively, R can be the ancestral relation, as defined in Principia Mathematica (Whitehead-Russell 1910–13), from the relation $y = (gf)(x)$. In simple terms, an equivalence class of R is simply all the members of A that belong to a J. Kőnig string ('A-string'). Now, the fixed-points in any CBT setting must contain complete A-strings. The A-strings are thus divided into three types: those that contain a member from A-B; those that contain a member from B-gf(A); those that do not contain a member from either of these sets. In the chain terminology the A-strings of the first type belong to the chain of A-B; the A-strings of the second type belong to the chain of B-f(A); the A-strings of the third type belong to the residue of A after removing the frames of the mentioned chains.

A fixed-point must contain either all of the first type A-strings or all of the second type A-strings, and any amount of A-strings of the third type. Therefore, we

[41] Apparently Borel perceived his context as the field of sets generated from A, B by equivalent frames and closed under difference (cf. Tarski 1948 p 98, 1949a p 216) and not as the power-set of A.

[42] Strangely, BB prefer to present J. Kőing's gestalt through its vague use by Korselt (1911, see Chap. 25) instead of J. Kőing's original presentation.

have two possible families of fixed-points, according to the definition of Φ, which follow the way of either Dedekind or Zermelo, as presented above. BB did not make, unfortunately, these remarks. Still they analyzed correctly J. Kőing's gestalt as using Dedekind's Φ, but with the largest fixed-point.[43]

BB's analysis is an example how the light of a new gestalt (fixed-point) directs a re-assessment of old results. However, such re-assessment should not distort the reading of the old results in their context by projecting on them the contours of the new gestalt.

BB's paper continues with generalizations of CBT to various categories and topoi. Though in Chap. 39 we discuss a paper with partly similar objectives, following BB's research program is outside the scope of our book.

[43] By switching to the language of category theory BB claim to have obtained that there is only one natural fixed-point in A. This result seems dubious to us.

Chapter 36
Reichbach's Proof of CBT

In 1955 Marian Reichbach published a one-page paper titled "A simple demonstration of the Cantor-Bernstein Theorem" which we review now. The proof contains very little that is new and one wonders why it was published at all. The journal in which it was published, Colloquium Mathematica,[1] does publish "interesting new proofs of important theorems" but this criterion can hardly fit Reichbach's proof. Still, since the paper contains a direct proof of CBT we have included it in our excursion and it concludes our visit at the Polish school.

In his 1958 paper Reichbach demonstrated, by leveraging on his CBT proof of 1955, that CBT holds in 0 dimensional metric spaces for homeomorphisms under which the diameters of the frames tend to 0. Thus he improved the results obtained under the research project initiated by Sikorski's open problem from 1948 (see Sect. 33.2), in which results by Kuratowski (1950), Kinoshita (1953) and Hanf (1957) were previously obtained. It is outside the scope of this work to detail this result of Reichbach.

Reichbach received his Ph. D. in 1956 under B. Knaster, one of the original editors of the Colloquium Mathematica. In the same year his 3 years younger brother Juliusz Reichbach also received his Ph. D. in logic. They were both born and educated in Wro¢law, Poland, and survived the Second World War, no doubt by concealing their Jewish identity, living in Gorlice, in south Poland. In 1958 they immigrated to Israel. Marian changed his name to Meir Reichaw and was teaching at the Technion in Haifa. His brother used the name Juliusz Podgor and was teaching in Tel-Aviv, perhaps in a high-school. Marian published mainly on topology and Juliusz on logic, though he also published a book on probability (Reichbach 1995; Reichbach 2000).[2]

Reichbach presents CBT (Reichbach uses this name) in the following variation of the single-set formulation:

[1] Or Mathematicum.

[2] I am grateful to Dr. Roman Mańka, see Chap. 31, who provided me with these references.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_36,

> Given a 1–1 (*biunivoque*) function f defined in a set M, such that f(M)⊂M, there exists, for every set E⊂M-f(M), a 1–1 function f* such that f*(M) = E + f(M).

Reichbach's formulation of the theorem differs from the standard single-set formulation in that it does not mention the middle set M′ (see Introduction) but refers to the frame E = M′-M″. With this change it is accentuated that E can be any set floating in M-M″. The gestalt behind this formulation was already present in Zermelo's 1901 proof, but there, because of the language of cardinal numbers used, this gestalt was not immediately exposed. Reichbach's proof is in the language of sets and mappings, avoiding the notion of cardinal numbers. Reichbach's formulation is similar to Zermelo's version of the single-set formulation (see Sect. 24.1) in that there the middle set is taken to be any set contained in M and containing M″. However, it was still a middle set that was the focus of the gestalt and not its frame E as in Reichbach's theorem.

Reichbach continued with the following proof-plan:

> It will be demonstrated that the function (1) f*(x) = x for x ∈ S and = f(x) for x ∈ M-S, where S is defined by (2) S = E + f(E) + f[f(E)] + . . ., satisfies the conditions of the thesis.

Reichbach's S is Dedekind's chain of E (see Sect. 9.1) so the proof is classified as following Dedekind's gestalt.[3] The chain is defined by complete induction and not impredicatively, as it was defined in the proofs of Dedekind and Zermelo. Reichbach does not mention Dedekind, Zermelo or the term "chain".

Reichbach's proof differs slightly in its metaphor, from Dedekind's proof: Dedekind took S to be the chain of M-M′ and *f** to be *f* on S and the identity on M-S while Reichbach takes *f* to be the identity on his S and *f* on his M-S. So the two proofs differ in the definition of f* over the residue of M after removal of the two chains.

> Demonstration. By the hypothesis on E and on f(M), we have f(E) ⊂ M, f[f(E)] ⊂ M and so on. So we have, after (2), S ⊂ M,[4] so (3) M = S + M-S. It follows similarly from (2) that f(S) = f(E) + f[f(E)] + . . .,[5] whence E + f(S) = S, which entails on the one hand
> (4) S + f(M-S) = E + f(S) + f(M-S) = E + f[S + (M-S)] = E + f(M), where the last equality results from (3), and on the other hand
> (5) f(M)-S = f(M)-[E + f(S)] = f(M)-E-f(S) = f(M)-f(S) = f(M-S),[6] where the one before last equality results from the hypothesis on E and the final from the 1–1 character of f.

---

[3] It is not clear to us why Fraenkel (1966 p 77 footnote 1) classified Reichbach's proof with the proofs of Banach and Whittaker that are entirely different (see Chaps. 29 and 31).

[4] Actually it can only be concluded that S⊆M but the case S = M is trivial and Reichbach disregards it.

[5] The frames that constitute S are disjoint because E and *f*(E) are disjoint and thus so are *f*(E) and *f*[(*f*(E))], etc. This observation is not applied in the proof except in (5) where E is omitted at the third step because it is disjoint from *f*(M).

[6] From (4), (5) we can seemingly draw a contradiction: from (5) *f*(M-S) = *f*(M)-S so substitute the right side for the left side in the left side of (4) to obtain S + *f*(M)-S = *f*(M) = E + *f*(M) and hence E = ∅. But the contradiction is a result of an illegal operation: after substitution we actually get S + (*f*(M)-S) and we are not allowed to remove the brackets and change the place of -S, as was

By (1), the function f* is 1–1 in the disjoint sets S and M-S, because f is such in M-S by hypothesis.[7] The sets f*(S) and f*(M-S) are also disjoint, because, after (1), the first coincides with S and the second with f(M-S), hence with f(M)-S because of (5). It follows by (3) that the function f* is 1–1 in the entire set M. Finally, The successive applications of (3), (1) and (4) give

f*(M) = f*(S) + f*(M-S) = S + f(M-S) = E + f(M). QED

Reichbach described his proof as simple. In comparison to Borel's proof it is perhaps simpler because it does not develop two sets of frames and it need not refer to the residue. However, the proof is not simpler than Dedekind's proof. Its algebraic nature is certainly a burden. Another weakness of Reichbach's style is that he likes to make definitions or prove lemmas which appear at first sight arbitrary (for example: the definition of f* before that of S, the introduction of S and the proof of (4) and (5)). In other words: his presentation lacks in heuristics and is very authoritarian and patterned according to the Euclidean style (Lakatos 1976).

In Reichbach's proof the "pushdown" metaphor we associated with Dedekind's proof is avoided because it is concealed under f*(M-S). Therefore the metaphor that we suggest for Reichbach's proof is "the fall of the gown", because as M-S is transferred to f(M-S), S poses untransformed, naked.

---

done in the proof of (4) in the term *f*[S + M-S], because S is not a subset of *f*(M). In general A + (B-C) is not equal to A-C + B.

[7] And so is the identity on S. But to conclude that *f** is 1–1 on M it is still necessary to show that *f**(S) and *f**(M-S) are also disjoint, which is what Reichbach does next.

Chapter 37
Hellmann's Proof of CBT

Marshall S. Hellmann published in the American Mathematical Monthly of October 1961, a short paper (only 26 lines long) with a proof of CBT. The paper came over 60 years after the first publications that provided proofs for the theorem. We could not locate any biographical facts about Hellmann but from the opening sentences of his paper (see Sect. 37.3) it appears that he was teaching mathematics to undergraduate students. His affiliation in the subtitle is to University of Maryland. Hellmann calls the theorem the Schröder-Bernstein Theorem and he brings it in its two formulations. The proof's gestalt is new in the history of CBT, and it provokes some thoughts regarding teaching CBT.

37.1 Theorems and Proofs

The single-set formulation is presented and proved as follows:

> Theorem 1. If A, B, C are pairwise disjoint sets and f a one-to-one mapping from A onto $A \cup B \cup C$, then there exists a one-to-one mapping g from A onto $A \cup B$.
> *Proof*: Let f^k be the composite mapping defined as follows:
> $f^1(x) = f(x)$ for $x \in A$, $f^k(x) = f[f^{k-1}(x)]$ if $f^{k-1}(x) \in A$.
> Let $\bar{A} = \{x \mid \text{for some } k,\ f^k(x) \in C\}$ and $\bar{B} = \{x \mid \text{for all } k,\ f^k(x) \notin C\}$. Then $\bar{A} \cap \bar{B} = \varnothing$ and $\bar{A} \cup \bar{B} = A$. Now define a mapping g on A by $g(x) = x$ for all $x \in \bar{A}$, $g(x) = f(x)$ for all $x \in \bar{B}$. A straightforward verification shows that g is the desired mapping.

There seems to be a lacuna in the definition of $\bar{B}$: If $x \in A$ and $f^k(x) \in B$, x is neither in $\bar{A}$ nor in B because $f^{k+1}(x)$ is not defined. A corrected definition of $\bar{B}$ should be $\bar{B} = \{x \mid \text{for all } k, \text{ if } f^k(x) \text{ is defined}, f^k(x) \notin C\}$, or simply, $\bar{B} = A - \bar{A}$. For the said verification it is necessary to show that g is 1–1 and onto. That g is 1–1 is trivial to prove, as it is the union of two 1–1 mappings on disjoint sets and their ranges are disjoint too; but to prove that g is onto $A \cup B$, some argumentation is necessary: Let y be a member of $A \cup B$. If $y \in B$ let x be the member in A such that $f(x) = y$. x cannot be in $\bar{A}$ because $f^2(x)$ is not defined. Therefore $g(x) = y$ and y is in the range of g. If $y \in A$ then if $y \in \bar{A}$ we have $g(y) = y$ so again y is

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_37,

in the range of g. Otherwise, $y \in \bar{\mathrm{B}}$. Let then x be the member of A such that $f(x) = y$. x is not in $\bar{\mathrm{A}}$ for if for some k, $f^k(x) \in \mathrm{C}$, then clearly $k > 1$ and therefore $f^{k-1}(y) \in \mathrm{C}$ y would have been in $\bar{\mathrm{A}}$ contrary to our assumption. Hence $g(x) = f(x) = y$ and y is in the range of g.

For the two-set formulation, which Hellmann characterizes as the theorem's "usual form", Hellmann gave the following:

> Theorem 2. If f and h are one-to-one mappings such that f maps B onto $D_1 \subset D$ and h maps D onto $B_1 \subset B$, there exists a mapping that maps B onto D in a one-to-one fashion.
> *Proof*: $h(D_1) = A \subset B_1 \subset B$. Therefore $hf(B) = A$ and since A, B_1-A, B-B_1 are pairwise disjoint and $B = A \cup B_1\text{-}A \cup B\text{-}B_1$ we may now employ the preceding theorem to establish the existence of a one-to-one mapping g such that $g(A) = B_1$. Hence $h^{-1}g(A) = D$ and since $A = hf(B)$, $h^{-1}ghf(B) = D$. Therefore $h^{-1}ghf$ is a one-to-one mapping of B onto D.

37.2 Discussion of the Proofs

Hellmann's proof is classified among the inductive, not the impredicative, proofs of CBT, for its use of the number concept and complete induction in the generation of the sequence f^k. The novelty of the proof, for the single-set formulation, is that in it the gestalt of frames and residue is not introduced.

Another novelty in Hellmann's proof is that the situation is viewed from the inner set. Usually, following Cantor, the sets A∪B∪C, A∪B, A are named, e.g., M, M′, M″, and the frames M-M′, M′-M″ are then introduced. The naming of the outer set first reflects its seniority in the gestalt. Hellmann's order of naming reflects the seniority of the innermost set. This seniority is expressed in the definition of the mapping f, which is taken from the inner set onto the outer set, instead of the customary way from the outer set to the inner. Hellmann used the Roman tower gestalt that we identified in the context of Tarski's Mean-Value Theorem (see Sect. 35.3 and Sect. 38.1).

Taking f from the inner set to the outer set provokes the gestalt that a Dedekind infinite set can be characterized as a set that is equivalent to a proper whole-set of it. Along this line CBT is seen as the expansion of the inner set rather than the collapse of the outer set and instead of saying that the collapse can be stopped in halfway, at A∪B, the theorem is now interpreted as saying that the expansion can be stopped in halfway. Some duality is sensed here, reminding us of the convex-concave aspect we discussed with regard to Zermelo's 1901 CBT proof.

Hellmann's gestalt is to partition A rather than A∪B∪C as in all previous proofs of CBT. The partitions are obtained by discerning two types of elements in A, distinguished by the metaphor 'do they have an image in C', and not by unifying certain frames. Here the naming of the frames B, C rather than A∪B and A∪B∪C, becomes efficient.

As in all previous proofs, g is defined by Hellmann as a combination of f and the identity on the two partitions. Hellmann's definition of g as the identity on $\bar{\mathrm{A}}$ is similar to Zermelo's definition but Hellmann includes in the domain of the identity only the chain of C, without the residue as in Zermelo's definition.

The proof of Theorem 2 lacks in style when compared to the proof of Theorem 1. The expression $h(D_1) = A \subset B_1 \subset B$ serves to both define A and state its subset relation to previously defined sets. The use of the letters B and D is awkward and not compatible with the notation of Theorem 1 into which the proof of Theorem 2 glides. Finally, in the application of Theorem 1 the mapping hf is used, which is from the outer set B to the inner set A, contrary to the metaphor of its proof.

37.3 Teaching Concerns

In the first paragraph of the paper Hellmann says:

> The usual proof of the Schröder-Bernstein theorem involves the construction of a mapping by means of an infinite process which many students find hard to grasp.[1] In the following form of the Schröder-Bernstein theorem the desired mapping is defined in an explicit fashion which conceals this process and facilitates an easy understanding of the theorem.

Hellmann's description of the infinite process involved in the definition of *g* in other proofs of CBT surely refers to the process by which the partitioning of A is obtained, the chain of the frames, not the definition of *g* itself, which is usually straightforward once the partitioning is given. Hellmann is correct to note that the said process is concealed, not absent, from his proof because, in fact, $\bar{A}$ is the image-chain of the chain of C.

But is it correct to simplify the proof of CBT when it is presented to students? CBT stands on the border-line between very elementary results in set theory, such as Venn diagrams type of relations, and complex results such as the structure of the scale of number-classes. The background knowledge necessary to introduce students to CBT is small: In a few sentences a novice can be introduced to the concept of set, of equivalence, of infinite sets and to the theorem with its proof. Instead of trying to conceal the intricacies of the proof, it is perhaps better to expand on it. CBT is a great opportunity to demonstrate to the student the drama of a mathematical proof. A proof is not only a means to verify a theorem, it is the location where mathematics happens. CBT provides also an opportunity to teach heuristics and to advocate the need to pull mathematics into general culture.

[1] Mańka-Wojciechowska (1984 p 191) express a similar view.

Chapter 38
CBT and Intuitionism

We review four sources that presented CBT in an intuitionist context, from which certain modes of reasoning, applied in the proofs of CBT, are expelled: A counter-example by Brouwer from his 1912 inauguration address, a theorem by Myhill from his 1955 paper, counterexamples by Dirk van Dalen from his 1968 paper and a theorem by Troelstra from his 1969 monograph.

In 1912, on the occasion of his nomination as an extraordinary professor at the university of Amsterdam, Brouwer gave his inauguration address titled "Intuitionism and Formalism" (Benacerraf-Putnam 1983 p 77). One of the examples that Brouwer brought there, to the divide between formalism and intuitionism, was CBT. In this he may have been following Poincaré (see Chap. 19). Brouwer argued, by an example, that for non-denumerable sets, CBT does not always allow an intuitionist interpretation (pp 87–89). We will present his argument in Sect. 38.1.

For denumerable sets Brouwer said that CBT "is self-evident". Indeed, if it is known that both sets in the conditions of CBT, in its two-set formulation, are denumerable, it is by definition that they are equivalent. If only one of the sets is known to be denumerable, the theorem is reduced to the theorem that an infinite subset of a denumerable set is denumerable – the denumerable CBT (see Sect. 1.2), which is proved by complete induction (Fraenkel 1966 p 35). This proof, however, does not yield the desired mapping constructively. Therefore, when the subject of constructive mathematics was developed, nearly 50 years after Brouwer's paper, his statement on denumerable sets was trialed. In 1955, Myhill gave a proof of a version of CBT for recursive functions. The proof was original in the history of proofs of CBT. We will discuss Myhill's paper in Sect. 38.2.

In his 1968 paper, van Dalen gave examples of denumerable sets, fulfilling the conditions of CBT and having intuitionistically valid partitioning, which fail to exhibit the conclusion of the theorem. Van Dalen's criticism was thus directed not towards the partitioning part of the proofs of CBT, but towards the construction of the 1–1 mapping suggested in these proofs. According to van Dalen's testimony, he was searching to translate to fans (see below) Myhill's result. Only when he failed to find such a proof did he turn to search for counterexamples (p 23). In Sect. 38.3

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_38,

we will present van Dalen's counterexamples. The exposition is unfortunately incomplete due to our failure to grasp certain elements in the original paper.

In 1969, Troelstra, guided by a metaphor derived from Myhill's proof, managed to obtain another original proof of CBT, for lawlike sequences, which he extended to empirical sequences. However, his claim that also Myhill's proof can be adapted to sequences, does not appear to be justified. We will describe Troelstra's results in Sect. 38.4.

38.1 Brouwer's Counterexample

Brouwer begins his discussion of CBT, which he calls, after Poincaré, 'the theorem of Bernstein' (p 87), by presenting its two canonical formulations, the two-set formulation and the single-set formulation[1]:

> 'If A has the same power as a subset[2] of B and B has the same power as a subset of A, then A and B have the same powers' or, in an equivalent form: 'if $A = A_1 + B_1 + C_1$ has the same power as A_1 then it also has the same power as $A_1 + B_1$'.

Brouwer defined (p 84) 'same power' to hold between two sets when their elements can be brought into 1–1 correspondence. This is Cantor's definition in his 1878 *Beitrag*. Though he did not define what 'power' is, Brouwer addressed the denumerable power as an entity, for he remarked (p 84) that "the power of such sets [denumerably infinite] is called aleph-null… [it] is the only infinite power of which the intuitionists recognize the existence". His criticism of CBT then goes hand in hand with his rejection of the notion "sets of higher power" (p 87), and he says that for the theorem to have any intuitionist meaning for such sets, it "will have to be interpretable as follows"[3]:

> If it is possible, *first* to construct a law determining a one-to-one correspondence between the mathematical entities of type A[4] and those of type A_1, and *second* to construct a law determining a one-to-one correspondence between the mathematical entities of type A and those of types A_1, B_1 and C_1, then it is possible to determine from these two laws by means of a finite number of operations a third law, determining a one-to-one correspondence between the mathematical entities of type A and those of types A_1 and B_1.

The notion of 'law' is not systematically introduced in the paper. First the context of the notion is "laws of nature of causal sequences of phenomena" (p 77). Then, for formalist mathematics, the laws are "laws of reasoning" (p 78).

---

[1] In the passages we quote from the paper, Brouwer is using quotation marks, which we reduced to '. Perhaps it was some kind of idiosyncrasy or he was quoting from some source, perhaps from his own earlier writings.

[2] In the 1912 paper Brouwer is still talking about 'sets', which he later termed 'species' (Heyting 1971 p 37, van Heijenoort 1967 p 448, Troelstra 1969 p 14).

[3] Brouwer addresses the second, single-set formulation of the theorem.

[4] It appears that Brouwer is using here 'type' as synonym of 'set' and not for equivalence classes, as the term began to be used in later writings (Rogers 1967 p 51).

Finally, for intuitionism, there are "finite laws of construction" (p 85), from which, as is apparent in the quoted passage, other laws of construction can be obtained. Brouwer is not using here 'law' in the sense he would later, as the defining procedure of spreads (Heyting 1971 p 34f, 39).

Brouwer's criticism of CBT was launched with a quote of a typical proof of the theorem[5]:

> From the division of A into $A_1 + B_1 + C_1$ we secure[6] by means of the correspondence γ_1 between A and A_1 a division of A_1 into $A_2 + B_2 + C_2$ as well as a one-to-one correspondence γ_2 between A_1 and A_2.[7] From the division of A_1 into $A_2 + B_2 + C_2$ we secure by means of the correspondence between A_1 and A_2 a division of A_2 into $A_3 + B_3 + C_3$ as well as a one-to-one correspondence γ_3 between A_2 and A_3. Indefinite repetition of this procedure will divide the set A into an elementary series of subsets $C_1, C_2, C_3, \ldots B_1, B_2, B_3, \ldots$, and a remainder set D. The correspondence γ_C between A and $A_1 + B_1$ which is desired is secured by assigning to every element of C_v the corresponding element of C_{v+1} and by assigning every other element of A to itself.

"In order to test this proof on a definite example", Brouwer takes A to be "the set of all real numbers between 0 and 1, represented in decimal fractions".[8] A_1 is the set of those decimal fractions in which the $(2n\text{-}1)$th digit is equal to the $2n$th digit, B_1 the set of the decimal fractions that are not in A_1 but that still have infinitely many cases of the mentioned equality of digits and C_1 those fractions in which the number of the said occurrences is finite [or 0, presumably]. By replacing every digit of an element of A by a pair of digits equal to it "we secure at once a law determining a one-to-one correspondence γ_1 between A and A_1". But to determine the element corresponding to π-3 according to the correspondence γ_C, it is necessary to decide if π-3 belongs to B_1 or C_1. This, however, is a problem for which there is no ground for believing that it can be solved. "Such belief", Brouwer adds in a footnote (p 88),

---

[5] The proof cited is similar to the one given in Poincaré 1906b (Poincaré's second inductive proof, see 19.2). Like Poincaré Brouwer uses the Roman tower notation (see 35.3) but Brouwer is using different letters to denote the sets. Brouwer does refer explicitly to Poincaré in the context, but to a later publication (1908), from which he quotes a dictum of Poincaré, regarding the schism that divides the different approaches to the foundation of mathematics: "the men do not understand [*s'entendent*] each other, because they do not speak the same language".

[6] The term 'secure' has specific meaning in intuitionism emerging from the fan theorem (van Heijenoort 1967 p 449). Here, however, the term appears to only imply the possibility of the construction of A_2, B_2, C_2, etc., as an alternative to the standard 'define' or 'exist' (compare footnote 3 p 85). Troelstra (1969 p 15) says that a species is secured when it is not empty, but this meaning seems only tangential to our context.

[7] Standards of the language of set theory, such as union, subset and 1–1 mapping, are available in the intuitionist language (Heyting 1971 §3.2.4, 3.2.5), though with different properties. These differences are not relevant to our discussion. Incidentally, stressing that γ_2 has to be constructed implies that restricting γ_1 to A_1 is not an intuitionist act; the law for the construction of γ_1 must be reformulated to give the law for the construction of γ_2.

[8] Brouwer assumes that there are infinitely many digits different from 9 in these decimal fractions (p 85 footnote 4). Actually the definition of A in intuitionist terms is rather complex (Heyting 1971 §2.2.6, 3.3.2) and Brouwer hints at it (p 85). But the argument presented here is independent of the way A is defined.

"could be based only on an appeal to the principle of excluded third,[9] i.e., to the axiom of the existence of the 'set of all mathematical properties', an axiom of far wider range even then the [unlimited] axiom of inclusion [subsets, comprehension]".

It turns out that Brouwer does not speak against the gestalt of the proof, the repeated projection of the partitions of A into A_1; rather he speaks against the very possibility of partitioning A![10] Thus the example[11] does not refute the proof (Troelstra 1969 p 24) and so it does not invoke a Lakatosian proof-analysis to search for a hidden lemma refuted by the example.

38.2 Myhill's CBT for 1–1 Recursive Functions

Theorem[12]*Let α, β be two subsets of N, the set of natural numbers [including 0]. If there is a 1–1 recursive function [Rogers* 1967 *Chap. 1] f such that $x \in \alpha$ iff $f(x) \in \beta$ and a 1–1 recursive function g such that $y \in \beta$ iff $g(y) \in \alpha$, then there is a 1–1 recursive permutation h such that $h(\alpha) = \beta$.*

A recursive permutation is a 1–1 recursive function with range the whole of N. Apparently, Myhill considers recursive functions to be total, namely, defined over all of N. It is for this reason, as will become clear from the proof, that the assumption here, that $x \in \alpha$ iff $f(x) \in \beta$ and $y \in \beta$ iff $g(y) \in \alpha$, is stronger than the assumption in CBT, which is $f(\alpha) \subseteq \beta$ and $g(\beta) \subseteq \alpha$, for it includes $f(\alpha') \subseteq \beta'$ and $g(\beta') \subseteq \alpha'$ (α', β' are the complements of α, β). The stronger assumption enables the definition of h over the entire N, with its definition outside α (β) using f (g) is outside β (α).

Proof: In the proof a sequence $\{(x_i, y_i)\}$, i < k+1, is constructed by complete induction, such that for every $i <$ k+1, $x_i \in \alpha$ iff $y_i \in \beta$, and for every $i, j <$ k+1, $x_i = x_j$ iff $y_i = y_j$. Such a sequence is called "a finite correspondence between α and β". Notice that a finite correspondence between α and β, turns into a finite correspondence between β and α simply by interchanging the places of the x's and the y's. Because of the stated properties of the sequence it follows that the mapping $y_k = h(x_k)$ is 1–1; that h is a recursive permutation will become clear from the proof. Hence, leaning again on the properties of the sequence, it follows that $h(\alpha) = \beta$.

The desired sequence rolls out by setting $x_0 = 0$, $y_0 = f(0)$. Notice that if $0 \notin \alpha$ then $y_0 \notin \beta$, that 0 is a constant and that $f(0)$ is known by the conditions

[9] See a footnote in Sect. 3.1.

[10] It is the touchstone of intuitionism that the partitioning of A cannot be made acceptable by placing in C_1 those members of A which at a given time are not known to belong to B_1 or C_1 (Heyting 1971 p 38f, Troelstra 1969 p 25f, Heyting 1971 p 119).

[11] It can be classified among Brouwer's weak counterexamples. http://plato.stanford.edu/entries/intuitionism/.

[12] This is theorem 18 of Myhill's paper, stripped of the notions not relevant to our focus on CBT, such as 'reducibility'. It is clearly similar to the two-set CBT.

of the theorem, so the first step of the definition of the sequence is recursive. After $(x_0, y_0), \ldots, (x_k, y_k)$ have been constructed recursively, (x_{k+1}, y_{k+1}) is defined as follows:

Case I. k is even. By a lemma to be presented below, applied to $(y_0, x_0), \ldots, (y_k, x_k)$, x_{k+1} can be defined for $y_{k+1} = k/2$, so that the sequence $(y_0, x_0), \ldots, (y_{k+1}, x_{k+1})$ is a finite correspondence between β and α.

Case II. k is odd. By the mentioned lemma, applied to $(x_0, y_0), \ldots, (x_k, y_k)$, y_{k+1} can be defined for $x_{k+1} = (k\text{-}1)/2$, so that the sequence $(x_0, y_0), \ldots, (x_{k+1}, y_{k+1})$ is a finite correspondence between α and β.

The sequence of x's and the sequence of y's, both cover N. Indeed, for any $n \in$ N, $x_{2n+2} = n$, and $y_{2n+1} = n$. Thus h is a recursive permutation, provided the construction in the lemma is assured to be recursive.

Lemma[13] *Given a 1–1 recursive function f*[14] *such that $x \in \alpha$ iff $f(x) \in \beta$, a finite correspondence $(x_0, y_0), \ldots, (x_k, y_k)$ between α and β and any number m, one can effectively find a number n such that $(x_0, y_0), \ldots, (x_k, y_k), (m, n)$ is likewise a finite correspondence between α and β.*

Proof: Define $\xi = \{x_0, \ldots, x_k\}$; $\eta = \{y_0, \ldots, y_k\}$; $\varphi(y) = x_i$ if $y = y_i$,[15] otherwise, namely when $y \notin \eta$, $\varphi(y)$ is left undefined; $a_0 = f(m)$, $a_{i+1} = f(\varphi(a_i))$ if $a_i \in \eta$, otherwise, namely when $a_i \notin \eta$, a_{i+1} is undefined.

Let us assume first that $m \notin \xi$. If there are more than k+1 a_i's, it must be that there are repetitions among the a_i, namely, there are i, j, $i < j <$ k+1, such that $a_i = a_j$. The composite function $f\varphi$ being 1–1, we obtain that, if $i > 0$, $a_{i\text{-}1} = a_{j\text{-}1}$. By repeating this argument i times we obtain that $a_0 = a_{r+1}$, $r \geq 0$. But $a_0 = f(m)$ and $a_{r+1} = f(\varphi(a_r))$ so $m = \varphi(a_r)$. Because $\varphi(a_r) \in \xi$ and $m \notin \xi$, this is a contradiction. So the assumption that there are more than k+1 a_i's must be rejected. Let then a_l be the last of the a_i's. Clearly $a_l \notin \eta$, for else a_{l+1} would be defined.

The sequence $(x_0, y_0), \ldots, (x_k, y_k)$ can be given the following Gödel number: $p_k = \Pi_{i=0, \ldots, k} p(i)^{2^{x_i} 3^{y_i}}$.[16] Now a function t can be defined as follows: $t(m, p_k) = a_l$ if $m \notin \xi$; $t(m, p_k) = y_i$ if $m = x_i$. Evidently t is a partial recursive function. We claim that $t(m, p_k)$ is the n sought in the lemma. To prove this contention we need to demonstrate that the properties for finite correspondence sequences hold for $(x_0, y_0), \ldots, (x_k, y_k), (m, t(m, p_k))$.

To prove that for every $i \leq k$, $x_i = m$ iff $y_i = t(m, p_k)$, we note first that if $m \notin \xi$, $x_i \neq m$ and $t(m, p_k) = a_l \notin \eta$, so $y_i \neq t(m, p_k)$. Second, if $m \in \xi$, say $m = x_j$, then $t(m, p_k) = y_j$ by definition.

To prove that $m \in \alpha$ iff $t(m, p_k) \in \beta$, we first assume that $m \notin \xi$ and thus $t(m, p_k) = a_l$. We have $m \in \alpha$ iff $a_0 = f(m) \in \beta$, by the nature of f. If $a_0 \in \eta$

---

[13] We have grouped in this lemma Myhill's theorem 17 and its lemmas A, B.

[14] When the lemma is called from the theorem in case II the f is taken from the conditions of the theorem; in case I it is g from the conditions of the theorem.

[15] Myhill denoted φ by g which could lead to confusion.

[16] $p(i)$ is the ith prime number, 2 being the 0 prime number.

then $a_0 \in \beta$ iff $\varphi(a_0) \in \alpha$, as this property holds by the induction hypothesis for the sequence $(x_0, y_0), \ldots, (x_k, y_k)$ that this sequence has the discussed property. But then $\varphi(a_0) \in \alpha$ iff $a_1 = f(\varphi(a_0)) \in \beta$, again by the nature of f. In summary, $m \in \alpha$ iff $a_1 \in \beta$. The same mode of reasoning leads us to conclude that $m \in \alpha$ iff $a_l \in \beta$, namely, $m \in \alpha$ iff $t(m, p_k) \in \beta$.

Now assume that $m \in \xi$ and say $m = x_j$. Then $t(m, p_k) = y_j$ and $m \in \alpha$ iff $t(m, p_k) \in \beta$ by the induction hypothesis for the sequence $(x_0, y_0), \ldots, (x_k, y_k)$.

We have thus completed the proof of the lemma and with it the proof of the theorem.

What are the proof-processing origins of Myhill's ingenious proof we do not know. The sequence of finite correspondence is built by a back and forth argument (metaphor). Every n appears at stage $2n$ as the left side of the $2n$ couple of the sequence and as the right side of the $2n$+1 couple. The impression is of a double decker tower that is being constructed inductively. This seems to us to have been the gestalt that guided the proof. It is possible that further analysis will reveal how Myhill's construction charts a path through all of J. Kőnig's strings built by f, g.

Another possible act of proof-processing lies perhaps in the following: According to J. Kőnig's gestalt, given f, g, N is partitioned into equivalence classes of the relation 'linked by a finite string of images under f or g, in alternating order'. There is thus some similarity of N so partitioned and the classical matrix of rational numbers. The metaphor of Myhill's proof may then be grasped as an enumeration of J. Kőnig's partitions, much as was Cantor's enumeration of the rationals. Anyway, this construction of a path that winds among the equivalence classes to linearize them, is the basic metaphor of Myhill's proof.[17]

38.3 Van Dalen's Counterexamples to CBT for Fans

Van Dalen's counterexamples were constructed at the heartland of intuitionism: the theory of fans. Like the paper that we survey we refer the reader to Heyting 1971 for a systematic presentation of intuitionism. We have also used Troelstra 1969 and Troelstra and van Dalen 1988. We will introduce the intuitionist notions that we use in shorthand only.

[17] We call attention here to results of Remmel 1981 among which is the claim that Banach's partitioning theorem is not effective, contrary to CBT. The results of Remmel are obtained "using the tools of recursion theory". It seems that Remmel's result implies that the fixed-point theorem, which implies Banach's theorem, is likewise not effective. It is beyond the scope of this book to follow on Remmel's results.

38.3.1 Basic Notions

A tree is a non-empty set (or species in proper intuitionist terminology) of finite sequences of natural numbers such that for every finite sequence n it is decidable if n belongs to the tree (is admissible) or not and for every n in the tree all its segments, including the empty segment $<>$, are in the tree. The sequences in a tree are called nodes and the numbers in a sequence are called components. The length of a node n we denote by $l\mathrm{th}(n)$. If $i < lth(n)$ we denote by $n|i$ the segment of n of length i. For $0 \leq i < lth(n)$ we denote by $n(i)$ the ith component of n. Note that the last component of $n|i$ is $n(i\text{-}1)$. Of course, $n(\text{-}1)$ does not exist and $n|0$ has no components.

A spread is a tree in which every node has a continuation in the tree. A spread is never given in its entirety; it is a process. A branch is a sequence of compatible nodes. Like the tree, a branch is a process in progress from node to node. Thus a branch is also called an infinitely proceeding sequence (ips, we use ips both singular and plural). Though the ips are never given in their entirety we refer to them by variables, such as α, β, γ, and we use the notation $\alpha|i$, $\alpha(i)$ for the segment of length i and ith component of α. In so doing we assume, no doubt, that these parts of the ips, which we reference, are already available when referenced.

A fan is a spread with finite branching at each node. A subfan is simply a subset of a fan. A (proper) subfan is 'removable' (or 'detachable') when it is decidable, for every node in the fan, whether it belongs to the subfan or not. The subfan of all nodes of a fan, which have a segment in a finite subset A of the fan, is clearly removable. It is the subfan determined by A. The opposite observation, that for every removable subfan there is a finite set of nodes, called a 'bar', such that every node in the subfan has a segment in the set, is a fundamental observation in the theory of fans, a variant of the Fan Theorem. It is justified by noting that decidability must rest upon finite data.

A fan F can be provided with a topology. For every node n, let V_n be the set of all nodes that continue n. Taking the V_n (including the empty set) as a basis we obtain a topology for F. The removable subfans of F are thus compact sets. The fan itself is compact because it is $V_{<>}$.

The binary fan, BF, is the set of all nodes with components in {0, 1}. For a natural number i, we write i^* for the set of all finite sequences of i, including the empty sequence.[18]

Between two fans there can be mappings just as between sets. However, an intuitionist mapping must be given by a finite law. Therefore, the image or source by a 1–1 mapping, of a removable subfan, is also removable. We simply combine the law that determines the removable subfan with the law giving the mapping (or its inverse).

[18] * is called the Kleene star (http://en.wikipedia.org/wiki/Kleene_star).

38.3.2 The First Counterexample

The first of van Dalen's counterexamples is the following: Let G be the fan that consists of BF and an additional ips. For example, the ips whose set of nodes is 2* which will be denoted by ζ.[19] BF and G, claims van Dalen, give an example where CBT fails.[20] Indeed, an injection g:G → BF is given by assigning to ζ the ips in BF whose nodes are in 0*, and to every BF ips ξ of G, the ips of BF obtained from ξ by appending to it 1 at the beginning ($<1>*\xi$).[21] The definition of g is given by a finite law so it is constructive. Thus the conditions of CBT in its single-set formulation, holds between G and BF. However, the conclusion of CBT does not, for if we assume that a constructive bijection h:G → BF exists, we obtain a contradiction[22] as a result of the Fan Theorem: the set $\{\zeta\}$ forms a removable subfan in G. The decision law is simple: an ips belongs to $\{\zeta\}$ iff its 0 component is 2. Thus the set $\{h(\zeta)\}$ must also be removable in BF. Thus there is a finite set A = $\{a_1, \ldots, a_n\}$ such that an ips belongs to $\{h(\zeta)\}$ iff it has a segment in $\{a_1, \ldots, a_n\}$. But clearly this is a contradiction for there are infinitely many ips with segments in A, while in $\{h(\zeta)\}$ there is only one ips.

A question that arises with respect to van Dalen's first example is whether it can be given in the form of Brouwer's counterexample discussed in Sect. 38.1. The answer is positive, for let G be partitioned into three sets: $\{\zeta\}$, BF-g(G) and g(G). Let then ξ be the ips with all its components 0 except that its nth component is 1, where n is the first place in the decimal expansion of π where a sequence of 10 consecutive 9s ends. It is undecided whether ξ belongs to the second or the third partition and thus whether in the definition of the 1–1 mapping between G and BF it should be defined by g or by the identity.

Van Dalen also claimed that the example with BF and G is a counterexample to the dual CBT.[23] Indeed, if surjections can be defined between BF and G, the impossibility of a bijection between these fans would entail that this is a counterexample to the dual CBT for fans. Such surjections are the following f', g': g' assigns every ips of BF in G, to itself in BF, and ζ, to the ips in BF whose segments are in 0*; f' assigns the latter ips to ζ, every ips which begins with a finite segment from 1* and

[19] van Dalen does not specify the additional ips of G.

[20] van Dalen did not detail an argument to support his claim.

[21] Here * stands for concatenation.

[22] Intuitionists may assume the existence of a mathematical entity to derive a contradiction. But from the contradiction, if it is derived under the assumption of a negation of a certain assertion, it is not the assertion that can be concluded but only the negation of the negation. Also, proving the existence of a certain mathematical entity without demonstrating a construction for it, is not acceptable in an intuitionist proof.

[23] The dual CBT asserts that when two sets are correlated by surjections there is a bijection between them. The dual CBT is proved in the general case, using inverses of the given surjections, defined via the axiom of choice (Banaschewski-Moore 1990).

continues with zeros only, to the ips obtained by omitting the first 1, and every other ips of BF to itself.

The above counterexample may give the impression that a fan can never be equivalent to a subfan. Designate by ζ_n the ips whose components are all 0 except the first n which are 1. Then the fan of all ζ_n with $n \geq 1$, is equivalent to the same fan less ζ_1. In classical terms the equivalence is obtained by mapping ζ_n onto ζ_{n+1} but from an intuitionist point of view it is rather the mapping from ζ_{n+1} to ζ_n that has to be taken because in this way we do not get lost in an infinite process.

In the previous example, however, the subfan is removable in the fan, while it is to be noticed that g(G) is not removable in BF. For otherwise, there would be a finite set of nodes from BF, A = $\{a_1, \ldots, a_n\}$, such that an ips of BF would belong to g(G) iff it has a segment in A. Since the ips with nodes in 0* belongs to g(G), there are members of A that have all of their components 0. Let k be the length of the longest such member of A. Then any ips in BF with k leading 0s and 1 at its kth place would have a segment in A but would not belong to g(G), by the definition of g, contrary to the assumption that A is a bar of g(G). So g(G) is not removable in BF. Thus the question remains after the first counterexample is exhibited, whether two fans can be exhibited that are each mapped 1–1 onto a removable subfan of the other, while still the two fans are not equivalent. Van Dalen answers this question positively in another example to be discussed next.

38.3.3 The Main Counterexample

Let H_n ($n > 1$) be the fan whose nodes are either from BF or from 1*2*... of degree n.[24] 1*2* is the set of all finite sequences obtained by concatenating to a sequence from 1* a sequence from 2*. 1* is 1*2*... of degree 1; 1*2* is 1*2*... of degree 2. 1*2*1* is the set of all finite sequences obtained by concatenating to a sequence from 1* a sequence from 2* and then again a sequence from 1*. It is of degree 3. In a likewise manner, 1*2*... of any degree n is constructed. The nodes of H_n can be classified as nodes of type 0, if they belong to BF, and nodes of type i, $1 \leq i \leq n$, if they belong to 1*2*... of degree i. In H_n there are no nodes that contain both 0 and 2.

Lemma *If $n \neq m$ then it is impossible to have an injection f:$H_n \rightarrow H_m$ such that $f(H_n)$ would be removable in H_m.*

Proof:[25] Let us first assume that $n > m$.

[24] Van Dalen uses 'length' for what we call 'degree'.

[25] Van Dalen gave only a sketch of the proof and only for $n = 2$, $m = 3$.

A node of type n from H_n has only one ips continuation. Since this ips is a removable subfan in H_n, its image must be removable in H_m. Therefore the nodes in this image must belong to one ips and must contain therefore a node of type m.

A node ξ with type n-1 in H_n has denumerably many continuations with type n and only one ips with nodes only of type n-1. It is a removable subfan and so its image in H_m must be removable too. In this image are contained all the images of the continuations of type n of ξ in H_n and it must contain one single ips that corresponds to the continuation of ξ with 1 only. Therefore to ξ must correspond a node of type m-1. Continuing in this way we find that every node of type n-i is mapped to a node of type m-i, with $0 \leq i \leq m\text{-}2$.

Now a ξ node of type n-m + 1 with all its continuations is a removable subfan in H_n and thus so must be its image in H_m. The images of the ips continuing ξ with larger types have already been identified, except the image of ξ that is continued by 1s only. So in the image of the neighborhood V_ξ there must be contained a single ips other than the already assigned ips. But this is impossible because the only nodes available in H_m are from BF and those have branching not previously assigned. So the possibility of an injection from H_n into H_m is refuted.[26]

Let us next assume that $n < m$. Let $F = f(H_n)$ be a removable subfan of H_m.

If there is a node ξ in F with type $n > 1$, then there are in F nodes with types j for every $1 < j \leq m$. In this case we can obtain a contradiction as before, for: Let ξ be a node of type m in F, it has only one ips in H_m and the set of the nodes of this ips is removable in H_m. Hence it is the image of a removable subfan of H_n. Hence this subfan must contain only the nodes of one ips and so it must contain a node of type n. Therefore all the nodes of F of type m must be the images by f to nodes of H_n of type n. Continuing in this way we find that every node of type m-i in F is the image of a node of type n-i, with $0 \leq i \leq n\text{-}2$. Clearly if we continue with nodes of F of type $m\text{-}n + 1$ we would obtain a contradiction as before.

If, on the other hand we have no nodes of type $n > 1$ in F, then F must contain only nodes from BF. This, however, is not possible: In H_n the ips of 1* has nodes that branch both into 1*2*... and into BF. Such a separation of any ips in BF is not possible.[27]

For the next step of the contemplated construction it is required to take disjoint unions[28] of a family of fans. So van Dalen introduces two operations for fans' unions:

[26] Nodes of type $n-1$ are accumulation points of nodes of type n, and so on for lower types nodes. The metaphor of the proof is similar to the generation of derived sets, perhaps a case of proof-processing.

[27] The argument here appears quite clear intuitively but would benefit from a rigorous makeover. Unfortunately, such was not provided by van Dalen and we did not have the resources to supply it.

[28] http://en.wikipedia.org/wiki/Disjoint_union.

- For a sum of two fans: Let S, T be two fans. W = S + T will be the fan with the following nodes: If $< x_1, \ldots, x_n >$ is a node of S (T) then $<2x_1, \ldots, 2x_n>$ ($<2x_1 + 1, \ldots, 2x_n + 1>$) will be a node of W. Note that S + T ~ T + S, namely, a bijection can easily be constructed between S + T and T + S. Thus we can say that + is commutative. Furthermore, R + (S + T) ~ (R + S) + T, so that + can be described as associative.
- For a sum of a sequence of fans (fan-sum): Let $\{F_i\}$ be a sequence of fans, then $\{F_i'\}$ will be the sequence of fans with F_i' composed of the ips generated from the ips of F_i by increasing each of its components by 1 (thus no zero appears at the beginning of the ips) and appending to its beginning i zeros.[29] The union of the F_i' is the fan-sum of $\{F_i\}$, denoted by $FS(F_1, F_2, \ldots)$. Note that by commuting the order of a finite number of the fans in a fan-sum, or an infinite number according to a law (say commuting even and odd elements of the sequence), we obtain an equivalent fan (namely, there is a bijection between the two fans). We will refer to this property of FS as commutativity. Note further that $FS(F_1, F_2, \ldots) \sim F_1 + FS(F_2, \ldots)$. Note that in a fan-sum, where the F_i include the empty node – as is the case with the H_n, a new ips is added to the transformed ips of the F_i: it is the ips with all its components equal to 0. Van Dalen calls it the spine of the fan-sum.[30] The metaphor is clear: all the ips of the fan-sum split off from the spine at a certain node.

Now van Dalen defines the following fans:

$A_p = FS(H_p, H_{p+1}, H_{p+2}, \ldots)$
$B_p = FS(H_p, H_{p+2}, H_{p+4}, \ldots)$
$C = FS(A_1, A_1, A_1, \ldots)$
$D_0 = FS(B_1, B_2, A_3, B_3, B_4, A_5, \ldots, B_{2k-1}, B_{2k}, A_{2k+1}, \ldots)$
$D_1 = FS(B_2, A_3, B_3, B_4, A_5, \ldots, B_{2k-1}, B_{2k}, A_{2k+1}, \ldots)$[31]
$D_2 = FS(A_3, B_3, B_4, A_5, \ldots, B_{2k-1}, B_{2k}, A_{2k+1}, \ldots)$
$F = C + D_0$
$K = C + D_1$
$G = C + D_2$

[29] Apparently, i is the place of the fan in the list and not its index in the original family. This means that if we take the fan-sum of a sequence from which some of the terms were omitted, the ips from the remaining terms will get a different number of leading zeros from what they would have in the original family.

[30] van Dalen introduces the spine by the following: "for convenience let us call $\lambda x[0]$ the spine of F". $\lambda x[0]$ must be a form of the lambda expression for the constant function $x = 0$; it conveys that for every n the nth member of the nth node, is 0.

[31] There seems to be a typo here in the original, where instead of our B_3 and B_4, B_4 and B_5 appear. Also, what we denote by D_1 is denoted in the original by D_2, and vice versa. Our reason for the switched notation is that D_1 and D_2 are obtained from D_0 by omitting first B_1 and then B_2. The D after the first omission is added to C to obtain K and the D after the second omission is added to C to obtain G. To us it seems more reasonable to call D_1 the D after the first omission and D_2 the D after the second omission. With regard to the formation of K, G we stayed aligned with van Dalen.

Note that in F, K, G, the spines of the A's grouped in C will have their nodes in 2*; the new spine of C will have its nodes in 0*. The spines of the A's and the B's grouped in the D's will have their nodes in 3* and the new spines of the D's will have their nodes in 1*.

Van Dalen notes that G is a removable subfan of K and K a removable subfan of F. This is not exact because in D_1, for example, there is one less leading 0 in its ips, compared to corresponding ips in D_0. The correct statement is that an injection can be constructed that takes D_1 into D_0. Likewise an injection can be constructed that takes D_2 into D_1. Using these injections, also injections from G into K (φ) and from K into F (ψ) can be constructed. The ranges of these injections are indeed removable in their target fans because we can construct a law that decides for every ips in F, for example, if it belongs to ψ(K) or not: If the ips belongs to B_1, as transformed in the formation of F, then it is not in ψ(K); otherwise it is in ψ(K). The ips of B_1 of D_0 are the only ips of F that begin with a single 1. So we have constructive law regarding the decision of whether an ips of F belongs to ψ(K) or not (Fig. 38.1).

It is easy to prove that F ~ G:

- First note the following equivalences that follow from the commutativity and associativity of +:

$$\begin{aligned} B_{2k-1} + B_{2k} + A_{2k+1} &\sim H_{2k-1} + B_{2k+1} + H_{2k} + B_{2k+2} + A_{2k+1} \\ &\sim A_{2k-1} + B_{2k+1} + B_{2k+2}. \end{aligned}$$

- Setting $D_0' = FS(A_1, B_3, B_4, A_3, B_5, B_6, A_5, \ldots, A_{2k-1}, B_{2k+1}, B_{2k+2}, \ldots)$ it is easy, based on the first step, to construct a bijection between D_0 and D_0' and hence F ~ C + D_0'.
- Extracting A_1 from D_0' we get:

  $F \sim C + A_1 + FS(B_3, B_4, A_3, B_5, B_6, A_5, \ldots, B_{2k+1}, B_{2k+2}, A_{2k+1}, \ldots)$.
- Now we can absorb A_1 into C since C + A_1 ~ C by the argument we used above with regard to the sequence ζ_n.[32]
- So we get: $F \sim C + FS(B_3, B_4, A_3, B_5, B_6, A_5, \ldots, B_{2k+1}, B_{2k+2}, A_{2k+1}, \ldots)$.
- By shifting the A_{2k+1} two places to the left using the commutativity of the fan-sum, we obtain an equivalent fan, so finally,

  $F \sim C + FS(A_3, B_3, B_4, A_5, B_5, B_6, A_7, \ldots, A_{2k+1}, B_{2k+1}, B_{2k+2}, \ldots) = G$.

[32] The family of fans unified by FS to obtain C remains unchanged if another A_1 is added to it in the beginning. Alternatively, one can argue that the new C is seen to be equivalent to the old C by mapping the new A_1 to the first old A_1, this old first to the old second and so on. In intuitionistic terms the mappings should go the other way: the first old A_1 to the new A_1, the second old to the first old and so on, beginning the process with an arbitrary *n* but staying always within finite reasoning. C thus acts as a buffer. The construction of C goes back to Dedekind's chain theory and its "pushing down the chain" metaphor (see Sect. 9.1). Proofs by the first argument, that C remains unchanged with a new A_1 appended to its beginning, appear in ordinal arithmetic (see Chap. 32).

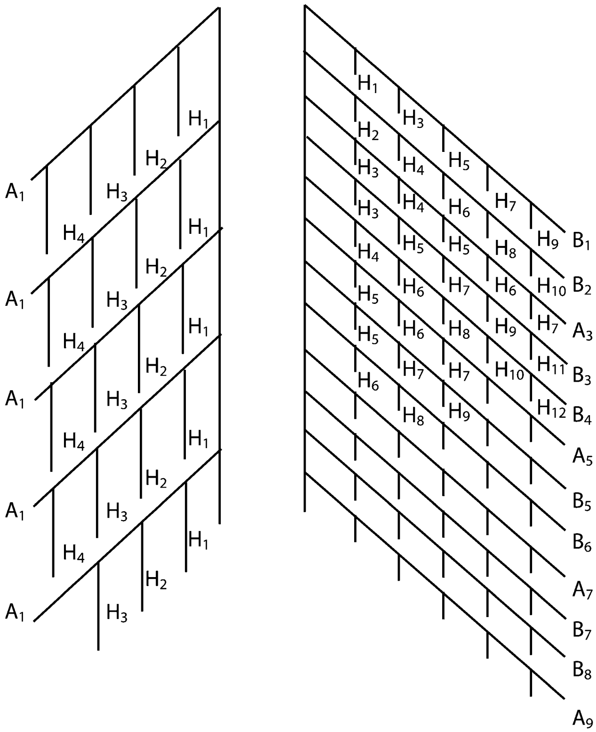


Fig. 38.1 van Dalen's drawing of F

Having obtained that the conditions of CBT in its single-set formulation, hold between F, K, G, with the additional property that G is removable in K and K in F, van Dalen now claims that F cannot be equivalent to K.[33] For otherwise, if we assume that a bijection *f* exists between F and K, a contradiction can be obtained. Van Dalen says that the proof applies often the fan theorem and as "the actual working out of the details is rather tedious and uninteresting we will omit them. Any reader familiar with the fan theorem can readily supply the missing details." As the reader will see our success was rather limited.

The proof rests on the following leammas which van Dalen did not prove:

Lemma 1. *The spine of a subfan A_i is [necessarily] mapped [by f] onto the spine of another subfan A_j.*

[33] van Dalen's result is reminiscent of the results obtained for Boolean algebras and topological spaces (see Sect. 33.2). It would be interesting to learn whether there is internal similarity to all these counterexamples; perhaps also a connection with the topological notion of dendrite. However, such study is outside the scope of this work.

Proof: We go to our topological point of view again[34]: All spines of A_i have the same neighborhood, from a certain segment on. No other ips of K or F have the same neighborhood, so the lemma holds.

Lemma 2. *The spine of a subfan B_i is [necessarily] mapped [by f] onto the spine of another subfan B_j.*

Proof: Same proof. Notice that both odd and even B_i's skip an H_n so they are of a similar structure in that, but it appears to make more sense that the spines of odd (even) B_i's would correspond to each other and the two types of B_i's will not be mixed by *f*.

Lemma 3. *The spine ξ of D_0 is mapped onto itself.*

Though in K, D_1 replaces D_0 of F, the spine of both D_0 and D_1 is the same, having its nodes in 1*.

Proof: Again, it is the topology that explains the claim. The two spines have the same neighborhoods after the first node of D_0 and they are the only ips with these neighborhoods.

Van Dalen ignores the spine of C, which, obviously, must be mapped to itself as well.

Lemma 4. *On the strength of Lemma 3, we can find a node b of ξ such that every ips through b is mapped onto an ips through the topnode a of ξ and likewise there is a node c of ξ such that every ips through c is mapped onto an ips through b. It is no restriction to suppose that b dominates c.*

Proof: The neighborhood attached to any node of ξ in F must be mapped to the neighborhood attached to some node of ξ in K. Apparently, *b* is the node in F that has its neighborhood mapped to the neighborhood of the topnode of K; *c* then is the node in F that has its neighborhood mapped to the neighborhood of *b* in K. If *c* dominates *b*, *f* can be changed so that *c* is mapped to *a* and *b* to *c*. Thereby *c* will become *b* and *b* will become *c*. Therefore we can, without loss of generality, assume that *b* dominates *c*.[35]

Lemma 5. *By Lemma 1, we can find for every subfan A_i a node a_i such that every ips through a_i, is mapped to an ips through a fixed node of the corresponding subfan both by f and f^{-1}. A similar lemma holds for the B_i's with nodes b_i's.*

Proof: If A_i corresponds by *f* to A_j then by Lemma 1 there is a node from which on, the two subfans have the same neighborhood. Likewise for f^{-1} which makes correspond to A_i some A_j'. The longer of the two nodes is the required node a_i. Similarly for B_i by Lemma 2.

[34] Why is topological heuristic advantageous in this context? Cf. Bernstein's use of abstraction in his proof of CBT (see Sect. 11.2) and our remark towards the end of Sect. 35.2.

[35] A node dominates another if it precedes it in the fan. The topnode is the first node in the fan.

We have now reached the final stage of the construction. We quote van Dalen in full:

> Now consider all B and A spines through nodes of ξ that dominate c. On each of the B spines locate the first subfan H_{ki} that passes through b_i and on each of the A spines locate the first subfan H_{ki} that passes through a_i. Thus we effectively determine a finite species of indices k_i with the required property. Let k be the maximum of this species, suppose k is odd, otherwise take k + 1. Consider the species S of subfans H_k that are dominated by b and the subfans H_k of the subfans B_i. This clearly is a finite species and we can determine the images of the concerned subfans under f. Since f maps the B-subfans onto B-subfans and K does not contain the subfan B_1[,] S cannot be mapped onto itself. Owing to the fact that f is bijective there are less subfans H_k mapped into S from the complements[36] of S than there are mapped from S into its complement.
>
> By (iv) and (v) [Lemmata 4, 5] the only possible subfans to be mapped from S into its complement are those dominated by b and not by c. Call this species of subfans $S_c{}^b$. The members of $S_c{}^b$ which are mapped into the complement of S must by (v) [Lemma 5] be subfans of A_i-subfans.
>
> Also the members of $S_c{}^b$ which occur as images of subfans from the complement of S are subfans of A_i-subfans. Hence the analogous species $T_c{}^b$ of subfans H_{k+1} has the same properties as $S_c{}^b$, i.e., there are more elements of T (the species of H_{k+1} analogous to S) mapped into the complement than from the complement of T into T.
>
> This clearly contradicts the fact that all members of T must occur as images under f.
>
> Hence we showed that no bijective f exists such that f:F → K.

It is outside our resources to explicate this passage.

38.4 Troelstra's CBT for Lawlike Sequences

Troelstra's main result (1969 p 103) is the following lemma (16.7.1) which is clearly a variant of CBT in its two-set formulation:

> Lemma. Let a, b be two lawlike sequences such that there exist bi-unique mapping φ, ψ satisfying
>
> φ[Range(a)] ⊆ Range(b), ψ[Range(b)] ⊆ Range(a).
>
> Then there exists a bi-unique ξ defined on Range (a) such that
>
> ξ[Range(a)] = Range(b).[37]

The terms used in the statement of the lemma are defined by Troelstra as follows:

- "N denotes the collection of natural numbers [including 0]" (p 13).
- "Species may be regarded as the intuitionist analogues of the classical sets" (p 14).
- "A mapping φ from a species X into a species Y is any kind of process which assigns to any $x \in$ X a $y \in$ Y"[38] (p 16).

---

[36] Here van Dalen has a footnote: "relative to the species of all subfans H_k".

[37] Troelstra adds that "if ξ is taken to be defined on Range(a) only, Range(ξ) = Range(b)" which means that in general ξ may be defined for other values not in Range(a).

[38] Troelstra adds: "and such that $x = x' \rightarrow \varphi(x) = \varphi(x')$. (This stipulation is necessary whenever = does not denote basic definitional equality, e.g., in the case where the elements of X are

- "A mapping φ from X into Y is said to be of type (X)Y (also $\varphi \in$ (X)Y)" (p 16).
- "A sequence is a mapping of type (N)X" (p 16).
- A lawlike sequence is "a sequence which is completely fixed in advance by a law, i.e., a prescription (algorithm) which tells us how to find for any $n \in$ N the nth member of the sequence" (p 17).[39]
- "A mapping $\varphi \in$ (X)Y is said to be bi-unique (an injection) if $(\forall x \in \text{X})(\forall x' \in \text{X})(\varphi(x) = \varphi(x') \rightarrow x = x')$" (p 16).[40]
- Range(a) is not explicitly defined but appears to be the species of all entities that are the image of a natural number under the sequence a.
- As an example of a lawlike sequence Troelstra mentions (p 17) "primitive recursive functions" but from Rogers (1967 p 1) it appears that any recursive function is a lawlike sequence.

The reason the Lemma stated above is not trivial is that a, b are not necessarily bi-unique, so the mapping $a(n) \rightarrow b(n)$ is not necessarily coherent: if $a(n) = a(n')$ not necessarily $b(n) = b(n')$. The metaphor of the proof is to ramify the ranges of the sequences so that a proper bi-unique mapping can be defined between them. Whence Troelstra obtained this metaphor we do not know.

Proof of the Lemma:

> We define a$'$ such that $\mathrm{a}'(\mathrm{x}) = \mathrm{y} \leftrightarrow (\mathrm{a}(\mathrm{x})$ is the $(\mathrm{y} + 1)^{\mathrm{th}}$ new number generated by a). Or, $\mathrm{a}'(\mathrm{x}) = \mathrm{y} \leftrightarrow (\exists \mathrm{z} < \mathrm{x})(\mathrm{a}(\mathrm{z} + 1) = \mathrm{a}(\mathrm{x}) \wedge \mathrm{a}(\mathrm{x}) \notin \{\mathrm{a}(0), \ldots, \mathrm{a}(\mathrm{z})\} \wedge (\{\mathrm{a}(0), \ldots, \mathrm{a}(\mathrm{z})\}$ contains y different elements$) \vee (\mathrm{a}(\mathrm{x}) = \mathrm{a}(0) \wedge \mathrm{y} = 0))$.
>
> b$'$ is defined likewise from b.

Let us spell out what the definition of a' means:

$a'(0) = 0$, because whatever $a(0)$ is, it is the first number generated by a, and so in this case $y + 1 = 1$ and $y = 0$;

If $a(1) \neq a(0)$ then $a'(1) = 1$ because in this case $a(1)$ is a new number generated by a and all together two numbers were generated by a, so $y + 1 = 2$ and $y = 1$. If, however, $a(1) = a(0)$ then $y + 1$ remains 1 and $a'(1) = y = 0$.

If $a(2) \neq a(1)$ and $a(2) \neq a(0)$ take $z = 1$ and then $y = 2$ and $a'(2) = 2$. Otherwise, $a(2) = a(0)$ or $a(2) = a(1)$. In the first case there is z as required and $a'(2) = 0$. In the second case necessarily $a(1) \neq a(0)$, or else the first case would apply. Take then $z = 0$ and then $a'(2) = 1$.

In general, if $a(n + 1)$ is not in $\{a(0), \ldots, a(n)\}$, take $z = n$ and $a'(n + 1) = n + 1$. Otherwise, if $a(n + 1) = a(0)$ we set $a'(n + 1) = 0$. Otherwise, there is a minimal z, $z > 0$, such that $a(n + 1) = a(z)$ and $a(z) \notin \{a(0), \ldots, a(z-1)\}$ and we

themselves species.)" Definitional equality is intentional whereas the equality between species is extensional. This distinction is not relevant to our discussion in which x is a natural number.

[39] In a lawless sequence, at any stage we know a finite sequence of the first values and nothing on future values (p 17).

[40] Troelstra uses $\vee$ for $\exists$ and $\wedge$ for $\forall$. Troelstra uses φx or ax for our $\varphi(x)$ and $a(x)$. Troelstra denotes the image by φ of any subset X of the domain of φ by φ[X]; we may use φ(X) instead.

define $a'(n + 1) = a'(z)$. $a'(x)$ then gives the first n for which $a(x)$ was generated.

The proof continues thus:

> Now we define ξ as follows: Consider a(x), and let a′(x) = y. Then we can find a b(z) such that b′(z) = y, because {φ(a(0)), . . ., φ(a(x))} must contain at least y different elements. Now we can put ξ(a(x)) = b(z).

Define $\xi(a(0)) = b(0)$. If $a'(x) = y > 0$, there are y different elements in $\{a(0), \ldots, a(x)\}$; so, by the bi-uniqueness of φ, there are also y different elements in $\{\varphi(a(0)), \ldots, \varphi(a(x))\}$. Let $b(z_0) = \varphi(a(0)), \ldots, b(z_x) = \varphi(a(x))$, and let $z_{\max}$ be the maximum of $z_0, \ldots, z_x$. Then $b'(z_{\max}) \geq y$ and there is $z \leq z_{\max}$ such that $b'(z) = y$ and the definition of ξ can be applied. Note that $b(z)$ is not necessarily in $\{\varphi(a(0)), \ldots, \varphi(a(x))\}$. Note further that if there is $x' < x$ such that $a(x') = a(x)$ then $a'(x') = a'(x)$ and then there is $z' < z$ such that $b'(z') = b'(z)$ which implies that $b(z') = b(z)$ and so $\xi(a(x')) = \xi(a(x))$. The definition of ξ is thus coherent.

To conclude the proof Troelstra states:

> We then use the existence of ψ to demonstrate
> ∀z∃x(ξ(a(x)) = b(z)).

Indeed, for $b(x)$, if $b'(x) = y$, then in $\{b(0), \ldots, b(x)\}$ there are at least y different elements. Therefore there are also y different elements in $\{\psi(b(0)), \ldots, \psi(b(x))\}$. There are therefore $z_0, \ldots, z_x$ such that $\psi(b(0)) = a(z_0), \ldots, \psi(b(x)) = a(z_x)$. Let $z_{\max}$ be the maximum of $z_0, \ldots, z_x$. Then $a'(z_{max}) \geq y$ and there is $z \leq z_{\max}$ such that $a'(z) = y$. Note that $a(z)$ is not necessarily in $\{\psi(b(0)), \ldots, \psi(b(x))\}$. Anyway, in $\{a(0), \ldots, a(z)\}$ there are y different elements and so there are $y + 1$ different elements in $\{\varphi(a(0)), \ldots, \varphi(a(z))\}$. Then there is z' such that $b(z') = y$ and by the definition of ξ, $\xi(a(z)) = b(z')$. Since $b(z') = y = b(x)$, $b(z') = b(x)$ and so ξ is on the range of b.

Note that unlike most proofs of CBT, in Troelstra's proof ξ is not defined by way of φ and ψ.[41] In fact, Troelstra even points out (p 104) that this is impossible and he gives the following counterexample (where $\pi_m(x)$ (p 23) means that x is the decimal place of the last 7 in the mth sequence of 10 consecutive 7's in the development of π):

$$\mathrm{X} = \mathrm{N};\ \ \mathrm{Y} = \{2\mathrm{n}+1 |\ \mathrm{n} \in \mathrm{N}\} \cup \{\mathrm{m}\ |\neg\exists \mathrm{x}(\pi_{\mathrm{m}}(\mathrm{x})\}, \varphi(\mathrm{n}) = 2\mathrm{n}+1, \psi(\mathrm{n}) = \mathrm{n}.$$

Troelstra's proof is another original in the history of CBT. Its metaphor is perhaps akin to the metaphor of Myhill's proof in that it enumerates J. Kőnig's strings by a winding path along their members. But Troelstra's path is different from that of Myhill: While Myhill progresses using f and g, Troelstra makes use of the well-ordering property of the natural numbers – the existing of a maximum and minimum in finite subsets.

To extend the lemma Troelstra introduces (p 95f) the relation $\vdash_m$P, where m is a natural number and P a proposition. It reads: at stage m the ideal mathematician has

[41] Such ξ is called by Troelstra "mathematical in φ, ψ" (p 104, 97).

a proof for proposition P.[42] It is assumed as an axiom that $\vdash_m$P is a decidable relation, namely, $\vdash_m P \vee \neg\vdash_m P$ holds for every m and P. Further, the following axiom (p 95f) is assumed: $P \leftrightarrow \exists m(\vdash_m P)$.

A sequence defined by way of $\vdash_m$P Troelstra calls an empirical sequence (p 96). As an example of an empirical sequence Troelstra takes X to be an inhabited species of natural numbers. "A species X is said to be inhabited or secured if $\exists x(x \in X)$" (p 15). It appears that Troelstra requires that a member of X be given, when X is inhabited.[43] For the sake of simplicity Troelstra assumes that this member is 0, an assumption expressed by $\vdash_0 0$. An empirical sequence a he now defines on N×N as follows: $\neg\vdash_m x \in X \rightarrow a(m, x) = 0$; $\vdash_m x \in X \rightarrow a(m, x) = x$. For this a we have (using the specification of the above axiom for $x \in X$, namely, $x \in X \leftrightarrow \exists m(\vdash_m x \in X)$), Range$(a) = X$ (Lemma 16.7.2 p 103).[44] Accepting empirical sequences as lawlike (p 97), Troelstra concludes, leaning on the Lemma proved above, the "intuitionistic analog of Cantor-Bernstein" (Theorem 16.7.3 p 104):

> Let X, Y be inhabited species of natural numbers, and let φ, ψ be bi-unique mappings, $\varphi \in (X)Y$, $\psi \in (Y)X$. Then there exists a bi-unique mapping $\xi \in (X)Y$ such that $\xi(X) = Y$.[45]

Troelstra then compares this theorem with the following:

> Theorem [16.7.5 p 104]: Let X, Y be inhabited subspecies of the natural numbers, and let φ, ψ be bi-unique mappings of X into Y, Y into X respectively, such that φ(X) is detachable in Y, ψ(Y) detachable in X. Then there exists a bi-unique mapping ξ from X onto Y.

Troelstra's use here of 'subspecies' instead of 'species' seems to be insignificant; "'subspecies' is defined like 'subsets'" (p 14). 'Detachable subspecies' is perhaps not defined in Troelstra 1969, but it is a common notion in intuitionism (Heyting 1971 p 39): Y is detachable in X if it is decidable for every $x \in X$ if $x \in Y$ or $x \in (X-Y)$, i.e., $(\forall x \in X)(x \in Y \vee x \notin Y)$.

For proof of Theorem 16.7.5 Troelstra references Myhill's theorem (in Dekker and Myhill 1960 Theorem 13(b) or Rogers 1967 §7.4 Theorem VI; see Sect. 38.2) indicating that its 'recursive' ought to be replaced by 'lawlike'. While it is acceptable that lawlike sequence and recursive function can be identified, because of the use of the notion of algorithm in the definition of both, Troelstra's contention is

---

[42] The context of this relation is Brouwer's theory of the creative subject. This theory is not of interest to us here. We are only interested in the algorithmic aspect of the relation $\vdash_m$P. Troelstra distinguishes between 'having a proof' and 'having evidence for a proposition'. We ignore this subtlety, not relevant to our interest in proofs of CBT.

[43] 'Inhabited' should be distinguished from the notion "non-empty species" which demands only that it is impossible that the species be empty (Dummett 1991 p 27f).

[44] Incidentally, equality here is extensional since we are dealing with species.

[45] Troelstra points out that the requirement of the inhabitance of X, Y can be removed but he does not detail how. Apparently, we could define $X' = X\cup\{0\}$ and $Y' = Y\cup\{0\}$ and get the theorem for X′ and Y′. If $\xi(0) = 0$ we get that ξ is also a bi-unique mapping between X and Y. Otherwise, we define ξ' to be equal to ξ except that $\xi'(\xi^{-1}(0)) = \xi(0)$ and $\xi'(0) = 0$.

unclear for in 16.7.5, φ and ψ are not required to be lawlike. Neither does it make sense that he is talking about φa and ψb, where a and b are the lawlike sequences with ranges X and Y, for φa and ψb are not 1–1. Troelstra could have meant that φ and ψ should be assumed lawlike but then he must also assume that they are defined over N.

However, the main problem with Troelstra's contention is that Myhill's conditions ($x \in$ X iff $\varphi(x) \in$ Y, $y \in$ Y iff $\psi(y) \in$ X) are not captured by the assumption that the ranges of φ, ψ are detachable subspecies of their targets X, Y. For instance, with the conditions of 16.7.5, $x \notin$ X does not imply that $\varphi(x) \notin$ Y; it only implies that $\varphi(x) \notin f$(X). For this reason we believe that Troelstra's attempt to connect his theorem and that of Myhill fails. With it fails also his conclusion (16.7.6 p 104) that under the conditions of 16.7.5, ξ is mathematical in φ and ψ, which is correct with respect to Myhill's theorem. From yet another aspect Troelstra's adoption to intuitionist context of Myhill's proof seems to be unwarranted: there are places in Myhill's proof where the principle of excluded third is used, as in the paragraph where it is argued that when $m \notin \xi$, there are only a finite number of a_i's. Troelstra ignores these places and provides no suggestions to bypass them.

Note that Troelstra's theorem is not contradicted by van Dalen's counterexamples. The former is using, in the conditions of his theorem, only 1–1 mappings, with no requirements on structure preservation, and he obtains a 1–1 mapping at its conclusion. The latter, however, assumes mappings that are structure preserving injections with a removable range, and he attempts to obtain a structure preserving bijection.

For connections between Troelstra's results presented in this section and Kripke's scheme, see Schuster and Zappe 2008 p 327 and passim.

Chapter 39
CBT in Category Theory

We are approaching the end of our journey. In this chapter we will discuss the porting of CBT to category theory. In the final section we will point out certain developments of the period 1918–1924, in which appear the gestalt and metaphor of commutative diagrams, a basic tool of category theory.

The porting of a theorem from one context to another, which usually requires adapting the previous proof to the new context, is a central method for the development of mathematics. Porting is different from generalization. In the latter case, the conditions of the theorem are weakened, but the context remains steady; in the former case, there is a new context. As an example of generalization we can mention D. Kőnig's Factoring Theorem that was first proved for finite graphs and later generalized to arbitrary graphs. As an example of porting we can mention the Fixed-Point Theorem of Tarski-Knaster, which was established in the context of the power-set of a set, and later ported to the context of lattices by Tarski.

The main results that we present in this chapter are from the paper of V. Trnková and V. Koubek from 1973. In their paper, Trnková-Koubek (TK) introduced the notion of a Cantor-Bernstein category and proved that every Brandt category is a Cantor-Bernstein category while every non-Brandt category is not a Cantor-Bernstein category.

Interestingly, TK's characterization of a Cantor-Bernstein category is not by way of the objects and morphisms of the category but by properties of its functors' category. Other researchers did consider porting of CBT to specific categories, namely, to specific mathematical structures. We will mention this research program after deploying TK's results.

As we move along the exposition of TK, we discuss its proof-processing connection, pointed out in the paper, to the papers of Banach (1924, see Chap. 29), Tarski (1928, 1928a) and Knaster (1928, see Sect. 31.2).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6_39, © Springer Basel 2013

39.1 Terminology

Let us introduce first the terminology used in the TK paper. Most of it is not explicitly defined there.

A category ***K*** is composed of two classes: the class of its objects, denoted by $\boldsymbol{K}^{o}$, and the class of its morphisms, denoted by $\boldsymbol{K}^{m}$. A morphism has a domain and range, which are objects. A morphism *m* with domain D and range R, is said to be from D to R and is denoted by $m : D \to R$. Between morphisms there exists an operation called composition and denoted by $\circ$, which is associative. For every object o of a category there exists the unit morphism $1_o : o \to o$ which, when composed to a morphism, leaves it unchanged. The category ***S*** of sets is composed of all sets[1] as objects and all functions between sets as morphisms.

If $f : X \to Y,\ g : Y \to X$ are two morphisms such that $f \circ g = 1_X$, *f* is called a left inverse of *g*, or retraction, and *g* a right inverse of *f*, or coretraction. *f* is called a monomorphism if $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$ for all morphisms g_1, g_2. Clearly if *f* has a left inverse it is a monomorphism but the converse is not necessarily true. *f* is called an epimorphism if $g_1 \circ f = g_2 \circ f$ implies $g_1 = g_2$. If *f* has a right inverse it is an epimorphism but the converse is not necessarily true. If *f* is both a monomorphism and an epimorphism it is called a bimorphism.

If a morphism *f* has both a left-inverse and a right-inverse, then the two inverses are equal and this morphism is the inverse of *f*. Inverse morphisms, if they exist, are unique. Clearly, if *g* is the inverse of *f*, *f* is the inverse of *g*. If *f* has an inverse *f* is called an isomorphism. Two objects with an isomorphism between them are called isomorphic or equivalent. In ***S***, a monomorphism is an injection, an epimorphism is a surjection and a bimorphism, which is always an isomorphism, is a bijection.

A functor $\mathcal{A}$ from the source category ***K***, to the target category $\boldsymbol{K}'$, denoted by $\mathcal{A} : \boldsymbol{K} \to \boldsymbol{K}'$, is a mapping that maps every object *o* of the source to an object $\mathcal{A}(o)$ of the target, and every morphism *m* of the source to a morphism $\mathcal{A}(m)$ of the target. Functors are the morphisms of the category of all categories,[2] and indeed functors can be composed. In this category the identity morphisms are functors from a category to itself that assign every object and morphism to itself. Given two categories ***K*** and $\boldsymbol{K}'$, we say that they are equivalent if there exist a functor $\mathcal{A} : \boldsymbol{K} \to \boldsymbol{K}'$, a functor $\mathcal{B} : \boldsymbol{K}' \to \boldsymbol{K}$, such that $\mathcal{B} \circ \mathcal{A}(o)$ is isomorphic to *o* and $\mathcal{A} \circ \mathcal{B}(o')$ is isomorphic to o', when *o* is an object of ***K*** and o' of $\boldsymbol{K}'$.

In the category of functors, the objects are the functors and the morphisms are natural transformations. A natural transformation η from $\mathcal{A}$ to $\mathcal{B}$, $\eta : \mathcal{A} \to \mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ are functors with the same source and target categories, associates to every object *o* of the source, a morphism η_o of the target, $\eta_o : \mathcal{A}(o) \to \mathcal{B}(o)$, called the component of η at *o*, such that for every morphism $f : o \to o'$ of the source, the following commutative diagram holds in the target: $\eta_{o'} \circ \mathcal{A}(f) = \mathcal{B}(f) \circ \eta_o$. We may

---

[1] Some restrictions apply to prevent $\boldsymbol{S}^{o}$ sets from including paradoxical sets.

[2] In this category the above mentioned restrictions apply too.

omit the subscripts of η when they can be derived from the context. The mono, epi, bi and iso variants of morphisms apply to natural transformations to provide mono, epi, bi or iso transformations. Two isomorphic functors, by an isotransformation, are called 'naturally equivalent'.

When $\mathcal{A}$ and $\mathcal{A}'$ are two functors from the same source category to the category of sets, $\mathcal{A}'$ is a subfunctor of $\mathcal{A}, \mathcal{A}' \subseteq \mathcal{A}$, when for every object o in the source, $\mathcal{A}'(\text{o}) \subseteq \mathcal{A}(\text{o})$, and for every morphism m of the source category, $\mathcal{A}'(\text{m})$ is a domain-range restriction of $\mathcal{A}$(m). If η is a natural transformation from $\mathcal{A}$ to $\mathcal{B}$, both functors from the same source category to the category of sets, and $\mathcal{A}' \subseteq \mathcal{A}$, then by $\eta(\mathcal{A}')$ is denoted the subfunctor $\mathcal{B}'$ of $\mathcal{B}$ which assigns to every object o of the source category, the set $\eta(\mathcal{A}'(\text{o}))$ which is a subset of $\mathcal{B}$(o), and to every morphism $m : \text{o} \to \text{o}'$, the restriction of $\mathcal{B}$(m) to $\mathcal{B}'(\text{o})$ and $\mathcal{B}'(\text{o}')$.

When $\mathcal{A}$ and $\mathcal{A}'$ are two functors from the same source category to the category of sets, we cannot assign to every object o of the source, the set $\mathcal{A}(\text{o}) \cap \mathcal{A}'(\text{o})$, and to every morphism m of the source, the morphism $\mathcal{A}(\text{m}) \cap \mathcal{A}'(\text{m})$, to obtain the functor $\mathcal{A} \cap \mathcal{A}'$, because the intersection of the morphisms may have a smaller domain than the intersection of their domains. Still, in one case the definition is rigorous: when all $\mathcal{A}(\text{o}) \cap \mathcal{A}'(\text{o})$ are equal to the empty set. In this case $\mathcal{A}(\text{m}) \cap \mathcal{A}'(\text{m})$ is empty too. This case is denoted by $\mathcal{A} \cap \mathcal{A}' = \boldsymbol{C}_0$, where $\boldsymbol{C}_0$ is the trivial functor that assigns to every object and morphism of the source the empty set. In this case we say that $\mathcal{A}$ and $\mathcal{A}'$ are disjoint and then we can create the union functor, $\mathcal{A} \cup \mathcal{A}'$, assigning to every object of the source the union of $\mathcal{A}(\text{o})$ and $\mathcal{A}'(\text{o})$, and to every morphism m, the union of $\mathcal{A}$(m) and $\mathcal{A}'(\text{m})$.

For any category $\boldsymbol{K}$ and two objects A, B in it, $\text{Hom}_{\boldsymbol{K}}(\text{A, B})$ is the set of all morphisms in $\boldsymbol{K}$ with domain A and range B. $\text{Hom}_{\boldsymbol{K}}(\text{A, })$ is the functor from $\boldsymbol{K}$ to $\boldsymbol{S}$ that assigns to every object B in $\boldsymbol{K}$ the set $\text{Hom}_{\boldsymbol{K}}(\text{A, B})$ and to every morphism $\text{f}: \text{X} \to \text{Y}$ of $\boldsymbol{K}$, the function $\text{Hom}_{\boldsymbol{K}}(\text{A, f}) : \text{Hom}_{\boldsymbol{K}}(\text{A, X}) \to \text{Hom}_{\boldsymbol{K}}(\text{A, Y})$ that assigns to every morphism g in $\text{Hom}_{\boldsymbol{K}}(\text{A, X})$ the morphism $\text{f} \circ \text{g}$ in $\text{Hom}_{\boldsymbol{K}}(\text{A, Y})$.

39.2 The Cantor-Bernstein Category and Relatives

TK call a category $\boldsymbol{K}$ a Cantor-Bernstein category (CBC), if each two functors $\mathcal{A}$, $\mathcal{B}$ from $\boldsymbol{K}$ to the category of sets $\boldsymbol{S}$, that are each a subfunctor of the other, are equivalent. TK give also an alternative definition: if whenever there is a monotransformation from $\mathcal{A}$ to $\mathcal{B}$ and a monotransformation from $\mathcal{B}$ to $\mathcal{A}$, the functors are naturally equivalent. The association of the second formulation with CBT, in its two-set formulation, is obvious. Clearly, the first formulation is a special case of the second when the monotransformations are the identity mapping. In this case, because of extensionality, $\mathcal{A}(\text{o}) = \mathcal{B}(\text{o})$ and $\mathcal{A}(\text{m}) = \mathcal{B}(\text{m})$ and trivially the two functors are naturally equivalent – by the identity. So the point of bringing the first definition of a CBC category is not clear. Anyway, this definition given in the summary to the paper, is never addressed again.

As an example of a CBC, TK suggest the category with only one morphism. Apparently they mean the category denoted by **2**, with two objects, o and o′, and one morphism m : o → o′, except for the two identities 1_o, $1_o{}'$. A functor from this category to the category of sets takes the two objects into two sets and the morphism into a function between these sets. If $\mathcal{A}$ and $\mathcal{B}$ are two functors from **2** to ***S***, a monotransformation $\eta : \mathcal{A} \to \mathcal{B}$ is composed of an injection $\eta_o : \mathcal{A}(o) \to \mathcal{B}(o)$ and an injection $\eta_{o'} : \mathcal{A}(o') \to \mathcal{B}(o')$. If there exists also a monotransformation $\eta' : \mathcal{B} \to \mathcal{A}$, then by CBT there are bijections (isotransformations) $\eta^*{}_o : \mathcal{A}(o) \to \mathcal{B}(o)$ and $\eta^*{}_{o'} : \mathcal{A}(o') \to \mathcal{B}(o')$. The bijections are composed of the injections and therefore the commutative diagram for the injections implies the commutative diagram for the bijections and so $\mathcal{A}$ and $\mathcal{B}$ are naturally equivalent.

TK call a category ***K***, a Banach category, if each two functors $\mathcal{A}$, $\mathcal{B}$ from ***K*** to the category of sets ***S***, for which there are monotransformations[3] $f : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{B} \to \mathcal{A}$, there are functors $\mathcal{A}_f$, $\mathcal{A}_g$, $\mathcal{B}_f$, $\mathcal{B}_g$ such that

(1) $\mathcal{A}_f \cap \mathcal{A}_g = C_0, \mathcal{A}_f \cup \mathcal{A}_g = \mathcal{A}, \mathcal{B}_f \cap \mathcal{B}_g = C_0, \mathcal{B}_f \cup \mathcal{B}_g = \mathcal{B}$ and

(2) $f(\mathcal{A}_f) = \mathcal{B}_f,\ g(\mathcal{B}_g) = \mathcal{A}_g$.[4]

These conditions immediately associate to Banach's Partitioning Theorem (see Sect. 29.1), from which Banach derived CBT. TK say that evidently, every Banach category is a Cantor-Bernstein category. Indeed, every object o of ***K*** is transferred by $\mathcal{A}$ to $\mathcal{A}_f(o) \cup \mathcal{A}_g(o)$ and by $\mathcal{B}$ to $\mathcal{B}_f(o) \cup \mathcal{B}_g(o)$. Because *f*, *g* are monotransformations, the commutative diagram holds between them and each morphism *m* of ***K***. But then, if we take f' to be f restricted to $\mathcal{A}_f$ and g' to be g restricted to $\mathcal{B}_g$, g' is an isotransformation (actually a bijection since our discussion is in ***S***), so that g'^{-1} exists and therefore $f' \cup g'^{-1}$ is an isotransformation (a bijection), so that $\mathcal{A}, \mathcal{B}$ are equivalent.[5]

TK also introduced the Tarski-Knaster category. It is a category ***K*** for which the following is fulfilled: if each two functors $\mathcal{A}$, $\mathcal{B}$ from ***K*** to the category of sets ***S***, for which there are monotransformation[6] $f : \mathcal{A}_0 \to \mathcal{B}$ and $g : \mathcal{B}_0 \to \mathcal{A}, \mathcal{A}_0 \subseteq \mathcal{A}$, $\mathcal{B}_0 \subseteq \mathcal{B}$ there are functors $\mathcal{A}_f$, $\mathcal{A}_g$, $\mathcal{B}_f$, $\mathcal{B}_g$ such that

(1) $\mathcal{A}_f \cap \mathcal{A}_g = C_0, \mathcal{A}_f \cup \mathcal{A}_g = \mathcal{A}, \mathcal{B}_f \cap \mathcal{B}_g = C_0, \mathcal{B}_f \cup \mathcal{B}_g = \mathcal{B}$ and

(2) $f^{-1}(\mathcal{B}_f) = \mathcal{A}_f,\ g^{-1}(\mathcal{A}_g) = \mathcal{B}_g$.

TK's definition of the Tarski-Knaster category contains no trace of the changes suggested by Tarski or Knaster to the context of CBT, namely, the fixed-point gestalt. In fact, there is only notational difference between their definition of Banach category and of Tarski-Knaster category: the functions of the one are the inverses of the function of the other. Since Tarski-Knaster categories play only a marginal role in the paper (results obtained for Banach categories are shown to hold for Tarski-Knaster categories too) we will in general omit referencing them.

---

[3] TK talk about transformations but from the context it appears they mean monotransformations.

[4] TK added a third condition: (3) If ($\mathcal{A}'_f$, $\mathcal{A}'_g$, $\mathcal{B}'_f$, $\mathcal{B}'_g$) also satisfy (1) and (2), then $\mathcal{A}'_f \subseteq \mathcal{A}_f$. TK disregarded the third condition in their discussion and so we have omitted it from the definition of a Banach category. (3) seems to be related to the possibility of partitioning the two sets in different ways, remarked in a footnote to Sect. 29.1.

[5] Recall that this was not how Banach proved CBT from his Partitioning Theorem. Banach proved an analog of CBT for relations R that fulfill certain properties.

[6] Again, TK talk about transformations but the context suggests monotransformation.

TK finally introduced the notion of Brandt category which is a category each of whose morphisms is an isomorphism. The notion seems to have emerged in Kurosh et al. 1960. At this point TK state their main result, that the following properties of a category ***K*** are equivalent: ***K*** is a Cantor-Bernstein category; ***K*** is a Banach category; ***K*** is a Brandt category. The contended equivalence is interesting because a CBC and a Banach category are characterized by the existence of certain natural transformations, which are mappings between sets, while a Brandt category is characterized by properties within the category, of its morphisms. This link between internal and external properties is unexpected and does not echo any of the gestalt yet encountered in our study.

Following the proof of the above theorem, TK conclude their paper with three more definitions of categories analogous to the Cantor-Bernstein category: A category ***K*** is called c-category (respectively: r-category, e-category) if for every two functors $\mathcal{A}$ and $\mathcal{B}$ from ***K*** to ***S***, such that there exists a contraction (retraction, epitransformation) from $\mathcal{A}$ to $\mathcal{B}$ and another one from $\mathcal{B}$ to $\mathcal{A}$, then $\mathcal{A}$ and $\mathcal{B}$ are naturally equivalent. TK claim, without proof, that the c-categories and r-categories, like the Cantor-Bernstein categories, coincide with the Brandt categories, while the e-categories coincide with the thin Brandt categories in which there exists at most one morphism, an isomorphism, between any two objects. We will not discuss these appended results.

39.3 Every Brandt Category is a Banach Category

This assertion is proved in Lemma 1 of the paper as follows:

Let ***K*** be a Brandt category, $\mathcal{A}$, $\mathcal{B}$ be two functors from ***K*** to ***S***, *f*, *g* two monotransformations $\mathrm{f} : \mathcal{A} \to \mathcal{B}$, $\mathrm{g} : \mathcal{B} \to \mathcal{A}$.[7] Define by complete induction the following sequences of sets:

$\mathcal{A}_0(\mathrm{o}) = \mathcal{A}(\mathrm{o}); \mathcal{B}_0(\mathrm{o}) = \mathcal{B}(\mathrm{o})\text{-}\mathrm{f}(\mathcal{A}_0(\mathrm{o}));$

$\mathcal{A}_{\mathrm{i}+1}(\mathrm{o}) = \mathcal{A}_0(\mathrm{o})\text{-}\Sigma_{\mathrm{j}\leq \mathrm{i}}\mathrm{g}(\mathcal{B}_\mathrm{j}(\mathrm{o})); \mathcal{B}_{\mathrm{i}+1}(\mathrm{o}) = \mathcal{B}(\mathrm{o})\text{-}\mathrm{f}(\mathcal{A}_{\mathrm{i}+1}(\mathrm{o})).$

Clearly, the $\mathcal{A}_\mathrm{i}(\mathrm{o})$ is a sequence of nesting sets and $\mathcal{B}_\mathrm{i}(\mathrm{o})$ is a sequence of encompassing sets. Define: $\mathcal{A}_\mathrm{f}(\mathrm{o}) = \cap \mathcal{A}_\mathrm{i}(\mathrm{o}), \mathcal{A}_\mathrm{g}(\mathrm{o}) = \mathcal{A}(\mathrm{o})\text{-}\mathcal{A}_\mathrm{f}(\mathrm{o}), \mathcal{B}_\mathrm{g}(\mathrm{o}) = \cup\mathcal{B}_\mathrm{i}(\mathrm{o})$, $\mathcal{B}_\mathrm{f}(\mathrm{o}) = \mathcal{B}(\mathrm{o})\text{-}\mathcal{B}_\mathrm{g}(\mathrm{o})$. It is easy to see that:

$\mathcal{A}_\mathrm{f}(\mathrm{o}) = \mathcal{A}(\mathrm{o})\text{-}\mathrm{g}(\mathcal{B}_\mathrm{g}(\mathrm{o})), \mathcal{B}_\mathrm{g}(\mathrm{o}) = \mathcal{B}(\mathrm{o})\text{-}\mathrm{f}(\mathcal{A}_\mathrm{f}(\mathrm{o}));$ hence

$\mathcal{A}_\mathrm{g}(\mathrm{o}) = \mathrm{g}(\mathcal{B}_\mathrm{g}(\mathrm{o})), \mathcal{B}_\mathrm{f}(\mathrm{o}) = \mathrm{f}(\mathcal{A}_\mathrm{f}(\mathrm{o}))$

---

[7] TK omit mentioning that ***K*** is a Brandt category and that *f*, *g* are monotransformations. They must have taken these points for granted because of the preliminaries to the lemma. We change slightly TK's notation in what follows and supplement their presentation with necessary details.

To contemplate $\mathcal{A}_f$, $\mathcal{A}_g$, $\mathcal{B}_f$, $\mathcal{B}_g$ as functors we need to define them for the morphisms of ***K*** (they are defined for the objects of ***K*** by their construction). So let m:o $\rightarrow$ o′ be a morphism of ***K***. Define:

$\mathcal{A}_f$(m) is $\mathcal{A}$(m) restricted to $\mathcal{A}_f$(o) and $\mathcal{A}_f(o')$;
$\mathcal{A}_g$(m) is $\mathcal{A}$(m) restricted to $\mathcal{A}_g$(o) and $\mathcal{A}_g(o')$;
$\mathcal{B}_g$(m) is $\mathcal{B}$(m) restricted to $\mathcal{B}_g$(o) and $\mathcal{B}_g(o')$;
$\mathcal{B}_f$(m) is $\mathcal{B}$(m) restricted to $\mathcal{B}_f$(o) and $\mathcal{B}_f(o')$.

To have this definition work we must prove that if m : o $\rightarrow$ o′ is a morphism of ***K***, then $\mathcal{A}$(m), which is a function in ***S***, takes $\mathcal{A}_f(o)(\mathcal{A}_g(o))$ to $\mathcal{A}_f(o')(\mathcal{A}_g(o'))$, and likewise for $\mathcal{B}$, $\mathcal{B}_f$, $\mathcal{B}_g$. Remember that since ***K*** is a Brandt category, $\mathcal{A}$(m) and $\mathcal{B}$(m) are bijections.

To prove these assertions we proceed by complete induction to prove that $[\mathcal{A}(m)](\mathcal{A}_i(o))$[8] $= \mathcal{A}_i(o')$ and $[\mathcal{B}(m)](\mathcal{B}_i(o)) = \mathcal{B}_i(o')$. Because $\mathcal{A}$(m) is a bijection, $\mathcal{A}$(m) transfers $\mathcal{A}$(o) onto $\mathcal{A}(o')$, which provides the induction's first step for the $\mathcal{A}$'s. Likewise $\mathcal{B}$(m) transfers $\mathcal{B}$(o) onto $\mathcal{B}(o')$ and in addition *f* transfers $\mathcal{A}$(o) onto $f(\mathcal{A}(o)) \subseteq \mathcal{B}(o)$ and $\mathcal{A}(o')$ onto $f(\mathcal{A}(o')) \subseteq \mathcal{B}(o')$. As *f* is a natural transformation, $f\circ\mathcal{A}(m) = \mathcal{B}(m)\circ f$ and so $[\mathcal{B}(m)](f(\mathcal{A}(o))) = f(\mathcal{A}(o'))$. Therefore, $[\mathcal{B}(m)](\mathcal{B}_0(o)) = [\mathcal{B}(m)](\mathcal{B}(o)\text{-}f(\mathcal{A}(o))) = \mathcal{B}(o')\text{-}f(\mathcal{A}(o')) = \mathcal{B}_0(o')$. So the first step of the induction is proved for the $\mathcal{B}$'s too.

Assuming the induction hypothesis, that $[\mathcal{A}(m)](\mathcal{A}_i(o)) = \mathcal{A}_i(o')$ and $[\mathcal{B}(m)](\mathcal{B}_i(o)) = \mathcal{B}_i(o')$, we wish to prove that $[\mathcal{A}(m)](\mathcal{A}_{i+1}(o)) = \mathcal{A}_{i+1}(o')$ and that $[\mathcal{B}(m)](\mathcal{B}_{i+1}(o)) = \mathcal{B}_{i+1}(o')$. First for the $\mathcal{A}$'s:

$$[\mathcal{A}(m)](\mathcal{A}_{i+1}(o)) = [\mathcal{A}(m)](\mathcal{A}_0(o)\text{-}\Sigma_{j\leq i}g(\mathcal{B}_j(o))) =$$
$$[\mathcal{A}(m)](\mathcal{A}_0(o))\text{-}\Sigma_{j\leq i}[\mathcal{A}(m)\circ g](\mathcal{B}_j(o)) = \mathcal{A}(o')\text{-}\Sigma_{j\leq i}[g\circ\mathcal{B}(m)](\mathcal{B}_j(o)) =$$
$$\mathcal{A}(o')\text{-}\Sigma_{j\leq i}g([\mathcal{B}(m)](\mathcal{B}_j(o))) = \mathcal{A}(o')\text{-}\Sigma_{j\leq i}g(\mathcal{B}_j(o')) = \mathcal{A}_{i+1}(o').$$

So the proof is complete for the $\mathcal{A}$'s. Now for the $\mathcal{B}$'s:

$$[\mathcal{B}(m)](\mathcal{B}_{i+1}(o)) = [\mathcal{B}(m)](\mathcal{B}_0(o)\text{-}f(\mathcal{A}_{i+1}(o))) =$$
$$[\mathcal{B}(m)](\mathcal{B}_0(o))\text{-}[\mathcal{B}(m)\circ f](\mathcal{A}_{i+1}(o)) = \mathcal{B}(o')\text{-}[f\circ\mathcal{A}(m)](\mathcal{A}_{i+1}(o)) =$$
$$\mathcal{B}(o')\text{-}f([\mathcal{A}(m)](\mathcal{A}_{i+1}(o))) = \mathcal{B}(o')\text{-}f(\mathcal{A}_{i+1}(o')) = \mathcal{B}_{i+1}(o').$$

So now we have

$$[\mathcal{A}(m)](\mathcal{A}_f(o)) = [\mathcal{A}(m)](\cap\mathcal{A}_i(o)) = \cap[\mathcal{A}(m)](\mathcal{A}_i(o)) = \cap\mathcal{A}_i(o') = \mathcal{A}_f(o'),$$

with similar results for the other of the four new functors.

So the theorem is proved.

If we restrict *f* to the sets $\mathcal{A}_f$(o) as domains and to the sets $\mathcal{B}_f$(o) as ranges, and *g* to the sets $\mathcal{B}_g$(o) as domains and the sets $\mathcal{A}_g$(o) as ranges, the restricted injections are bijections and it follows that $\mathcal{A}$ and $\mathcal{B}$ are equivalent.

With regard to the proof of Lemma 1 TK say: "it is only a simple modification of the Banach theorem or of the Tarski-Knaster theorem". This statement is, however,

[8] $[\mathcal{A}(m)](\mathcal{A}_i(o))$ means the morphism $\mathcal{A}$(m) acting on the object $\mathcal{A}_i(o)$.

unjustified because it bundles the two proofs mentioned under one gestalt, whereas, in fact, Banach's proof rested on the gestalt originated in J. Kőnig's 1906 proof of CBT (see Chap. 29) while Knaster's proof defines a mapping in the power-set and takes its fixed-point as one of the partitions of Banach's theorem from which the others are defined (see Sect. 35.7).

TK's proof of Lemma 1 applies the metaphor from Tarski-Knaster's Fixed-Point Theorem for power-sets (see Sect. 31.2.), for the definition of $\mathcal{A}_1(\mathrm{o})$: it is the complement of the image of the complement of the image. However, instead of defining $\mathcal{A}_f(\mathrm{o})$ by the impredicative definition metaphor, as did Whittaker, the $\mathcal{A}_f(\mathrm{o})$ is defined by the metaphor of taking the intersection of the sequence of $\mathcal{A}_i(\mathrm{o})$ which is defined by the complete induction metaphor. At the same time, the sequence of $\mathcal{B}_i(\mathrm{o})$ is also defined by the complete induction metaphor and $\mathcal{B}_g(\mathrm{o})$ is defined by taking the dual metaphor of union on the $\mathcal{B}_i(\mathrm{o})$ sequence.[9] Thus, unlike the proofs of Banach and Whittaker, in which only one of the partitions is defined directly, in TK's proof two partitions are defined directly, a new metaphor in the partitioning gestalt of Banach. These partitions are fixed-points of f, g respectively. The advantage for TK's procedure is that the defined functors are required to preserve certain structuralistic properties and those are easy to verify by complete induction on the sequences. Here the metaphor of the proof is in that the functors take the morphisms of the category into bijections; this is how the internal property of the category projects an external property of the image of the category in $\boldsymbol{S}$. At this stage of the proof, the commutative diagram of the given monomorphisms is applied to establish the commutative diagram of the constructed isomorphism. To us it seems possible to re-formulate the proof in such a way that a certain relation would emerge that fulfills Banach's properties (α), (β) and thus the result would be obtained by Banach's 1924 Theorem 3.

39.4 A Non-Brandt Category is Not a CBC

TK's proof-plan for this result is to construct, for every non-Brandt category $\boldsymbol{K}$, a pair of functors from $\boldsymbol{K}$ to $\boldsymbol{S}$, with monotransformations from each to the other, which are not isomorphic.

As a preliminary step for the construction, the following lemma is proved:

Lemma 2. *In a non-Brandt category* $\boldsymbol{K}$*, there exists a morphism which is not a coretraction.*

Proof: Let $\mu : \mathrm{o} \to \mathrm{o}'$ be a morphism of $\boldsymbol{K}$ which is not an isomorphism.[10] If μ is a coretraction there exists a morphism $\nu : \mathrm{o}' \to \mathrm{o}$ in $\boldsymbol{K}$ such that $\nu \circ \mu = 1_{\mathrm{o}}$. If ν is a coretraction there exists a morphism $\kappa : \mathrm{o} \to \mathrm{o}'$ such that $\kappa \circ \nu = 1_{\mathrm{o}'}$. Then

[9] The gestalt behind the inductive procedure applied to define the sequences is the gestalt of frames, which appears in Borel-like proofs of CBT.

[10] There exists such, otherwise the category is a Brandt category.

$\kappa = \kappa \circ 1_o = \kappa \circ \nu \circ \mu = 1_{o'}\mu = \mu$, so $\mu \circ \nu = 1_{o'}$ and μ is an isomorphism, contrary to the assumption. So either μ or ν is not coretraction.

The metaphor of the proof must have its origin in the context of algebraic structures, which, however, we have not attempted to trace.

TK now give the construction of the two functors $\mathcal{A}$ and $\mathcal{B}$ as follows[11]:

Let $\boldsymbol{K}$ be a non-Brandt category and $\mu : \mathrm{A} \to \mathrm{B}$ a morphism which is not a coretraction.[12] Define the following functors $\mathcal{A}$, $\mathcal{B}$ from $\boldsymbol{K}$ to $\boldsymbol{S}$: For every object C of $\boldsymbol{K}$, $\mathcal{A}(\mathrm{C}) = \{<\varphi, i, n> \mid \varphi \in \mathrm{Hom}_{\boldsymbol{K}}(\mathrm{A}, \mathrm{C}), i \leq n, n \geq 1\}$,
$\mathcal{B}(\mathrm{C}) = \{<\varphi, i, n> \mid \varphi \in \mathrm{Hom}_{\boldsymbol{K}}(\mathrm{A}, \mathrm{C}), i \leq n, n \geq 2\}$, with the additional provision that $<\varphi, i, n> = <\varphi', i', n'>$ iff $n = n'$, $\varphi = \varphi'$ and $i = i'$ or φ is not a coretraction, and for every morphism $f : \mathrm{X} \to \mathrm{Y}$ of $\boldsymbol{K}$ and
$<\varphi, i, n> \in \mathcal{A}(\mathrm{C})$ $(\in \mathcal{B}(\mathrm{C}))$,
$\mathcal{A}(\mathrm{f})(<\varphi, i, n>) = <\mathrm{f} \circ \varphi, i, n> (= \mathcal{B}(\mathrm{f})(<\varphi, i, n>)$ for $\boldsymbol{n} > 1)$.

The following two lemmas complete the proof that $\boldsymbol{K}$ is not CBC:

Lemma 3. *There exist monotransformations* $f : \mathcal{A} \to \mathcal{B}$ *and* $g : \mathcal{B} \to \mathcal{A}$.

Proof g can be the identity and f defined by $f(<\varphi, i, n>) = <\varphi, i, n + 1>$.

Lemma 4. *The functors* $\mathcal{A}$ *and* $\mathcal{B}$ *are not naturally equivalent.*

Proof:[13] We suppose that there exists an isotransformation $\mathrm{h} : \mathcal{A} \to \mathcal{B}$ and derive a contradiction. h should take the object $\mathcal{A}(\mathrm{A})$ and map it to $\mathcal{B}(\mathrm{A})$. Thus h is assumed to map all members of $\mathcal{A}(\mathrm{A})$, which are triplets $<\varphi, i, n>$, where φ is a morphism from A to A and $i \leq n$, onto the triplets of $\mathcal{B}(\mathrm{A})$, which are of the same form with $n > 1$. Moreover, if f is a morphism of $\boldsymbol{K}$, then the following diagram (Fig. 39.1), where f acts as per the definition of $\mathcal{A}(f)$, should commute, namely:
$\mathrm{h}(<\mathrm{f} \circ \varphi, \mathrm{i}, \mathrm{n}>) = <\mathrm{f} \circ \varphi', \mathrm{i}', \mathrm{n}'>$.

Let $h(<1_{\mathrm{A}}, 1, 1>) = <\varphi, i, n>$. i, n here are specific values and not variables, and we have that $i \leq n$. φ belongs to $\mathrm{Hom}_{\boldsymbol{K}}(\mathrm{A}, \mathrm{A})$, namely, $\varphi : \mathrm{A} \to \mathrm{A}$. From the following diagram (Fig. 39.2), which implements the previous one, we have for every $\psi : \mathrm{A} \to \mathrm{A}$ that $\mathrm{h}(<\psi, 1, 1>) = <\psi \circ \varphi, \mathrm{i}, \mathrm{n}>$.

We distinguish two cases:

(1) φ is not a coretraction. As h is assumed to be an isotransformation, there is $<\chi, k, m>$, $k \leq m$, such that $h(<\chi, k, m>) = <1_{\mathrm{A}}, i, n>$. Then the adjacent diagram entails that $<\varphi \circ \chi, k, m> = <1_{\mathrm{A}}, 1, 1>$, which entails that $m = 1 = k$ and $\varphi \circ \chi = 1_{\mathrm{A}}$ (Fig. 39.3).
So now we have that $\mathrm{h}(<\chi, 1, 1>) = <1_{\mathrm{A}}, i, n>$ by definition, while from the previous diagram we obtained that $\mathrm{h}(<\chi, 1, 1>) = <\chi \circ \varphi, i, n>$; so

[11] $\mathcal{A}$, $\mathcal{B}$ are built through two applications of the operation of disjoint union on the functor $\mathrm{Hom}_{\boldsymbol{K}}(\mathrm{A},)$ (http://en.wikipedia.org/wiki/Disjoint_union).

[12] The first clause that $\boldsymbol{K}$ is non-Brandt, is redundant because of the second clause about μ.

[13] We detail and slightly change TK's proof.

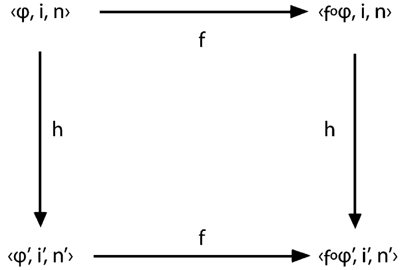


Fig. 39.1 First commutative diagram

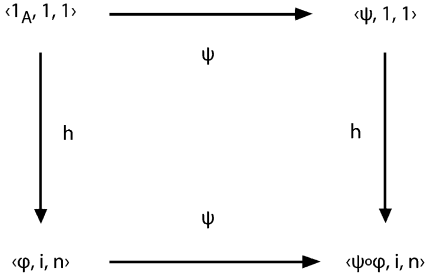


Fig. 39.2 Second commutative diagram

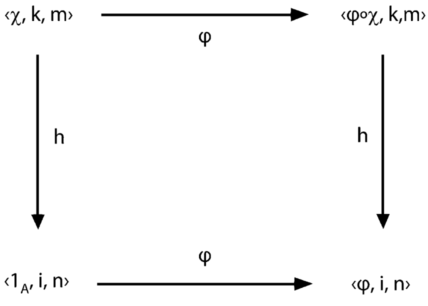


Fig. 39.3 Third commutative diagram

$<1_A$, i, n$> = <\chi\circ\varphi$, i, n$>$ and it follows that $\chi\circ\varphi = 1_A$, contrary to the assumption that φ is not a coretraction. So this case is contradicted.

(2) φ is a coretraction. Then there is ψ such that $\psi\circ\varphi = 1_A$. By a previous remark, $h(<\psi, 1, 1>) = <\psi\circ\varphi$, i, n$>$ so $h(<\psi, 1, 1>) = <1_A, i, n>$. The adjacent diagram (Fig. 39.4) then gives that $h(<\varphi\circ\psi, 1, 1>) = <\varphi$, i, n$>$. But we have by definition $h(<1_A, 1, 1>) = <\varphi, i, n>$, so $<\varphi\circ\psi, 1, 1> = <1_A, 1, 1>$ that implies $\varphi\circ\psi = 1_A$. Thus φ and ψ are inverses of each other.

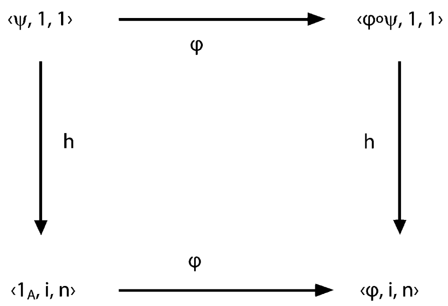


Fig. 39.4 Fourth commutative diagram

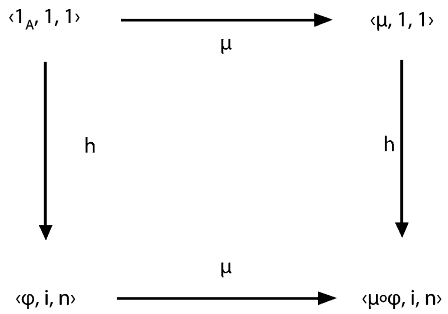


Fig. 39.5 Fifth commutative diagram

Now, because φ is a coretraction, there is $i' \leq$ n such that $<\varphi, i', n> \neq <\varphi, i, n>$. Also, because h is an isomorphism, there is $<\chi, k, m> \in \mathcal{A}$(A) such that $h(<\chi, k, m>) = <\varphi, i', n>$.

Remembering that we have μ : A $\rightarrow$ B which is not a coretraction, we consider the following diagram (Fig. 39.5): $\mu \circ \varphi$ is not a coretraction, for otherwise, there is a ζ : B $\rightarrow$ A such that $\zeta \circ \mu \circ \varphi = 1_A$. Multiplying this equation on the right by ψ we get $\zeta \circ \mu = \psi$. Multiplying this equation on the left by φ we get $\varphi \circ \zeta \circ \mu = 1_A$ and μ appears to be a coretraction contrary to our assumption. As $\mu \circ \varphi$ is not a coretraction, $<\mu \circ \varphi, i', n> = <\mu \circ \varphi, i, n>$.

By the previous diagram, $h(<\mu, 1, 1>) = <\mu \circ \varphi$, i, n$>$, so, in light of the preceding result and the previous diagram, $<\mu, 1, 1> = <\mu \circ \chi$, k, m$>$ and hence $m = 1 = k$. By a previous observation, $h(<\chi, 1, 1>) = <\chi \circ \varphi$, i, n$>$. On the other hand, by the definition of $<\chi, k, m>$ and the assignment $k = m = 1$, we have $h(<\chi, 1, 1>) = <\varphi, i', n>$. So $<\chi \circ \varphi$, i, n$> = <\varphi$, i$'$, n$>$. Since φ is a coretraction this can happen only if $i = i'$, contrary to the definition of i'. So case (2) is contradicted as well (Fig. 39.6).

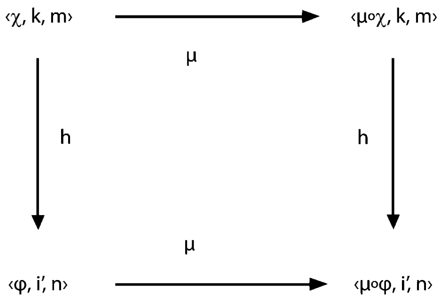


Fig. 39.6 Sixth commutative diagram

Thus h cannot be an isotransformation and the proof that a non-Brandt category is not a Cantor-Bernstein category, is complete.

39.5 Another Approach to CBT in Category Theory

The following question is posted on the web[14]: "Can we characterize Cantor-Bernsteiness in terms of other categorical properties?" Apparently, this question relates CBT to the objects of a category; in fact, just as the original CBT relates to the objects and morphisms of the category of sets. The question then reflects an approach different from the approach in TK's paper, where CBT is related to the category of functors from a given category to the category of sets. The result obtained by TK links a property of the category (Brandt) to a property of its category of functors to ***S***, but it does not answer the web question.

At the referenced site, examples of categories in which CBT holds and others in which it does not hold are given, and the discussion never shifts to the category of functors as in TK's paper. A rudimentary general answer to the question is suggested at the above site. It begins as follows: "Whenever the objects in your category can be classified by a bounded collection of cardinal invariants, then you should expect to have the Schroeder-Bernstein property." In that answer more examples of categories where CBT holds are mentioned. At the site, a link to another site is provided,[15] where "a game of proving or disproving Schroeder-Bernstein in other categories" played in college is mentioned. The posting has several comments added, in which other examples are discussed, along the line of the question, without any mention of TK's results or approach. We do not mean here to devalue the result

[14] http://mathoverflow.net/questions/1058/when-does-cantor-bernstein-hold/1101#1101.

[15] http://sbseminar.wordpress.com/2007/10/30/theme-and-variations-schroeder-bernstein.

of TK, only to point out that the other approach took apparently stronger hold in mathematical research in certain communities.[16] Unfortunately, it is outside the scope of our book to follow this direction any further.[17]

39.6 On Possible Origins of the Commutative Diagram

In his 1928 paper, Tarski suggested giving the relation of equivalence between sets the following definition: A ~ B when there exist two functions of sets F and G such that we have A = F(B) and B = G(A) and for each two sets X and Y: $F(G(X) \cap Y) = X \cap F(Y)$ and $G(F(Y) \cap X) = Y \cap G(X)$. Taking this definition as point of departure", said Tarski, "one can develop a considerable part of the theory of the equality of powers (elementary theory, the Equivalence Theorem [CBT], different theorems of Bernstein [probably BDT] and Zermelo [probably his Denumerable Addition Theorem (1901)]), without appeal to the general theory of sets and applying only the notions and theorems of the algebra of sets.[18]

As reason for this definition Tarski gave his preceding Theorem 5[19]: "For there to exist, for given functions F and G, a relation R such that we have always have F(Y) = R(Y) and G(X) = R^{-1}(X), it is necessary and sufficient that the formulas $F(G(X) \cap Y) = X \cap F(Y)$ and $G(F(Y) \cap X) = Y \cap G(X)$, be satisfied for every two sets X and Y." Theorem 5 followed four other theorems all concerning the images of sets by relations. Tarski indicated that his point of departure is section *37 of Principia Mathematica. There the idea of plural descriptive function is introduced. Tarski notes that the proofs of all the theorems of the 1928 paper requires only the means of the algebra of sets, which he equates with the algebra of logic.

It is curious that Tarski added the note that his results are interesting "... for certain categories of images of sets" before the ideas of category theory emerged. Can Tarski's F, G be interpreted as functors from the regular theory of sets to the algebra of sets? We will not attempt to answer this question here or reconstruct Tarski's theory, which led Tarski-Knaster to their Fixed-Point Theorem (see Chap. 31). We just wish to point out that Tarski's direction can be followed and categories that satisfy the conditions of Tarski's new definition of equivalence be

[16] See Banaschewski-Brümmer 1986 and Sect. 35.7. There are many papers that discuss CBT for various structures. Jan Jakubík should be mentioned who, since 1973, published a number of papers on CBT for various types of lattices and related structures (cf. Jakubík 2002). His work gained a following, especially since the late 1990s, which expanded the discussion. Cf., De Simone et al. 2003, Ionascu 2006, Galego 2010, and the bibliography in those publications.

[17] See remarks 9, 11, 12 in the latter mentioned site on generalizations of the proofs of CBT. Some of the remarks reflects views that we share.

[18] In categories for which CBT fails, is there any meaning in speaking of objects that fulfill the conditions of CBT as having some kind of an equivalence relation? This suggestion may be inline with Tarski's comment cited here.

[19] Tarski gave no proofs of the theorems in his 1928 paper.

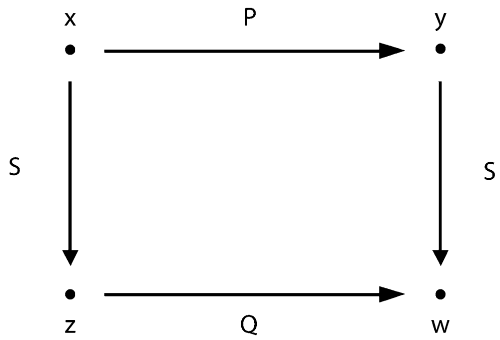


Fig. 39.7 Russell's diagram for the similarity of relations

identified. Perhaps in this direction TK's work can be expanded and other general criteria for categories to be Tarskian categories can be worked out.

Anyway, we brought Tarski's remark here only to show that there were some early conceptions compatible with the later category theory. And we turn to another such example. In 1924, a short (2 pages) note was published titled "On the concept of structure and the theorem of Cantor-Bernstein". The note itself bears only the name Rosenfeld and underneath the name it is written, in smaller font, Liége. In Fraenkel's 1966 bibliography, the note is attributed to Rosenfeld L. It appears thus plausible that the writer of the paper is Leon Rosenfeld, who studied in Liége and in 1930, became Bohr's closest collaborator. So a physicist in the backyard of mathematics!

Rosenfeld describes the purpose of his paper as follows: "The present note contains the results of our analysis of the concept of structure of a relation"

Rosenfeld references Russell 1919 (Chap. 6) as the origin of his thoughts. There (p 54) Russell speaks of the similarity of relations which he explicates by saying that a relation P is similar to a relation Q when there exists a relation S that makes the adjacent diagram commute (Fig. 39.7).

Russell is clearly expressing himself here in category theory terms, almost 20 years before this field had been initiated. Russell continues (p 60) to speak of the structure of the relation as also depicted by a diagram as in the following example (Fig. 39.8).

This presentation of a structure again anticipates the perception of a category as consisting of objects, represented by points, and morphisms, represented by arrows. Russell thus uses diagrams in two different ways: to represent the structure of a relation and to represent the natural transformation between relations, his notion of 'similarity' between relations.

Rosenfeld appears fascinated by the idea of structure, perhaps because of Russell's implied view that structures require interpretation just as scientific laws. He brought a diagram similar to the one given above, and he tried to formalize the notion of structure and to suggest several necessary and sufficient conditions for relations to have the same structure. However, here he ran into some dead-ends; because he was interested only in relations that have the same structure he de facto

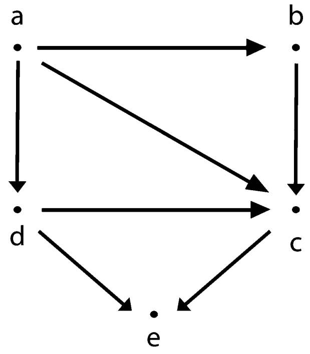


Fig. 39.8 Russell's diagram for the structure of a relation

restricted the theory of structures to Cantor's theory of order-types, which Russell named 'relational numbers'.

At a certain point Rosenfeld's discussion becomes obscure. It seems that he wanted to define the structure of a set α, perhaps as the set of all unit sets of members of α, so that two sets are similar iff they have the same structure. Under his notion of structure Rosenfeld believed CBT would result by extensionality. We do not know if Rosenfeld applied his structuralist vision in his studies of physics.

Conclusion

Our excursion ends here. We have visited many mathematical landscapes, which was the aim of our journey. We have also attempted to link the different sights by developing an eye for similarities and differences in gestalt and metaphors of the different contexts.

While it is not a new topic in the history of mathematics to point out the consequences of certain achievements, or in reverse, to search for the origin of certain developments, depicting such relations in terms of the proof descriptors that we used, seems to us a novelty that emerges with our theory of proof-processing. But in this theory, we believe, we have only marked the beginning. The theory ought to be developed by improving the definition of the proof descriptors that we suggested and by adding new proof descriptors that will enable the identification of deeper links between mathematical contexts. With this development perhaps a new branch, based on a new historiography, will emerge in the history of mathematics: the internal history of mathematics, which will undertake to chart the directed graph of proof-processing relations in the history of mathematical proofs.

The historiography that we suggest stems from the understanding that often proofs, rather than theorems, inspire new proofs and bring about new developments in mathematics. Clearly this view will stand to its Popperian fate like any other theory of phenomena. At this stage, however, the work of the internal historian is to find evidence corroborating the theory. This was our intention in bringing the stories of CBT and BDT.

Another type of corroborating evidence, according to Popper, comes when a theory explains something it was not intended to explain. Our theory explains the following problem, unrelated to the problem situation from which proof-processing emerged: Proofs form the bulk of mathematical texts; it is with proofs that a mathematician wrestles during his education stage and it is in the quest of proofs that he permeates the most fruitful part of his career. Thus it makes sense that existing proofs should influence the emergence of new proofs more than existing theorems, and likewise influence research programs. The proof-processing theory asserts that indeed such influence exists, and if it is not apparent in mathematical texts this is perhaps due to the still prevailing Euclidean style of presentation (Lakatos 1976).

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6, © Springer Basel 2013

Besides its contribution to history, the attempt to describe mathematical proofs by their informal descriptors could enhance the development of popularized expositions of mathematical results. Such development is long overdue, especially when compared to the achievements obtained in the popularizations of physics, biology and genetics. There is no reason why an intellectual is supposed to know of Kant's Copernican revolution, whereby perception conditions reality, and not of Dedekind's Copernican revolution (or, more modestly, view) that infinity conditions finitude. Unexpected links with other products of the human spirit, in art and poetry, could then emerge, enriching the experience of learning, so essential to modern living.

To identify proof descriptors, close-reading of the original texts must be applied. Close-reading, a method in hermeneutics, is common in reading religious scriptures, occult signifiers or poetic texts. Its application to the reading of mathematical texts contributes to the development of what Lakatos called "mathematical criticism" (Lakatos 1976 p 81 note 4, p 139), which he saw as a viable discipline, just as literary or art criticism are, which could contribute to an assessment of the comparative value of proofs.

But if proof-processing is an important tool in describing the historical development of mathematics, it must be present in the logic of discovery of proofs. Indeed we conjecture the theory that mathematicians generate gestalt and metaphoric descriptors, routinely, when they study proofs. It is, we believe, the most basic activity that mathematicians do when appropriating their knowledge (analytic phase). With this conjecture we step out of the research plan of this book into the methodology of the development of mathematics. We suggest the theory that gestalt and metaphoric descriptors obtained through proof-processing form the arsenal of mathematicians for problem solving. When a problem is addressed, it is processed in a similar fashion to the processing of a proof; the descriptors obtained from proof-processing a problem are scanned against the descriptors in the arsenal of the mathematician involved in solving the problem. When an association occurs,[1] it is budding a solution (synthetic phase) to the problem. Perhaps the ability to generate gestalt and metaphoric descriptors to mathematical contexts is the defining characteristic of a mathematician. It is further our conjecture that mathematicians use gestalt and metaphoric descriptors as standard components in the informal language they employ in peer discussions, though they omit them from formal presentations because of traditional reasons and because the language of presentation has not developed enough to enable the presentation of all proof-processing components in writing. Obviously, this aspect of the theory of proof-processing cannot be corroborated through historical study alone. Such study can provide only indirect evidence. An anthropological field study is required, that includes interviews with mathematicians[2] and recordings of informal communication among peers. The development of this research program is, however, outside the scope of this book.

---

[1] How associations occur is probably not part of the logic of discovery.

[2] Some "informant reports" are available in the literature. See Hadamard 1954, Thurston 1994 and Sfard 1994.

Bibliography

Abian A. On comparability of sets. Mathematische Annalen. 1963;149:276–7.

Abian A. The theory of sets and transfinite arithmetic. Philadelphia: Saunders; 1965.

Aczel AD. The mystery of the aleph. New York: Washington Square Press; 2000.

Anellis IH. Russell's earliest reactions to Cantorian set theory, 1896–1900. In: Baumgartner JE, Martin DA, Shelah S, editors. "Axiomatic set theory", Contemporary mathematics. 1984;31:1–11.

Anellis IH. Schröder material at the Russell archives. Modern Logic. 1990/1991;1:237–47.

Anellis IH. The first Russell paradox. In: Drucker T, editor. Perspectives on the history of mathematical logic. Boston: Birkhäuser; 1991: 33–46.

Anellis IH. Review of Mary Tiles, The philosophy of set theory: a historical introduction to Cantor's Paradise. Modern Logic. 1992;2:392–404.

Anellis IH. Memories of Medvedev – a kind and gentle bear. Modern Logic. 1994;4(2):196–206.

Banach S. Un théorème sur les transformations biunivoques. Fund Math. 1924;6:236–9.

Banach S, Tarski A. Sur la décomposition des ensembles de points en parties respectivement congruentes. Fund Math. 1924;6:244–77.

Banaschewski B, Brümmer GCL. Thoughts of the Cantor-Bernstein theorem. Quaestiones Math. 1986;9:1–27.

Banaschewski B, Moore GH. The dual Cantor-Bernstein theorem and the partition principle. Notre Dame J Formal Logic. 1990;31(3):375–81.

Baumgartner JE, Martin DA, Shelah S, editors. Axiomatic set theory, Contemporary mathematics, vol. 31. Providence: American Mathematical Society; 1984.

Bechler Z. Three Copernican revolutions, Zmora-Bitan Tel-Aviv (Hebrew). 1999.

Becker O. Grundlagen der Mathematik in geschichtlicher Entwicklung, München; 1954.

Belna JP. Cantor. Paris: Les Belles lettres; 2003.

Benacerraf P, Putnam H, editors. Philosophy of mathematics: selected readings. 2nd ed. Cambridge: Cambridge University Press; 1983.

Bernstein F. Bemerkung zur Mengenlehre, Göttinger Nachrichten. 1904;557–60.

Bernstein F. Untersuchungen aus der Mengenlehre. Mathematische Annalen. 1905;61:117–55.

Bernstein F. Sur la théorie des ensembles. Comptes Rendus Hebdomadaire des Séances de l'Académie des Science, Paris. 1906;143:953–5.

Bettazzi R. Gruppi finiti ed infiniti di enti, Atti della Reale Academia delle scienze di Torino, Classe di scienze fisiche, matematiche e naturali. 1896;31:506–12. http://www.archive.org/details/attidellarealeac31real.

Birkhoff G. Lattice theory. 3rd ed. New York: American Mathematical Society Colloquium Publications; 1967.

Bolzano B. Paradoxien des Unendlichen. Vdm Verlag, 2006. Partial translation in Ewald 1996 vol 1 p. 249.

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6,

Boniface J. Poincaré et le principe d'induction. Philosophiques. 2004;31(1):131–49.
Borel E. Leçons sur la théorie des fonctions, 1950 edition, Gauthiers-Villars: Paris; 1898.
Borel E. Sur les ensembles effectivement énumèrables et sur les definitions effectives. Rendiconti della Reale Accademia dei Lincei (Roma) (5)28 II(1919), 163–5.
Bourbaki N. Éléments d'histoire des mathématiques. Masson; 1984.
Boutroux P. Correspondance mathématique et relation logique. Revue de metaphysique et de moral. 1905;13:620–37.
Brady G. From Peirce to Skolem: a neglected chapter in the history of logic. Amsterdam: Elsevier; 2000.
Brouwer LEJ. Intuitionism and formalism. Inaugural address 1912. In: Benacerraf P, Putnam H, editors. Philosophy of mathematics: selected readings. 2nd ed. Cambridge: Cambridge University Press; 1983: 77–89.
Bruns G, Schmidt J. Eine Verschärfung des Bernsteinischen Äquivalenzsatzes. Mathematische Annalen. 1958;135:257–62.
Bunn R. Developments in the foundation of mathematics from 1870 to 1910. In: Grattan-Guinness I, editor. From the calculus to set theory, 1630–1910: an introductory history. Princeton: Princeton University Press; 1980b: 220–255.
Burali-Forti C. Una questione sui numeri transfiniti, Rendiconti del Circolo matematico di Palermo. 1897;11:154–64, English translation: a question on transfinite numbers, in van Heijenoort 1967 pp. 104–11.
Candy AL. The elements of plane and solid analytic geometry. Boston: D.C. Heath & co.; 1904. http://www.archive.org/details/elemplanesolidan00candrich.
Cantor G. Notiz zu dem Aufsatz: Beweis, dass eine für jeden reellen Wert von x durch eine in dieser Form gegebene Funktion f(x) sich nur auf eine einzige Weise in dieser Form darstellen lässt ('1871 Notiz'). Cantor 1932;84–6.
Cantor G. Über die Ausdehnung eines Satzes aus der Theorie der trigonometrische Reihen, ('1872 Ausdehnung'). Cantor 1932;92–102.
Cantor G. Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen, ('1874 Eigenschaft'). Cantor 1932;115–8. English translation: Ewald 1996 vol 2, 839–43.
Cantor G. Ein Beitrag zur Mannigfaltigkeitslehre, ('1878 Beitrag'). Cantor 1932;119–33.
Cantor G. Über unendliche, lineare Punktmannigfaltigkeiten, 2, Mathematische Annalen. 1880;17:355–8. Cantor 1932;145–8.
Cantor G. Über unendliche, lineare Punktmannigfaltigkeiten, 3, Mathematische Annalen. 1882;20:113–21. Cantor 1932;149–57.
Cantor G. Grundlagen einer allgemeinen Mannigfaltigkeitslehre, ('Grundlagen'). Cantor 1932;165–209. English translation: Ewald 1996 vol 2, 881–920. Cf. Parpart 1976.
Cantor G. Über unendliche, lineare Punktmannigfaltigkeiten, 6, Mathematische Annalen. 1884;23:453–88. Cantor 1932, 210–46.
Cantor G. Review of Frege [1884], Cantor 1932;440–1. Partial English translation: Ewald 1996 vol 2, 924.
Cantor G. Mitteilungen zur Lehre von Transfiniten, ('1887 Mitteilungen'). Cantor 1932. 378–439.
Cantor G. Über eine elementare Frage der Mannigfaltigkeitslehre, ('1892 Frage'). Jahresbericht der Deutschen Mathematiker-Vereinigung. 1892;1:75–8. Cantor 1932;278–80. English translation Ewald 1996 vol 2, 920–2.
Cantor G. Beiträge zur Begründung der transfiniten Mengenlehre, ('1895 Beiträge'). Cantor 1932;282–311. English translation: Cantor 1915. We refer to this and the following paper together as Cantor 1895/7.
Cantor G. Beiträge zur Begründung der transfiniten Mengenlehre, ('1897 Beiträge'). Cantor 1932;312–56. English translation: Cantor 1915.
Cantor G. Contributions to the founding of the theory of Transfinite Numbers, English version of Cantor 1895 and Cantor 1897, translated by Jourdain PEB. Dover Publications Inc. First published 1915.

Cantor G. Gesammelte Abhandlungen Mathematischen und philosophischen Inhalts, edited by Zermelo E. Springer: Berlin; 1932. http://infini.philosophons.com/.

Cantor G, Dedekind R. Briefwechsel Cantor - Dedekind. In: Noether E, Cavaillès J, editors. Paris: Hermann; 1937. A French translation is included in Cavailles 1962.

Cantor G. Principien einer Theorie der Ordnungstypen, erste Mitteilung, in Grattan-Guinness 1970.

Cavailles J. Philosophie mathématique. Paris: Hermann; 1962.

Charraud N. Infini et inconscient: essai sur Georg Cantor. Paris: Anthropos; 1994.

Coffa JA. The humble origins of Russell's paradox. Russell: the Journal of Bertrand Russell Studies. 1979;99:31–7.

Corry L. Modern algebra and the rise of mathematical structures. Basel/Boston: Birkhäuser Verlag; 1996.

Corry L. The origin of Hilbert's axiomatic method. In: Renn J, editor. The genesis of general relativity, Theories of gravitation in the twilight of classical physics: the promise of mathematics and the dream of a unified theory, vol. 4. Dordrecht: Springer; 2006. p. 139–236.

Couturat L. Les principes des mathématiques, Revue de metaphysique et de moral, 1904;12:19–50, 211–40, 664–98, 810–44, 1905;13:254–6. Collected in Les principes des mathématiques, Paris Alacan 1905, reissued 1965.

Couturat L. Pour la logistique. Revue de metaphysique et de moral. 1906;14:208–50.

Crossley J. A note on Cantor's theorem and Russell's paradox. Australas J Philos. 1973;51:70–1.

Dauben JW. Georg Cantor. His mathematics and the philosophy of the infinite. Cambridge MA: Harvard University Press; 1979. Reprinted by Princeton University Press; 1990.

Dauben JW. The development of Cantorian set theory. In: Grattan-Guinness I, editor. From the calculus to set theory, 1630–1910: an introductory history, Princeton University Press; 1980b: 181–219.

Dauben JW. Georg Cantor and the battle for transfinite set theory. http://www.acmsonline.org/Dauben-Cantor.pdf.

De Simone A, Mundici D, Navara M. A Cantor-Bernstein theorem for σ-complete MV-algebras. Czech Math J. 2003;53:437–47.

Dedekind R. Was sind und was sollen die Zahlen. Braunschweig; 1888. We refer to this publication as *Zahlen*. Cf. Dedekind 1963.

Dedekind R. Gesammelte Mathematische Werke. edited by Fricke R, Noether E, Ore O, editors. vol 3 Braunschweig; 1930–32.

Dedekind R. Essays on the theory of numbers, English translation of Dedekind 1888, by Berman WW, New York: Dover; 1963. Also in Ewald 1996 vol 2 p 787–833.

Dekker JCE, Myhill J. Recursive equivalence types. Berkley: University of California Press; 1960.

Diestel R, Thomassen C. A Cantor-Bernstein theorem for paths in graphs, Am Math Mon. 2006;113 2:161–5. http://www.math.uni-hamburg.de/home/diestel/papers/PathMatchings.pdf.

Dipert RR. The life and work of Ernst Schröder. Modern Logic 1990/1991;1:119–39.

Doyle PG, Conway JH. Division by three. 1994. http://www.math.dartmouth.edu/~doyle/docs/three/three.pdf.

Drake F. Set theory. Amsterdam: North Holland; 1974.

Drucker T, editor. Perspectives on the history of mathematical logic. Boston: Birkhäuser; 1991.

Dugac P. Richard Dedekind et les fondements des mathématiques. Paris: Vrin; 1976.

Dugac P. Richard Dedekind et l'application comme fondement des mathématiques, in Scharlau 1981;134–44.

Dummett MAE. The logical basis of metaphysics. Cambridge: Harvard University Press; 1991.

Ebbinghaus HD. Zermelo: definiteness and the universe of definable sets. History and Philosophy of Logic. 2003;24:197–219.

Ebbinghaus HD. Zermelo in the mirror of the Baer correspondence, 1930–1931. Hist Math. 2004;31(1):76–86.

Ebbinghaus HD. Ernst Zermelo. An approach to his life and work. New York: Springer; 2007.

Erdos P. Personal reminiscences and remarks on the mathematical work of Tibor Galai. Combinatorica. 1982;2(3):207–12.

Ewald W. editor. From Kant to Hilbert: a source book in the foundations of mathematics. 2 vols. Oxford: Clarendon Press; 1996.

Felscher W. Did Cantor prove the Schröder-Bernstein theorem? http://sunsite.utk.edu/math_archives/.http/hypermail/historia/mar99/0148.html.

Ferreirós J. On the relations between Georg Cantor and Richard Dedekind. Hist Math. 1993;20:343–63.

Ferreirós J. 'What fermented in me for years': Cantor's discovery of transfinite numbers. Hist Math. 1995;22:33–42.

Ferreirós J. Traditional logic and the early history of sets 1854–1908. Archive for history of exact sciences. 1996;50:1–67.

Ferreirós J. Labyrinth of thought. A history of set theory and its role in modern mathematics. Basel/Boston/Berlin: Birkhäuser; 1999.

Fraenkel AA. Georg Cantor. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1930; 39:189–266.

Fraenkel AA. Abstract set theory. 1st ed. Amsterdam: North Holland; 1953.

Fraenkel AA. Abstract set theory. 3rd ed. Amsterdam: North Holland; 1966.

Fraenkel AA, Bar-Hillel Y. Le Probleme des antinomies et ses développements recents. Revue de Metaphysique et de Morale. 1939;46(2):225–42.

Fraenkel AA, Bar-Hillel Y. Foundations of set theory. 1st ed. Amsterdam: North Holland; 1958.

Fraenkel AA, Bar-Hillel Y, Levy A. Foundations of set theory. 2nd ed. Amsterdam: North Holland; 1973.

Franchella M. On the origins of Dénes Kőnig's infinity lemma. Archive for History of Exact Sciences. 1997;51(1):3–27.

Franchella M. Towards a re-evaluation of Julius Kőnig's contribution to logic. The bulletin of symbolic logic. 2000;6(1):45–56.

Franklin RW. editor. The poems of Emily Dickinson, 3 vols. Cambridge: Belknap Press; 1999.

Frápolli MJ. Is Cantorian set theory an iterative conception of set? Modern Logic 1990/1991;1: 302–18.

Frege G. Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle/Saale: Verlag L. Nebert; 1879. English translation: van Heijenoort 1967 p 1–82.

Frege G. Die Grundlagen der Arithmetik. Eine Logisch-mathematische Untersuchung über den Begriff der Zahl, Breslau: Verlag von Wilhelm Koebner, 1884. English translation: Austin JL, Oxford: Basil Blackwell; 1950.

Frege G. Grundgesetze der Arithmetik. Begriffsschriftlich abgeleitet. 2 Bände. Jena; 1893. English translation: Montgomery Furth, University of California Press, Berkeley, 1967.

Frege G. Kritische Beleuchtung einiger Punkte in E. Schröder's Vorlesungen über die Algebra der Logik, Archiv fur systematische Philosophie. 1895;1:433–56. English translation: Geach P, Black M. Translation from the philosophical writings of Gottlob Frege, Oxford: Basil Blackwell; 1980. pp. 86–106.

Frege G. On the foundations of geometry and formal theories of arithmetic. English translation of papers from 1903 and 1906, with an introduction, by Kluge E-HW, New Haven/London: Yale University Press; 1971.

Frewer M. Felix Bernstein. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1981;83:84–95.

Galego EM. Cantor-Bernstein sextuples for Banach spaces. Can Math Bull. 2010;53:278–85.

Geach P, Black M. Translation from the philosophical writings of Gottlob Frege. Oxford: Basil Blackwell; 1980.

Gillies DA. Frege, Dedekind, and Peano on the foundations of arithmetic. Assen, Netherlands: Van Gorcum; 1982.

Gillispie CC editor. Dictionary of scientific biography, New York: Charles Scribner's Sons;1970–80.

Ginsburg S. Uniqueness in the left division of order types. Proc Am Math Soc. 1955;6:120–3.

Gödel K. Russell's mathematical logic. In: Benacerraf P, Putnam H, editors. Philosophy of mathematics: selected readings. 2nd ed. Cambridge: Cambridge University Press; 1983. p. 447–69. Originally published in 1944.
Goldfarb W. Poincaré against the logicists. In: Kitcher P, Aspray W, editors. History and philosophy of modern mathematics, Minnesota studies in the philosophy of mathematics XI. Minneapolis: University of Minnesota Press; 1985. p. 61–81.
Grattan-Guinness I. An unpublished paper of Georg Cantor. Acta Math. 1970;124:65–107.
Grattan-Guinness I. The correspondence between Georg Cantor and Philip Jourdain. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1971a;73(Part 1):111–39.
Grattan-Guinness I. Towards a biography of Georg Cantor. Ann Sci. 1971b;27:345–91.
Grattan-Guinness I. The rediscovery of the Cantor-Dedekind correspondence. Jahresbericht der Deutschen Mathematiker-Vereiningung. 1974;76:104–39.
Grattan-Guinness I. Wiener on the logics of Russell and Schröder. Annalen of Science. 1975;31:103–32.
Grattan-Guinness I. Dear Russell – dear Jourdain. A commentary on Russell's logic, based on his correspondence with Philip Jourdain. New York: Columbia University Press; 1977.
Grattan-Guinness I. How Bertrand Russell discovered his paradox. Hist Math. 1978;5:127–37.
Grattan-Guinness I. Georg Cantor's influence on Bertrand Russell. History and philosophy of logic. 1980a;1:61–93.
Grattan-Guinness I, editor. From the calculus to set theory, 1630–1910: an introductory history. Princeton: Princeton University Press; 1980b.
Grattan-Guinness I. How did Russell write 'The principles of mathematics' (1903)? Russell ns. 1997;16:101–27.
Grattan-Guinness I. The search for mathematical roots, 1870–1940: logics, set theories and the foundations of mathematics from Cantor through Russell and Gödel. Princeton: Princeton University Press; 2000.
Griffin N. The Cambridge companion to Bertrand Russell. Cambridge: Cambridge University Press; 2003.
Hadamard J. An essay on the psychology of invention in the mathematical field. New York: Dover; 1954.
Hallett M. Cantorian set theory and limitation of size. Oxford: Clarendon Press; 1984.
Hanf W. On some fundamental problems concerning isomorphism of Boolean algebras. Math Scand. 1957;5:205–17.
Hanf WP, Myers D. Boolean sentence algebras: isomorphism constructions. J Symbolic Logic. 1983;48(2):329–38.
Hardy GH. A theorem concerning the infinite cardinal numbers. The Quarterly Journal of Pure and Applied Mathematics. 1903;35:87–94.
Harward AE. On the transfinite numbers. Philosophical Magazine (6). 1905;10(58):439–60.
Hausdorff F. Grundzüge der Mengenlehre, Berlin; 1914a, reprinted by Chelsea, New York; 1949.
Hausdorff F. Bemerkung über den Inhalt von Punktmengen. Mathematische Annalen. 1914b;75:428–33.
van Heijenoort J. From Frege to Gödel. Cambridge, MA: Harvard University Press; 1967.
Hellmann M. A short proof of an equivalent form of the Schröder-Bernstein theorem. Am Math Mon. 1961;68:770.
Hessenberg G. Grundbegriffe der Mengenlehre. Abhandlungen der Friesschen Schule. 1906;2(1):479–706. reprinted Göttingen, Vardenhoeck & Ruprecht 1906.
Heyting A. Intuitionism: an introduction. 3rd ed. Amsterdam: North Holland; 1971.
Hilbert D. Über die Grundlagen der Logik und Arithmetic, Verhandlungen des Dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904;174–85. English translation: van Heijenhoort 1967. pp. 129–38.
Hobson EW. On the general theory of transfinite numbers and order types. Proc London Math Soc. 1905;3(2):170–88.

Hobson EW. The theory of functions of a real variable and the theory of Fourier's series. Cambridge University Press; 1907. Digital version: http://quod.lib.umich.edu/cgi/t/text/text-idx?c=umhistmath;idno=ACM2112.

Huntington EV. New sets of independent postulates for the algebra of logic. Trans Am Math Soc. 1933;35:274–304. 274–304, 557–558, 971.

Ionascu EJ. A new construction of wavelet sets. 2006. http://arxiv.org/PS_cache/math/pdf/0608/0608277v1.pdf. (2006)

Jakubík J. Cantor Bernstein theorem for lattices. Mathematica Bohemica. 2002;127(3):463–71.

Jané I. The role of the absolute infinite in Cantor's conception of set. Erkenntnis 1905;42:375–402.

Jourdain PEB. On unique non-repeating integer functions. Messenger of Mathematics. May 1901.

Jourdain PEB. On the transfinite cardinal numbers of well-ordered aggregates. Philosophical Magazine (6). 1904a;7(37):61–75.

Jourdain PEB. On the transfinite cardinal numbers of number-classes in general. Philosophical Magazine (6). 1904b;7(39):294–303.

Jourdain PEB. On a proof that every aggregate can be well-ordered. Mathematische Annalen. 1905a;60:465–70.

Jourdain PEB. The definition of a series similarly ordered to the series of all ordinal numbers. Messenger of Mathematics. 1905b;35(2):56–8.

Jourdain PEB. On the question of the existence of transfinite numbers. Proc London Math Soc. 1907a;4(2):266–83.

Jourdain PEB. On the comparison of aggregates. Quarterly Journal of Pure and Applied Mathematics. 1907b;38:352–67.

Jourdain PEB. The multiplication of alephs. Mathematische Annalen. 1908a;65:506–12.

Jourdain PEB. On infinite sums and products of cardinal numbers. The Quarterly Journal of Pure and Applied Mathematics. 1908b;39:375–84.

Kanamori A. Zermelo and set theory. Bulletin of symbolic logis. 2004;10(4):487–553.

Kennedy HC. Peano: life and works of Giuseppe Peano. Dordrecht: D. Reidel; 1980. http://lib.ololo.cc/gen/get/6843d80c6ab75a2002da7a53804b3314/73/6843d80c6ab75a2002da7a53804b3314.pdf.

Kennedy HC. Twelve articles on Giuseppe Peano. Peremptory publication; 2002. http://home.pacbell.net/hubertk/Twelve_Articles.pdf.

Kinoshita S. A solution of a problem of R. Sikorski. Fund Math. 1953;40:39–41.

Knaster B. Un théorème sur les fonctions d'ensembles. Annales de la Societe Polonaise des Mathématiques. 1928;6:133–4.

Kneale W, Kneale M. The development of logic. Oxford: Clarendon Press; 1962.

Kőnig D. Zur Theorie der Mächtigkeiten. Rendiconti Circulo Mathmatico di Palermo. 1908;26:1–4.

Kőnig D. Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre, Mathematische Annalen 1916; 77:453–65. Partial English translation: Biggs NL, Lloyd EK., Wilson RJ, editors. Graph theory 1736–1936. Oxford: Clarendon press; 1986. pp. 203–5.

Kőnig D. Sur un problème de la théorie générale des ensembles et la théorie des graphes. Revue de metaphysique et de morale. 1923;30(3):443–9.

Kőnig D. Sur la correspondence multivoque des ensembles. Fund Math. 1926;8:114–34.

Kőnig D, Valkó S. Über mehrdeutige Abbildungen von Mengen. Mathematische Annalen. 1926;95:136–8.

Kőnig J. Sur la théorie des ensembles. Comptes Rendus Hebdomadaire des Séances de l'Academie des Science, Paris. 1906;143:110–2.

Kőnig J. Neue Grundlagen der Logik, Arithmetik und Mengenlehre. Leipzig: Veit; 1914.

Korselt A. Paradoxien der Mengenlehre. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1906;15:215–9.

Korselt A. Über einen Beweis des Äquivalenzsatzes. Mathematische Annalen. 1911;70:294–6.

Kowalewski G. Bestand und Wandel. Munich: R. Oldenburg; 1950.

Kreiser L. Die Hörer Freges und sein Briefpartner Alwin Korselt. In: Wittgenstein Studies, Diskette; 1/1995. http://www.phil.uni-passau.de/dlwg/ws03/24-1-95.txt.

Kuratowski C. Une propriété de correspondances biunivoques. Fund Math. 1924;6:240–3.

Kuratowski C. On a topological problem connected with the Cantor-Bernstein theorem. Fund Math. 1950;37:21–218.

Kuratowski C. Topology, vol. I. New York/London/Warszawa: Academic Press/Polish Scientific Publishers; 1966.

Kuratowski C, Mostowski A. Set theory. Amsterdam: North Holland; 1968.

Kurosh AG, Livshits A, Shul'geifer EG. Foundations of the theory of categories. Russ Math Surv. 1960;15(6):1.

Lakatos I, editor. Problems in the philosophy of mathematics. Amsterdam: North Holland; 1967.

Lakatos I. Proofs and refutations. Cambridge: Cambridge University Press; 1976.

Lakatos I. What does a mathematical proof prove? In: Worral J, Curry G, editors. Mathematics, science and epistemology. Cambridge University Press; 1978a. p. 61–9. Also in Tymoczko T. editor. New directions in the philosophy of mathematics: an anthology. Birkhäuser; 1986. p. 153–62.

Lakatos I. The methodology of scientific research programmes. Cambridge: Cambridge University Press; 1978b.

Landau E. Richard Dedekind – Gedächtnisrede, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen Geschäftliche Mitteilungen 1917;50–70.

Levy A. Basic set theory, Dover Publications Inc 2002. Originally published by Springer in 1979.

Lindenbaum A, Tarski A. Communication sur les recherches de la théorie des ensembles. Comptes rendu des séances de la société Polonaise de Mathématique section Varsovie Annales de la Societe Polonaise Mathématique. 1926;19:299–330.

Mańka R, Wojciechowska A. On two Cantorian theorems. Annals of the Polish Mathematical Society, Series II: Mathematical News. 1984;25:191–8.

Mazurkiewicz S, Sierpiński W. Sur un ensemble superposable avec chacune de ses deux parties. C R Acad Sci Paris. 1914;158:685–6.

McCarty DC. The mysteries of Richard Dedekind. In: Hintikka J, editor. From Dedekind to Gödel. Dordrecht: Kluwer; 1995. p. 53–96.

McLarty C. Poincaré: mathematics & logic & intuition. Philosophia Math. 1997;5(2):97–115.

Medvedev FA. Ранняя история теоремы эквивалентности (Early history of the equivalence theorem), Ист.-мат. исслед. (Research in the history of mathematics) 1966;17:229–46.

Mehrtens H. Richard Dedekind – der Mensch und die Zahlen. Abhandlungen der Braunschweigischen Wissenschaftlichen Gesellschaft. 1982;33:19–33.

Meschkowski H, Nilson W. Georg Cantor: Briefe. Berlin: Springer; 1991.

Moore GH. Ernst Zermelo, A E Harward, and the axiomatization of set theory. Hist Math. 1976;3(2):206–9.

Moore GH. The origins of Zermelo's axiomatisation of set theory. J Philosophical Logic. 1978;7:307–29.

Moore GH. Zermelo's axiom of choice: its origin, development and influence. Berlin: Springer; 1982.

Moore GH. Towards a history of Cantor's continuum problem. In: Rowe DE, McCleary J, editors. The history of modern mathematics. Boston: Academic; 1989. p. 79–121.

Moore GH. The origins of Russell's paradox: Russell, Couturat, and the antinomy of infinite number. In: Hintikka J, editor. From Dedekind to Gödel. Dordrecht: Kluwer; 1995. p. 215–39.

Moore GH. Review of Ebbinghaus 2007. Not Am Math Soc. 2009;56(7):823–7.

Moore GH, Garciadiego A. Burali-Forti's Paradox: a reappraisal of its origins. Hist Math. 1981;8:319–50.

Morel CA. On the arithmetic of order types. Trans Am Math Soc. 1959;92(1):48–71.

Muddox RD. The origins of the relations algebras in the development and axiomatization of the calculus of relations. Studia Logica. 1991;50:421–55.

Myhill J. Creative sets. Zeitschrift für mathematische Logik und Grundlagen der Mathematik. 1955;1:97–108.

Newman JR, editor. The world of mathematics, vol. 4. New York: Simon and Schuster; 1956.

Otchan G. Quelques questions de l'equivalence des familles d'ensembles. Bulletin de l'Accademie des sciences de l'URSS, series Mathematiques. 1942;6:171–88.

Parpart U. Foundation of a general theory of manifolds, English translation of Cantor 1883, with introduction, The Campaigner (the theoretical journal of the National Caucus of Labor committees), 1976;9:6–68.

Peano G. Arithmetices principia, nova methodo exposita. Turin; 1899. English translation (which we reference): van Heijenoort 1967; pp. 83–97.

Peano G. Démonstration de l'intégrabilité des équations différentielles ordinaire. Mathematische Annalen. 1890;37:182.

Peano G. Super Theorema de Cantor-Bernstein, Rendiconti del Circulo Mathmatico di Palermo, 21:360–6. This printing is from May 1906. The paper was reprinted in Revista de Mathematica, 1906a;8:136–57. The second printing is dated August 1906. The second printing contains an additional part titled "additione".

Peckhaus V. Alwin Reinhold Korselt. Arbeitsberichte aus dem DFG-Projekt Fallstudien zur Begründung einer Sozialgeschichte der Logik. 1986;6:10–23.

Peckhaus V. 'Ich habe mich wohl gehütet, alle Patronen auf einmal zu verschiessen'; Ernst Zermelo in Göttingen. History and Philosophy of Logic. 1990;11:19–58.

Peckhaus V. Introduction, special issue on Schröder. Modern Logic 1990/1991a;1:113–6.

Peckhaus V. Ernst Schröder und die 'pasigraphischen Systeme' von Peano und Peirce. Modern Logic 1990/1991b;1:174–205.

Peckhaus V. Wozu Algebra der Logik? Ernst Schröders Suche nach einer universalen Theorie der Verknüpfungen. Modern Logic. 1994;4:356–81.

Peckhaus V. Ernst Schröder und der Logizismus. In: Stelzner W, editor. Philosophie und Logik. Berlin: de Gruyter; 1995a. p. 108–19.

Peckhaus V. Discussion: Rüstow's thesis on Russell's paradox. Modern Logic. 1995b;5:167–9.

Peckhaus V. Russell's paradox. Russell: the Journal of Bertrand Russell Studies. 1997;17.

Peckhaus V. Pro and Contra Hilbert: Zermelo's set theory. 2002. Lecture notes: http://www.uni-paderborn.de/en/fakultaeten/kw/institute-einrichtungen/institut-fuer-humanwissenschaften/philosophie/personal/peckhaus/texte-zum-download/.

Peckhaus V. Schröder's logic. In: Gabbay DM, Woods J, editors. The rise of modern logic: from Leibniz to Frege, Handbook of the history of logic, vol. 3. Amsterdam: North Holland; 2004a. p. 557–609.

Peckhaus V. Paradoxes in Göttingen, In: Link G editor. One hundred years of Russell's paradox: mathematics, logic, philosophy. Papers from the 2001 Munich Russell conference, Berlin: de Gruyter; 2004b. pp. 501–15.

Peirce CS. On the algebra of logic; A contribution to the philosophy of notation. Am J Math. 1885;7:180–202.

Pla i Carrera J. Dedekind and the theory of sets. Modern Logic. 1993;3(3):215–305.

Poincaré H. "Les mathématiques et la logique", part I. Revue de Métaphysique et de Morale. 1905;13:815–35.

Poincaré H. "Les mathématiques et la logique", part II. Revue de Métaphysique et de Morale. 1906a;14:17–34.

Poincaré H. "Les mathématiques et la logique", part III. Revue de Métaphysique et de Morale. 1906b;14:294–317.

The above articles can be found in: http://www.ac-nancy-metz.fr/enseign/philo/textesph/LES_MATHEMATIQUES_ET_LA_LOGIQUE.pdf. An English translation of the three articles appears in Ewald 1996.

Poincaré H. A propos de la logistique. Revue de Métaphysique et de Morale. 1906c;14:866–8.

Poincaré H. logiques de l'infini, Scientia (Revista di Scientia). 1908;12:1–11. Also in Poincaré 1913 chapter V.

Poincaré H. Réflexions sur les deux notes précédents. Acta Math. 1909;32:195–200.
Poincaré H. On transfinite numbers, in Ewald 1996, vol 2, 1071–4.
Poincaré H. Dernier Pensées. Paris: Flammarion; 1913. http://www.archive.org/details/dernirespens00poin.
Pólya G. How to solve it. Princeton University Press; 1945. 2004 edition.
Pólya G. Mathematics and plausible reasoning, vol. 2. Princeton: Princeton University Press; 1954.
Popper KR. Logik der Forschung. 1934. English translation: The logic of scientific discovery, 1959, Hutchinson.
Potter M. Set theory and its philosophy: a critical introduction. Oxford: Oxford University Press; 2004.
Purkert W. Cantor's philosophical views on the foundations of mathematics. In: Rowe DE, McCleary J, editors. The history of modern mathematics. Boston: Academic; 1989. p. 49–65.
Purkert W, Ilgauds HJ. Georg Cantor 1845–1918. Basel: Birkhäuser Verlag; 1987.
Pyrih P. An example of self-homeomorphic dendrite not strongly self-homeomorphic. 2000. http://www.karlin.mff.cuni.cz/kma-preprints/2000-pap/2000-23.ps.
Reichbach M. Une simple démonstration du théorème de Cantor-Bernstein. Colloquium Mathematica (Wroclaw). 1955;3:163.
Reichbach M. A topological theorem related to the theorem of Cantor-Bernstein, (Hebrew). Riveon Lemathematica. 1958;12:27–31.
Reichbach J. Memorial card. Wiadomości Matematyczne. 1995;31:193–6.
Reichbach M. Memorial card. Wiadomości Matematyczne. 2000;36:189–91.
Remmel JB. On the effectiveness of the Schröder-Bernstein theorem. Proc Am Math Soc. 1981;83 (2):379–86.
Richard J. Les Principes des Mathématiques et le Problème des Ensembles. Revue Générale des Sciences Pures et Appliquées. 1905;16:541–3. English translation: van Heijenoort 1967, 142–144.
Rogers Jr H. Theory of recursive functions and effective computability. New York: McGraw-Hill; 1967.
Rosenfeld L. Sur les concept de structure et le théorème de Cantor-Bernstein. Comptes Rendus de l'Association Francaise pour l'avancement des sciences. 1924;48:144–5.
Russell B. Recent work on the foundation of mathematics. Originally published in 1901. Reprinted in Russell B. Mysticism and logic, Chapter 5 titled Mathematics and the metaphysician, Barnes & Noble 1981 printing, pp. 59–74. Originally printed in 1917.
Russell B. Sur la logique des relation avec des application a la theorie des series, Revue de mathématique. 1900–01;7:115–48. English translation in Marsh RC. editor. Logic and knowledge. London: Georg Allen and Unwin; 1956. pp. 3–38, and Russell; 1993. pp. 310–49.
Russell B. Théorie générale des séries bien-ordonnées, Revue de mathematique 1902;8:12–43. English translation in Russell; 1993. pp. 384–421.
Russell B. The principles of mathematics, I. 1903. 2nd ed. London: George Allen & Unwin; 1937. We refer to the W. W. Norton & Company 1996 printing. We refer to this book as POM.
Russell B. Sur la relation des mathématiques à la logistique, Revue de métaphysique et de morale, 1905;13:906–17. Cf. Russell 1994, paper 23.
Russell B. On some difficulties in the theory of transfinite numbers and order types. Proc London Math Soc. 1906a;4(2):29–53.
Russell B. Les Paradox de la logique. Revue de Metaphysique et de Morale. 1906b;14:627–50.
Russell B. Introduction to mathematical philosophy. London: George Allen & Unwin; 1919.
Russell B. My mental development. Originally published in 1944. Reprinted in Newman 1956, pp. 381–94.
Russell B. The collected papers of Bertrand Russell: toward the "principles of mathematics", 1900–02, vol 3, edited by Moore GH, London: Routledge; 1993.
Russell B. The collected papers of Bertrand Russell: foundations of logic, 1903–1905, vol 4, edited by Urquart A, London: Routledge; 1994.

Scharlau W. editor. Richard Dedekind 1831–1981. Eine Würdigung zu seinem 150. Geburtstag. Braunschweig/Weisbaden: Vieweg; 1981.

Schoenflies A. Die Entwicklung der Lehre von den Punktmannigfaltigkeiten, I. Jahresbericht der Deutschen Mathematiker-Vereinigung 1900;8.

Schoenflies A. Über die logischen Paradoxien der Mengenlehre. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1906;15:19–25.

Schoenflies A. Entwicklung der Mengenlehre und ihrer Anwendung. Leipzig: B. G. Teubner; 1913.

Schoenflies A. Zur Erinnerung an Georg Cantor. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1922;31:97–106.

Schröder E. Vorlesungen über die Algebra der Logik (exakte Logik), vol 3, part 1: Algebra und Logik der Relative, B. G. Teubner: Leipzig; 1895. Reprinted in 1966, Chelsea.

Schröder E. Über G. Cantorsche Sätze. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1896;5:81–2.

Schröder E. Über Zwei Definitionen der Endlichkeit und G. Cantorsche Sätze, Nova Acta. Abhandlungen der Kaiserlichen Leopold-Carolinschen deutschen Akademie der Naturfoscher. 1898;71:301–62.

Schuster P, Zappe J. Über das Kripke-Schema und abzählbare Teilmengen. Logique & Analyse. 2008;204:317–29.

Sfard A. Reification as the birth of metaphor. Learn Math. 1994;14(1):44–55.

Sierpiński W. L'axiome de M. Zermelo et son rôle dans la théorie des ensembles et l'analyse, Bulletin international de l'Académie des Sciences de Cracovie, (A)1918;97–152. Also in Sierpiński W., "Oeuvres choisies", 3 Vols, Hartman S. et al. editors. Panistwowe Wydawnictwo Naukowe: Warsaw; 1974. pp. 208–55.

Sierpiński W. Les examples effectifs et l'axiom du choix. Fund Math. 1921;2:112–8.

Sierpiński W. Sur l'égalité 2m=2n pour les nombres cardinaux. Fund Math. 1922;3:1–6.

Sierpiński W. Sur les ensembles de la même puissance qui ne sont pas effectivement de la même puissance. Fund Math. 1932;18:189–92.

Sierpiński W. Les correspondances multivoques et l'axiome du choix. Fund Math. 1947a;34:39–44.

Sierpiński W. Un théorème sur les puissances des ensembles. Fund Math. 1947b;34:72–4.

Sierpiński W. Démonstration de l'egalite 2m-m=2m pour les nombres cardinaux transfini. Fund Math. 1947c;34:113–8.

Sierpiński W. Sur l'implication $(2m \leq 2n) \rightarrow (m \leq n)$ pour les nombres cardinaux. Fund Math. 1947d;34:148–54.

Sierpiński W. Sur la division des types ordinaux. Fund Math. 1948;35:1–12.

Sierpiński W. Algèbre des ensembles. Warszawa; 1951.

Sikorski R. On a generalization of theorems of Banach and Cantor-Bernstein. Colloquium Mathematicae. 1948;1:140–4.

Silver CL. Who invented Cantor's back and forth argument? Modern Logic. 1994;4:74–8.

Sinaceur MA. Appartenance et inclusion: un inedit de Richard Dedekind. Revue d'Histoire des Sciences. 1971;24:247–55.

Sinaceur MA. La méthode mathématique de Dedekind. Revue d'histoire des science et de leurs application. 1979;32(2):107–42.

Szpilrajn-Marczewski E. On the isomorphism and the equivalence of classes and sequences of sets. Fund Math. 1939;32:133–48.

Tait WW. Cantor's Grundlagen and the Paradoxes of set theory. 1998. http://home.uchicago.edu/~wwtx/cantor.pdf.

Tait WW. Book review on Potter 2004. History and philosophy of logic. 2005;26(2):164.

Tannery P. Note sur la Théorie des ensembles. Bulletin de la Société Mathématique de France. 1884;12:90–6.

Tarski A. Sur les ensemble finis. Fund Math. 1924;6:45–95.

Tarski A. Sur quelques properiétés caractéristiques des images d'ensembles, Comptes rendu des séances de la société Polonaise des Mathematiques 1928;6:127–8. Cf. "Sur quelques theorèmes généraux sur les images densembles", ibid., 132–3.

Tarski A. Über Äquivalenz der Mengen in Bezug auf eine beliebige Klasse von Abbildungen. International Congress of Mathematicians. 1928;2:243–52. Published in 1930 and thus referenced.

Tarski A. Axiomatic and algebraic aspects of two theorems on sums of cardinals. Fund Math. 1948;35:79–104.

Tarski A. Cardinal algebras. New York: Oxford University Press; 1949a.

Tarski A. Cancellation laws in the arithmetic of cardinals. Fund Math. 1949b;36:77–92.

Tarski A. A lattice theoretical fixpoint theorem and its applications. Pac J Math. 1955;5(2):285–309.

Taylor GR. Zermelo, Reductionism, and the Philosophy of Mathematics. Notre Dame Journal of Formal Logic. 1993;34(4):539–63.

Taylor GR. Zermelo's Cantorian theory of systems of infinitely long propositions. Bulletin of Symbolic logic. 2002;8:478–514.

Thurston WP. On Proof and progress in mathematics. Bulletin of the Am Math Soc. 1994;30 (2):161–77.

Tiles M. The philosophy of set theory: a historical introduction to Cantor's paradise. Oxford: Basil Blackwell; 1989.

Trnková V, Koubek V. The Cantor-Bernstein theorem for functors. Commentations Mathematicae universitatis Carolinae. 1973;14(2):197–204.

Troelstra AS. Principles of intuitionism. Lecture notes in mathematics, Springer; 1969.

Troelstra AS, van Dalen D. Constructivism in mathematics: an Introduction. Amsterdam: North Holland; 1988.

Tymoczko T. New directions in the philosophy of mathematics: an anthology. Boston: Birkhäuser; 1986.

Ulam S. Remark on the generalized Bernstein's theorem. Fund Math. 1929;13:281–3.

Unguru S. On the Need to Rewrite the History of Greek Mathematics. Archive for History of Exact Sciences. 1975;15:67–114.

Urquhart A. The theory of types. In: Griffin N, editor. The Cambridge companion to Bertrand Russell. Cambridge: Cambridge University Press; 2003. p. 286–309.

Van Dalen D. A note on spread-cardinals. Compo Math. 1968;20:21–8.

van Dalen D, Dijkman JG, Kleene SC, Troelstra AS, editors. Logic and foundations of mathematics: dedicated to A Heyting on his 70th birthday. Groningen: Wolters-Noordhoff; 1968.

Van Dalen D, Ebbinghaus HD. Zermelo and the Skolem paradox. Bulletin of Symbolic Logic. 2000;6(2):145–61.

Vitali G. Sul problema della misura dei gruppi di punti di una retta. Bologna: Tipografia Gamberini e Parmeggiani; 1905.

Wagon S. The Banach-Tarski paradox. Cambridge: Cambridge University Press; 1985.

Weber H. Elementare Mengenlehre. Jahresbericht der Deutschen Mathematiker-Vereinigung. 1906;15:173–84.

Weyl H. Das Kontinuum. English translation 1987 The continuum: a critical examination of the foundation of analysis. Dover; 1994.

Whitehead AN. On Cardinal numbers. Am J Math. 1902;24:365–94.

Whitehead AN, Russell B. Principia Mathematica, 3 vol, Cambridge University Press. Originally published in 1910–1913. 2nd ed. 1925–1927. We refer to this book as PM.

Whittaker JM. A note on the correlation of classes. Proc Edinburgh Math Soc. 1927;1(1):47–8.

Wiener N. A simplification of the logic of relations. Proceedings of the Cambridge Philosophical Society, 1911–1914;17:387–90. We refer to the publication in van Heijenoort 1967. pp. 224–7.

Winnicott DW. The capacity to be alone. Int J Psychoanal. 1956;39:416–20.

Yuxin Z. From the logic of mathematical discovery to the methodology of scientific research programmes. Brit J Philos Sci. 1990;41(3):377–99.

Zermelo E. Über die Addition transfiniter Cardinalzahlen, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse aus dem Jahre 1901;34–8.

Zermelo E. Beweis dass jede Menge wohlgeordnet werden kann. Mathematische Annalen 1904;59:514–6. English translation: van Heijenoort 1967. pp. 139–41.

Zermelo E. Neuer Beweis für die Möglichkeit einer Wohlordnung. Mathematische Annalen 1908a;65:107–28. English translation: van Heijenoort 1967. pp. 183–98.

Zermelo E. Untersuchungen über die Grundlagen der Mengenlehre. I. Mathematische Annalen 1908b;65:261–81. English translation: van Heijenoort 1967. pp. 199–215.

Zermelo E. Sur les ensembles finis et le principe de l'induction complète. Acta Math. 1909;32:185–93.

Zermelo E. Über den Begriff der Definitheit in der Axiomatik. Fund Math. 1929;14:339–44.

Zermelo E. Über Grenzzahlen und Mengenlehrebereiche, Fund Math. 1930;16. English translation: Ewald 1996 vol 2, 1208–18.

Zhegalkin II. Transfinite numbers, (Russian). 1907. I have not been able to locate further bibliographical information on this work.

Zlot WL. The principle of choice in pre-axiomatic set theory. Scripta Math. 1960;25:105–23.

Names Index

Subject Index

A. Hinkis, *Proofs of the Cantor-Bernstein Theorem*, Science Networks. Historical Studies 45, DOI 10.1007/978-3-0348-0224-6, © Springer Basel 2013