Contributions of

The Logicians.

Part II.
From Richard Dedekind to Gerhard Gentzen

Stanley N Burris

Draft: March 3, 2001

Author address:

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1
E-mail address: snburris@thoralf.uwaterloo.ca
URL: www.thoralf.uwaterloo.ca

Contents

Preface

Over the years I have written up brief surveys of the work done by various mathematical logicians in the period before 1940. This part of the notes gathers short items that are offered as supplementary reading on the web page for my book Logic for Mathematics and Computer Science. Some of the later items, such as Gödel's work, I barely sketch, mainly because they are so well known and well documented elsewhere.

Stanley Burris
Waterloo, Ontario
March, 2001

CHAPTER 1

Was sind und was sollen die Zahlen? Dedekind

Richard Dedekind (1831–1916)
1872 - Stetigkeit und irrationale Zahlen
1888 - Was sind und was sollen die Zahlen

Let us recall that by 1850 the subject of analysis had been given a solid footing in the real numbers — infinitesimals had given way to small positive real numbers, the ε's and δ's. In 1858 Dedekind was in Zürich, lecturing on the differential calculus for the first time. He was concerned about his introduction of the real numbers, with crucial properties being dependent upon the intuitive understanding of a geometrical line.[1] In particular he was not satisfied with his geometrical explanation of why it was that a monotone increasing variable, which is bounded above, approaches a limit. By November of 1858 Dedekind had resolved the issue by showing how to obtain the real numbers (along with their ordering and arithmetical operations) from the rational numbers by means of cuts in the rationals — for then he could prove the above mentioned least upper bound property from simple facts about the rational numbers. Furthermore, he proved that applying cuts to the reals gave no further extension.

These results were first published in 1872, in *Stetigkeit und irrationale Zahlen.* In the introduction to this paper he points out that the real number system can be developed from the natural numbers:

> I see the whole of arithmetic as a necessary, or at least a natural, consequence of the simplest arithmetical act, of counting, and counting is nothing other that the successive creation of the infinite sequence of positive whole numbers in which each individual is defined in terms of the preceding one.

In a single paragraph he simply states that, from the act of creating successive whole numbers, one is led to the concept of addition, and then to multiplication. Then to have subtraction one is led to the integers. Finally the desire for division leads to the rationals. He seems to think that the passage through these steps is completely straight-forward, and he does not give any further detail.

Given the rationals he comes to the conclusion that what is missing is *continuity*, where continuity for him refers to the fact that you cannot create new numbers by cuts. By applying cuts to the rationals he gets the reals, lifts the operations of addition, etc., from the rationals to the reals, and then shows that by applying cuts to the reals no new numbers are created.

In his penetrating 1888 monograph Dedekind returns to numbers. The nature of numbers was a topic of considerable philosophical interest in the latter half of the 1800's — we have already said much about Frege on this topic. In 1887 Kronecker published Begriff der Zahl, in which he does rather little of technical interest, but he does quote an interesting remark which Gauss made in a letter to Bessel in 1830. Gauss says that numbers are distinct from space and time in that the former are a product of our mind. Dedekind picks up on this theme in the introduction to his monograph when he says

[1]Recall that in geometry some mathematicians had already taken efforts to eliminate the dependence of the proofs on drawings.

> In view of this freeing of the elements from any other content (abstraction) one is justified in calling the numbers a free creation of the human mind.

This seems to contrast with Kronecker's later remark:

> God made the natural numbers. Everything else is the work of man.

Regarding the importance of the natural numbers, Dedekind says that it was well known that every theorem of algebra and higher analysis could be rephrased as a theorem about the natural numbers[2] — and that indeed he had heard the great Dirichlet make this remark repeatedly (*Stetigkeit*, p. 338). Dedekind now proceeds to give a rigorous treatment of the natural numbers, and this will be far more exacting than his cursory remarks of 1872 indicated. Actually Dedekind said he had plans to do this around 1872, but due to increasing administrative work he had managed, over the years, to jot down only a few pages. Finally, in 1888, he did finish the project, and published it under the title *Was sind und was sollen die Zahlen?*

Dedekind starts by saying that *objects* (Dinge) are anything one can think of; and collections of objects are called *classes* (Systeme), which are also objects. He takes as absolutely fundamental to human thought the notion of a *mapping*. He then defines a *chain* (Kette) as a class A together with a mapping $f : A \to A$, and proves that *complete induction* holds for chains, i.e., if A and f are given, and if B is a set of generators for A, then for any class C we have

$$A \subseteq C \qquad \text{iff} \qquad B \subseteq C \qquad \text{and} \qquad f(A \cap C) \subseteq C.$$

To say that B is a set of generators for A means that $B \subseteq A$ and the only subclass of A which has B as a subclass and is closed under f is A.

Next a class A is defined to be *infinite* if there is a one-to-one mapping $f : A \to A$ such that $f(A) \neq A$. Dedekind notes that the observation of this property of infinite sets is not new, but using it as a definition is new. He goes on to give a proof that there is an infinite class by noting that if s is a thought which he has, then by letting s' be a thought about the thought s he comes to the conclusion that there are an infinite number of possible thoughts, and thus an infinite class of objects.

A is said to be *simply infinite* if there is a one-to-one mapping $f : A \to A$ such that $A \setminus f(A)$ has a single element a in it, and a generates A. He shows that every infinite A has a simply infinite B in it. Combining this with his proof that infinite classes exist we have a proof that simply infinite sets exist. Any two simply infinite classes are shown to be isomorphic, so he says by *abstracting* from simply infinite classes one obtains the *natural numbers* N.

Let 1 be the initial natural number (which generates N), and let n' be the successor of a natural number n (i.e., n' is just $f(n)$). The *ordering* $<$ of the natural numbers is defined by $m < n$ iff the class of elements generated by n is a subclass of the class of elements generated by m'; and

tion by recur-

the *linearity* of the ordering is proved. Next he introduces *definition by recursion*, namely given any set A and any function $\theta : A \to A$ and given any $a \in A$ he proves there is a unique function satisfying the conditions

- $f(1) = a$
- $f(n') = \theta(f(n))$.

He proves this by first showing (by induction) that for each natural number m there is a unique f_m from N_m to A, where N_m is the set $\{n \in N : n \leq m\}$, which satisfies

- $f_m(1) = a$

[2] For example, the Riemann hypothesis is equivalent to the following statement about the reals (μ is the Möbius function):

$$\forall\varepsilon > 0 \exists x \forall y \left(y > x \Longrightarrow \left(|\sum_{n=1}^{y} \mu(n)| < y^{1/2+\varepsilon} \right) \right),$$

and this can in turn be reduced to a statement about the natural numbers.

- $f_m(n') = \theta(f_m(n))$ for $n < m$.

Then he defines

- $f(m) = f_m(m)$.

Now he turns to the definition of the basic operations. For each integer m he uses recursion to get a function $g_m : N \to N$ such that ←defining +

- $g_m(1) = m'$
- $g_m(n') = (g_m(n))'$.

Then + is defined by

- $m + n = g_m(n)$.

The operation + is then proved to be completely characterized by the following:

- $x + 1 \approx x'$ characterization
- $x + y' \approx (x + y)'$.

Likewise multiplication and exponentiation are defined and shown to be characterized by

- $x \times 1 \approx x$ characterization
- $x \times y' \approx (x \times y) + x$

- $x^1 \approx x$ characterization
- $x^{y'} \approx (x^y) \times x$.

Using induction the following laws are established:

- $x + y \approx y + x$
- $x + (y + z) \approx (x + y) + z$
- $x \times y \approx y \times x$
- $x \times (y \times z) \approx (x \times y) \times z$
- $x \times (y + z) \approx (x \times y) + (x \times z)$
- $(x \times y)^z \approx (x^z) \times (y^z)$
- $x^{y+z} \approx x^y \times x^z$
- $(x^y)^z \approx x^{y \times z}$.

In what follows we see the development of these fundamental laws: the presentation is very close to the original work of Dedekind.

Definition 0.1. [**Addition**] Let addition be defined by:

i. $n + 1 \approx n'$
ii. $m + n' \approx (m + n)'$.

Lemma 0.2. $m' + n \approx m + n'$

PROOF. By induction on n.

For $n \approx 1$	$m' + 1$	$\approx$	$(m')'$	by	$\boxed{}$
		$\approx$	$(m + 1)'$	by	$\boxed{}$
		$\approx$	$m + 1'$	by	$\boxed{}$
Induction Hypothesis:	$m' + n$	$\approx$	$m + n'$		
Thus	$(m' + n)'$	$\approx$	$(m + n')'$	by	$\boxed{}$
so	$m' + n'$	$\approx$	$m + n''$	by	$\boxed{}$

Lemma 0.3. $m' + n \approx (m + n)'$

PROOF.

$m' + n$	$\approx$	$m + n'$	by	$\boxed{}$
	$\approx$	$(m + n)'$	by	$\boxed{}$

Lemma 0.4. $1 + n \approx n'$

PROOF. By induction on n.

$$\begin{array}{lrcll}
\text{For } n \approx 1 & 1+1 & \approx & 1' & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & 1+n & \approx & n' & \\
\text{Thus} & (1+n)' & \approx & n'' & \text{by } \boxed{} \\
\text{so} & 1+n' & \approx & n'' & \text{by } \boxed{}
\end{array}$$

Lemma 0.5. $1 + n \approx n + 1$

PROOF.

$$\begin{array}{lcll}
1+n & \approx & n' & \text{by } \boxed{} \\
 & \approx & n+1 & \text{by } \boxed{}
\end{array}$$

Lemma 0.6. $m + n \approx n + m$

PROOF. By induction on n.

$$\begin{array}{lrcll}
\text{For } n \approx 1\text{:} & m+1 & \approx & 1+m & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & m+n & \approx & n+m & \\
\text{Thus} & (m+n)' & \approx & (n+m)' & \text{by } \boxed{} \\
\text{so} & m+n' & \approx & n+m' & \text{by } \boxed{} \\
 & & \approx & n'+m & \text{by } \boxed{}
\end{array}$$

Lemma 0.7. $(l + m) + n \approx l + (m + n)$

PROOF. By induction on n.

$$\begin{array}{lrcll}
\text{For } n \approx 1 & (l+m)+1 & \approx & (l+m)' & \text{by } \boxed{} \\
 & & \approx & l+m' & \text{by } \boxed{} \\
 & & \approx & l+(m+1) & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & (l+m)+n & \approx & l+(m+n) & \\
\text{Thus} & ((l+m)+n)' & \approx & (l+(m+n))' & \text{by } \boxed{} \\
\text{so} & (l+m)+n' & \approx & l+(m+n)' & \text{by } \boxed{} \\
 & & \approx & l+(m+n') & \text{by } \boxed{}
\end{array}$$

Definition 0.8. [**Multiplication**] Let multiplication be defined by:

iii. $n \cdot 1 \approx n$
iv. $m \cdot n' \approx (m \cdot n) + n$.

Lemma 0.9. $m' \cdot n \approx (m \cdot n) + n$

PROOF. By induction on n.

$$
\begin{array}{lrcll}
\text{For } n \approx 1 & m' \cdot 1 & \approx & m' & \text{by } \boxed{} \\
 & & \approx & m+1 & \text{by } \boxed{} \\
 & & \approx & (m \cdot 1)+1 & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & m' \cdot n & \approx & (m \cdot n)+n & \\
\text{Thus} & (m' \cdot n)+m' & \approx & (m \cdot n+n)+m' & \text{by } \boxed{} \\
\text{so} & m' \cdot n' & \approx & (m' \cdot n)+m' & \text{by } \boxed{} \\
 & & \approx & ((m \cdot n)+n)+m' & \text{by } \boxed{} \\
 & & \approx & (m \cdot n)+(n+m') & \text{by } \boxed{} \\
 & & \approx & (m \cdot n)+(m'+n) & \text{by } \boxed{} \\
 & & \approx & (m \cdot n)+(m+n') & \text{by } \boxed{} \\
 & & \approx & ((m \cdot n)+m)+n' & \text{by } \boxed{} \\
 & & \approx & ((m \cdot n')+n' & \text{by } \boxed{}
\end{array}
$$

Lemma 0.10. $1 \cdot n \approx n$

PROOF. By induction on n.

$$
\begin{array}{lrcll}
\text{For } n \approx 1 & 1 \cdot 1 & \approx & 1 & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & 1 \cdot n & \approx & n & \\
\text{Thus} & (1 \cdot n)+1 & \approx & n+1 & \text{by } \boxed{} \\
\text{so} & 1 \cdot n' & \approx & n' & \text{by } \boxed{}
\end{array}
$$

Lemma 0.11. $m \cdot n \approx n \cdot m$

PROOF. By induction on n.

$$
\begin{array}{lrcll}
\text{For } n \approx 1\text{:} & m \cdot 1 & \approx & m & \text{by } \boxed{} \\
 & & \approx & 1 \cdot m & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & m \cdot n & \approx & n \cdot m & \\
\text{Thus} & (m \cdot n)+m & \approx & (n \cdot m)+m & \text{by } \boxed{} \\
\text{so} & m \cdot n' & \approx & n' \cdot m & \text{by } \boxed{}
\end{array}
$$

Lemma 0.12. $l \cdot (m+n) \approx (l \cdot m)+(l \cdot n)$

PROOF. By induction on n.

$$
\begin{array}{lrcll}
\text{For } n \approx 1 & l \cdot (m+1) & \approx & l \cdot m' & \text{by } \boxed{} \\
 & & \approx & (l \cdot m)+l & \text{by } \boxed{} \\
 & & \approx & (l \cdot m)+(l \cdot 1) & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & l \cdot (m+n) & \approx & (l \cdot m)+(l \cdot n) & \\
\text{Thus} & l \cdot (m+n') & \approx & l \cdot (m+n)' & \text{by } \boxed{} \\
\text{so} & (l \cdot m) \cdot n' & \approx & l \cdot (m+n)' & \text{by } \boxed{} \\
 & & \approx & (l \cdot (m+n))+l & \text{by } \boxed{} \\
 & & \approx & ((l \cdot m)+(l \cdot n))+l & \text{by } \boxed{} \\
 & & \approx & (l \cdot m)+((l \cdot n)+l) & \text{by } \boxed{} \\
 & & \approx & (l \cdot m)+(l \cdot n') & \text{by } \boxed{}
\end{array}
$$

Lemma 0.13. $(m+n) \cdot l \approx (m \cdot l)+(n \cdot l)$

PROOF.

$$\begin{aligned}(m+n)\cdot l &\approx l\cdot(m+n) && \text{by } \boxed{}\\ &\approx (l\cdot m)+(l\cdot n) && \text{by } \boxed{}\\ &\approx (m\cdot l)+(n\cdot l) && \text{by } \boxed{}\\ &\approx (l\cdot m)+(l\cdot n) && \text{by } \boxed{}\end{aligned}$$

Lemma 0.14. $(l\cdot m)\cdot n \approx l\cdot(m\cdot n)$

PROOF. By induction on n.

$$\begin{array}{llcll}\text{For } n\approx 1 & (l\cdot m)\cdot 1 & \approx & l\cdot m & \text{by } \boxed{}\\ & & \approx & l\cdot(m\cdot 1) & \text{by } \boxed{}\\ \text{Induction Hypothesis:} & (l\cdot m)\cdot n & \approx & l\cdot(m\cdot n) & \\ \text{Thus} & (l\cdot m)\cdot n' & \approx & ((l\cdot m)\cdot n)+(l\cdot m) & \text{by } \boxed{}\\ & & \approx & (l\cdot(m\cdot n))+(l\cdot m) & \text{by } \boxed{}\\ & & \approx & l\cdot((m\cdot n)+m) & \text{by } \boxed{}\\ & & \approx & l\cdot(m\cdot n') & \text{by } \boxed{}\end{array}$$

Definition 0.15. [**Exponentiation**] Let exponentiation be defined by:

v. $a^1 \approx a$
vi. $a^{n'} \approx a^n \cdot a$.

Lemma 0.16. $a^{m+n} \approx a^m \cdot a^n$

PROOF. By induction on n.

$$\begin{array}{llcll}\text{For } n\approx 1 & a^{m+1} & \approx & a^{m'} & \text{by } \boxed{}\\ & & \approx & a^m\cdot a & \text{by } \boxed{}\\ \text{Induction Hypothesis:} & a^{m+n} & \approx & a^m\cdot a^n. & \\ \text{Thus} & a^{m+n'} & \approx & a^{(m+n)'} & \text{by } \boxed{}\\ & & \approx & a^{m+n}\cdot a & \text{by } \boxed{}\\ & & \approx & (a^m\cdot a^n)\cdot a & \text{by } \boxed{}\\ & & \approx & a^m\cdot(a^n\cdot a) & \text{by } \boxed{}\\ & & \approx & a^m\cdot a^{n'} & \text{by } \boxed{}\end{array}$$

Lemma 0.17. $(a^m)^n \approx a^{m\cdot n}$

PROOF. By induction on n.

$$\begin{array}{llcll}\text{For } n\approx 1 & (a^m)^1 & \approx & a^m & \text{by } \boxed{}\\ & & \approx & a^{m\cdot 1} & \text{by } \boxed{}\\ \text{Induction Hypothesis:} & (a^m)^n & \approx & a^{m\cdot n}. & \\ \text{Thus} & (a^m)^{n'} & \approx & (a^m)^n\cdot a^m & \text{by } \boxed{}\\ & & \approx & a^{m\cdot n}a^m & \text{by } \boxed{}\\ & & \approx & a^{(m\cdot n)+m} & \text{by } \boxed{}\\ & & \approx & a^{m\cdot n'} & \text{by } \boxed{}\end{array}$$

Lemma 0.18. $(a\cdot b)^n \approx a^n \cdot b^m$

Proof. By induction on n.

$$
\begin{array}{llcll}
\text{For } n \approx 1 & (a \cdot b)^1 & \approx & a \cdot b & \text{by } \boxed{} \\
 & & \approx & a^1 \cdot b^1 & \text{by } \boxed{} \\
\text{Induction Hypothesis:} & (a \cdot b)^n & \approx & a^n \cdot b^n. & \\
\text{Thus} & (a \cdot b)^{n'} & \approx & (a \cdot b)^n \cdot (a \cdot b) & \text{by } \boxed{} \\
 & & \approx & ((a \cdot b)^n \cdot a) \cdot b & \text{by } \boxed{} \\
 & & \approx & (a \cdot (a \cdot b)^n) \cdot b & \text{by } \boxed{} \\
 & & \approx & (a \cdot (a^n \cdot b^n)) \cdot b & \text{by } \boxed{} \\
 & & \approx & ((a \cdot a^n) \cdot b^n) \cdot b & \text{by } \boxed{} \\
 & & \approx & ((a^n \cdot a) \cdot b^n) \cdot b & \text{by } \boxed{} \\
 & & \approx & (a^{n'} \cdot b^n) \cdot b & \text{by } \boxed{} \\
 & & \approx & a^{n'} \cdot (b^n \cdot b) & \text{by } \boxed{} \\
 & & \approx & a^{n'} \cdot b^{n'} & \text{by } \boxed{}
\end{array}
$$

In 50 pages and 158 propositions Dedekind has developed the natural numbers. Now one can use the usual operations of $+$ and $\times$ on N and the ordering $\leq$ to define their extension first to the integers, then to the rationals, and finally to the reals. Consequently the basic study of the real line has been reduced to the study of natural numbers.

Exercises

0.1. [Dedekind] Fill in the details of Dedekind's proofs of the basic laws of arithmetic given above.

CHAPTER 2

Set Theory: Cantor

As we have seen, the naive use of classes, in particular the connection between concept and extension, led to contradiction. Frege mistakenly thought he had repaired the damage in an appendix to Vol. II. Whitehead & Russell limited the possible collection of formulas one could use by *typing*. Another, more popular solution would be introduced by Zermelo. But first let us say a few words about the achievements of Cantor.

Georg Cantor (1845–1918)	
1872 -	Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen
1874 -	Über eine Eigenschaft des Inbegriffes aller reelen algebraischen Zahlen
1879–1884 -	Über unendliche lineare Punktmannigfaltigkeiten (6 papers)
1890 -	Über eine elementare Frage der Mannigfaltigkeitslehre
1895/1897 -	Beiträge zur Begrundung der transfiniten Mengenlehre

We include Cantor in our historical overview, not because of his direct contribution to logic and the formalization of mathematics, but rather because he initiated the study of infinite sets and numbers which have provided such fascinating material, and difficulties, for logicians. After all, a natural foundation for mathematics would need to talk about sets of real numbers, etc., and any reasonably expressive system should be able to cope with one-to-one correspondences and well-orderings.

Cantor started his career by working in algebraic and analytic number theory. Indeed his PhD thesis, his Habilitation, and five papers between 1867 and 1880 were devoted to this area. At Halle, where he was employed after finishing his studies, Heine persuaded him to look at the subject of trigonometric series, leading to eight papers in analysis.

In two papers 1870/1872 Cantor studied when the sequence

$$a_n cos(nx) + b_n sin(nx)$$

converges to 0. Riemann had proved in 1867 that if this happened on an interval and the coefficients were Fourier coefficients then the coefficients converge to 0 as well. Consequently a Fourier series converging on an interval must have coefficients converging to 0. Cantor first was able to drop the condition that the coefficients be Fourier coefficients — consequently any trigonometric series convergent on an interval had coefficients converging to 0. Then in 1872 he was able to show the same if the trigonometric series converged on $[a, b] \setminus A$, provided $A^{(n)} = \emptyset$, where $A^{(n)}$ is the n^{th} derived set of A. The sequence of derived sets is monotone decreasing, and by taking intersections at appropriate points

$$A' \supseteq A'' \supseteq \cdots A^{(n)} \supseteq \cdots \bigcap_{n=1}^{\infty} A^{(n)} \cdots$$

he was led in 1879 to introduce the ordinal numbers $0, 1, \cdots \omega, \cdots$. The key property of ordinals is that they are *well-ordered* sets. (A well-ordered set can be order embedded in the real line iff it is countable.)

We have already mentioned Cantor's (brief) 1872[1] description of how to use Cauchy sequences of rationals to describe the reals. He says that identifying the geometric line with the reals is an axiom.

Cantor's first results on cardinality appear in an 1874 paper where he introduces the 1–1 correspondences, and uses them to show that the algebraic numbers can be put in 1–1 correspondence with the natural numbers; and in the same paper he proves that such a correspondence between any interval of reals and the natural numbers is not possible. Thus he has a new proof of Liouville's 1844 result on the existence of (infinitely many) transcendental numbers (in every interval).

espondences in-
d

oof of existence
cendental num-

In 1878 Cantor proved the (at that time quite paradoxical result) that $\mathbf{R}^n$ could be put into 1–1 correspondence with the reals. Dedekind's response was I see it, but I don't believe it. Cantor subsequently tried to show that no such correspondence could be a homeomorphism if $n > 1$, but a correct proof would wait till Brouwer (1910).

Next followed Cantor's publications of a series of six papers, *On infinite sets of reals,* written between 1879–1884, in which he refined and extended his previous work on infinite sets. He introduced countable ordinals α to describe the sequence of derived sets $A^{(\alpha)}$, and proved that the sequence would eventually stabilize in a perfect set. From this followed the result that any infinite closed subset of $\mathbf{R}$ is the union of a countable set and a perfect set. Next he proved that any nonempty perfect subset of $\mathbf{R}$ could be put in one-to-one correspondence with the real line, and this led to the theorem that any infinite closed set was either countable or had the cardinality of the real line. Cantor claimed that he would soon prove every infinite subset of $\mathbf{R}$ had the cardinality of the positive integers or the cardinality of $\mathbf{R}$, and thus the cardinality of $\mathbf{R}$ would be the second infinity. His proof of what would later be called the *Continuum Hypothesis* (more briefly, the CH) did not materialize. Later Souslin would be able to extend his ideas to show that analytic sets were either countable or the size of the continuum; attempts to settle the Continuum Hypothesis would lead to some of the deepest work in set theory — by Gödel (1940), who showed the consistency of the CH, and Cohen (1963) who invented forcing to show the independence of the CH.

A particularly famous result appeared in *Über eine elementare Frage der Mannigfaltigkeitslehre* (1890), namely the set of functions 2^A, i.e., the set of functions from A to $\{0,1\}$, has a larger cardinality than A, proved by the now standard diagonal method.

Cantor's last two papers on set theory, *Contributions to the foundations of infinite set theory,* 1895/1897, give his most polished study of cardinal and ordinal numbers and their arithmetic. He says that the cardinality of a set is obtained by using our mental capacity of *abstraction,* by ignoring the nature of the elements. By looking at the sequence of sizes of ordinals he obtains his famous $\aleph$'s ($\aleph_0, \cdots, \aleph_\omega, \cdots$) which, ordered by their size, form a well-ordered set *in the extended sense,* i.e., for any set of $\aleph$'s there is a smallest one, and a next largest one. He claims that the size of any set is one of his $\aleph$'s — as a corollary it immediately follows that the reals can be well-ordered. He tried several times to give a proof of this claim about the $\aleph$'s; but it was not until 1904, when Zermelo invoked the axiom of choice, that there would finally be a success.

For his development of ordinal numbers he first looks at *linearly ordered* sets and defines $+$ and $\times$ for the *order types* abstracted from them. Next he shows the order type of the rationals is completely determined by the properties of being

1. countable
2. order dense, and
3. without endpoints.

Then he characterizes the order type of the interval [0,1] of reals by

1. every sequence has a limit point, and
2. there is a countable dense subset.

[1]This is the same year he met Dedekind, while on vacation

Ordinals are then defined as the order types (abstracted from) well-ordered sets. Exponentiation of ordinals is defined, and the expansion of countable ordinals as sums of powers of ω is examined. The paper ends with a look at the countable ε ordinals, i.e., those α which satisfy $\omega^\alpha = \alpha$ (and hence their expansion is just ω^α).

By the end of the nineteenth century Cantor was aware of the paradoxes one could encounter in his set theory, e.g., the set of everything thinkable leads to contradictions, as well as the set of all cardinals and the set of all ordinals. Cantor solved these difficulties for himself by saying there were two kinds of infinities, the *consistent* ones and the *inconsistent* ones. The inconsistent ones led to contradictions. This approach, of two kinds of sets, would be formalized in von Neumann's set theory of 1925.

Cantor's early work with the infinite was regarded with suspicion, especially by the influential Kronecker. However, with respected mathematicians like Hadamard, Hilbert, Hurwitz, Mittag-Leffler, Minkowski, and Weierstrass supporting his ideas, in particular at the First International Congress of Mathematicians in Zürich (1897), we find that by the end of the century Cantor's set theory was widely known and publicized, e.g., Borel's *Lecons sur la théorie des fonctions* was mainly a text on this subject. When Hilbert gave his famous list of problems in 1900, the Continuum Problem was the first.

A considerable stir was created at the Third International Congress of Mathematicians in Heidelberg (1904) when König presented a proof that the size of $\mathbf{R}$ was not one of the $\aleph$'s of Cantor. Cantor was convinced that the cardinal of every set would be one of his $\aleph$'s. König's proof was soon refuted.

The first textbook explicitly devoted to the subject of Cantor's set theory was published in 1906 in England by the Youngs, a famous husband and wife team. The first German text would be by Hausdorff in 1914.

CHAPTER 3

Concept Script: Frege

Gottlieb Frege (1848–1925)
1879 - Begriffsschrift
1884 - Grundlagen der Arithmetik
1893 - Grundgesetze der Arithmetik I.
1903 - Grundgesetze der Arithmetik II.

Frege initiated an ambitious program to use a precise notation which would help in the rigorous development of mathematics. Although his efforts were almost entirely focused on the natural numbers, he discussed possible applications to geometry, analysis, mechanics, physics of motion, and philosophy.

The precise notation of Frege was introduced in *Begriffsschrift* (Concept Script) in 1879. This was a two-dimensional notation whose powers he compared to a microscope. The framework in which he set up his *Begriffsschrift* was quite simple — we live in a world of objects and concepts, and we deal with statements about these in a manner subject to the laws of logic. Thus Frege had only one model in mind, the real world. Let us refer to this as the *absolute universe.* From this he was going to distill the numbers and their properties. The absolute universe approach to mathematics via logic was dominant until 1930 — we see it in the work of Whitehead and Russell (1910–1913).

His formal system with two-dimensional notation had the universal quantifier, negation, implication, predicates of several variables, axioms for logic, and rules of inference. The explicit universal quantifier, predicates of several variables and the rules of inference were new to formal systems!

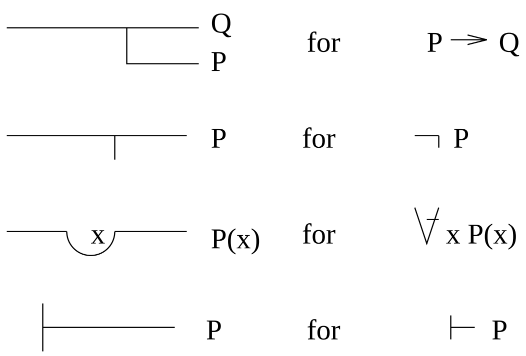

Figure 7: Begriffschrift notation

The rule of inference and axioms found at the beginning of Begriffsschrift are

RULE of INFERENCE

s ponens

$$\frac{P \to Q,\ P}{Q}$$

AXIOMS	
$P \to (Q \to P)$	$\neg\neg P \to P$
$(P \to (Q \to R)) \to ((P \to Q) \to (P \to R))$	$P \to \neg\neg P$
$(P \to (Q \to R)) \to (Q \to (P \to R))$	$x \approx y \to (P(x) \to P(y))$
$(P \to Q) \to (\neg Q \to \neg P))$	$x \approx x$
$\forall x P(x) \to P(x)$	

Higher order quantification was also permitted in the Begriffsschrift, but the axioms and rules for working with such were not presented.

The two-dimensional notation, the lack of new mathematical results, and the tedious work required ensured that *Begriffsschrift* would be almost totally ignored. Nonetheless there were major contributions in this paper, namely

i. An elegant propositional logic based on $\to$ and $\neg$,
ii. a notation for universally quantified variables,
iii. relation symbols of several variables,[1]
    and these were incorporated in
iv. a powerfully expressive higher order logic, the likes of which had never been seen before.

Frege's attempt to set up the natural numbers in logic is based on what he calls his definition of a *sequence* — this is his main application of his logic to mathematics. Although his ultimate goal is to abstractly describe the sequence of natural numbers N with the usual ordering $\leq$, he only succeeds in describing a broader class of "sequences".

The crucial point in his work is that from a notion of *successor* he wants to be able to capture the notion of "y follows x" without using the obvious "for some integer n, y is the n^{th} element after x" — for otherwise his development of the natural numbers will be circular.

In modern notation he proceeds as follows. A property P is *hereditary* for a binary relation r, written $\text{Hered}(P, r)$, if

$$\forall x \forall y\, [(P(x) \wedge r(x,y)) \to P(y)]$$

holds. The relation r is *one-one*, written $E(r)$, if

$$\forall x \forall y \forall z\, [r(x,y) \wedge r(x,z)) \to y \approx z].$$

Given any binary r he defines a binary relation $\leq$ by

$$x \leq y \qquad \text{iff} \qquad x \approx y \vee \forall P\, [\text{Hered}(P,r) \to \forall z\, ((r(x,z) \wedge P(z)) \to P(y))].$$

Then the final results of Begriffsschrift are the transitivity and comparability properties of $\leq$:

$$\forall x \forall y \forall z\, [(x \leq y \wedge y \leq z) \to x \leq z] \tag{1}$$

$$E(r) \to \forall x \forall y (x \leq y \vee y \leq x). \tag{2}$$

Frege does not specify a particular relation r which will lead to a sequence like the natural numbers — this will first appear in Dedekind, 1888, where r is selected to be one-one, not onto, function from a domain to itself. Dedekind will use only one property P, namely select any a from the domain, but not from the range, of r, and let $P(x)$ be "x is in the subuniverse generated by a". One can of course define the "subuniverse generated by a" without reference to n-fold applications

[1] Frege used the word 'function' where we now use the word 'relation'. This was again adopted by Hilbert and Ackermann in the second edition of their book (1938). Unfortunately we also use the words 'function symbol' in modern logic, but with quite a different meaning, namely such will be interpreted as a function on a domain to itself.

of r to a; and it is a hereditary property. Dedekind, in a later introduction to his work, will state that in his development of the natural numbers he was unaware of Frege's work of 1879.

Frege's later work on the foundations of arithmetic abandons this work on sequences and picks up Cantor's definition of cardinal number based on one-one correspondences.

Obviously one has only begun to investigate number systems at the end of *Begriffsschrift.* Indeed, numbers have not even been defined. The reviews ranged from mediocre to negative. In particular Schröder thought Boole's logical system was far superior — Boole used the arithmetic notation for plus and times, and had marvelled at how much the laws of logic were like the laws of arithmetic. Schröder showed how much easier it was to write out the propositional part of Frege's work in Boole's notation.

In 1884 Frege published his second book on his approach to numbers, *Grundlagen der Arithmetik* (Basic Principles of Arithmetic). However this time he tried for popular appeal by omitting any scientific notation and using prose to explain his ideas. Although he was quite content with the foundations of geometry,[2] his theme that no one had provided a decent foundation for numbers was discussed at length. He presents several explanations of the nature of numbers which could be found in the literature, pointing out the shortcomings of each in turn.

In the last part of the book he proposed to solve the foundational question by showing that numbers can be obtained from pure logic. His main tool was the then well-known notion of one-to-one correspondence. Using this he defined the *cardinal number* (Anzahl) of a property P as the collection of all properties Q such that the class defined by Q can be put in one-to-one correspondence with the class defined by P. Then he defined 0 to be the cardinal number of the property $x \not\approx x$. Next he defined what it means for one number *to immediately follow* another, defined the number 1 to be the cardinal number of the property $x \approx 0$, and stated some elementary theorems.

At the end of *Grundlagen,* Frege says that with his notation and definitions it will be possible to carry out a development of the basic laws of arithmetic without gaps. *Grundlagen* received three reviews, all negative (one was by Cantor). Nonetheless Frege turned to the writing of his final masterpiece, the two volumes of *Grundgesetze der Arithmetik* (Basic Laws of Arithmetic). As the work neared completion he had difficulty finding a publisher because of the poor showing of his previous books on this subject. Finally he found a publisher who agreed to publish *Grundgesetze I,* and would publish Vol. II provided Vol. I was successful.

Grundgesetze I appeared in 1893. Again Frege emphasizes the need for a rigorous development of numbers. And as usual he finds fault with what has been offered by others. He says that Dedekind's *Was sind und was sollen die Zahlen?* is full of gaps. We quote: (*Grundgesetze*, p. VII)

> To be sure this [Dedekind's] brevity is attained only because a great deal is really not proved at all $\cdots$ an inventory of the logical or other laws which he takes as basic is nowhere to be found.

Also he says that Schröder's *Algebra der Logik* (1890–1910) is more concerned with techniques and theorems than with foundations. As to Frege's standards and goals we can do no better than quote from the foreword of *Grundgesetze I*, p. VI:

> The idea of a strictly scientific method in mathematics, which I have attempted to realize, and which might indeed be named after Euclid, I should like to describe as follows. It cannot be demanded that everything be proved, because that is impossible; but we can require that all propositions used without proof be expressly declared as such, so that we can see distinctly what the whole structure rests upon. After that we must try to diminish the number of primitive laws as far as possible, by proving everything that can be proved.

[2]In a letter to Hilbert at a later date he said there was no reason to prove the axioms of geometry consistent because they were true.

> Furthermore, I demand — and in this I go beyond Euclid — that all methods of inference employed be specified in advance ···

In *Grundgesetze I* Frege introduces some items which did not appear in *Begriffsschrift*, namely

i. True and False[3]
ii. $\{x : P(x)\}$ for the class determined by the property P (his notation was like x'$P(x)$),
iii. $\backslash x\{y : P(y)\}$ for the unique element satisfying $P(y)$, if such exists; and otherwise this signifies just $\{y : P(y)\}$, and
iv. $x \in y$ for x is a member of the class y (his notation was $x \frown y$).

In *Grundgesetze I* he only needs the single rule of inference modus ponens. For the axioms he takes, in addition to those for the propositional calculus, the following:

i. $\forall x\, P(x) \rightarrow P(x)$,
ii. $\forall P\, P(x) \rightarrow P(x)$,
iii. $x \approx y \rightarrow \forall P\, (P(x) \rightarrow P(y))$
iv. $\{x : P(x)\} \approx \{x : Q(x)\} \leftrightarrow \forall x\, (P(x) \leftrightarrow Q(x))$, and
v. $x \approx \backslash y\{y : x \approx y\}$.

Grundgesetze I had only two reviews, again neither very favorable. Peano was one of the reviewers, and he used the review to point out the advantages of his own approach to number systems; in particular he thought the fact that he used fewer primitives (three) than Frege was a big plus for his system. In 1896 Peano received a detailed response from Frege, pointing out that in fact there were at least nine primitive notions in Peano's system, not to mention the incompleteness and confusion at certain points. Peano was able to incorporate some of Frege's suggestions into Vol. II of Formulario (1899), and, as requested, he also published Frege's letter with a recanting of statements in his review of *Grundgesetze I.*

Around 1900 the young Bertrand Russell was studying the work of Frege. Frege had lamented the lack of serious study of his foundations of numbers, but this was to change. Frege was very confident about his work, with perhaps one exception. We quote from the foreword of *Grundgesetze I*, p. VII:

> A dispute can arise, as far as I can see, only with my basic law concerning the domain of definition (V), which logicians perhaps have not yet expressly enunciated, and yet is what people have in mind, for example, when they speak of extensions of concepts. I hold that it is a law of pure logic. In any event, the place is noted where a decision must be made.

The questionable law[4] (V) (*Grundgesetze I*, p. 36) was

$$\{x : P(x)\} \approx \{x : Q(x)\} \leftrightarrow \forall x\, (P(x) \leftrightarrow Q(x)).$$

One of the propositions Frege derived using (V) was (*Grundgesetze I*, p. 117)

$$P(y) \leftrightarrow y \in \{x : P(x)\}.$$

In June of 1901 Russell discovered that Frege's logical system led to contradictions, for if one lets $P(x)$ be the property $x \notin x$ and lets y be $\{x : x \notin x\}$ we have[5]

$$\{x : x \notin x\} \notin \{x : x \notin x\} \leftrightarrow \{x : x \notin x\} \in \{x : x \notin x\}.$$

[3] Frege describes the truth value of $P \rightarrow Q$ for each of the four possible truth values of P, Q. Thus we almost have the first *truth table;* but there is no table, just a verbal description. Truth tables are usually attributed to Wittgenstein who was well versed in Frege's work, and had gone to study with Russell at Frege's suggestion.

[4] In Frege's book it was written

$$\vdash\!\!\!\!\text{———}\ (\varepsilon\text{'f}(\varepsilon) = \alpha\text{'g}(\alpha)) = (\text{—}\overset{\mathfrak{a}}{\smile}\text{——}\ \text{f}(\mathfrak{a}) = \text{g}(\mathfrak{a})).$$

[5] A slight variation on this applies to $\varphi(P) = \neg P(P)$, namely one gets $\varphi(\varphi) \leftrightarrow \neg\varphi(\varphi)$.

Russell communicated this to Frege in June, 1902, when *Grundgesetze II* was just about ready to appear (Frege had not found a publisher, so he was paying for the publication out of his own pocket). Working on the contradiction during the summer he was able to compose an appendix, pointing out the flaw in his system, and suggesting a remedy. (Years later it was discovered that the remedy reduced the universe to one object!)

The first 160 pages of *Grundgesetze II* (1903) are devoted to criticism of existing approaches to the real numbers. Frege discusses briefly defining the reals using the integer part plus a dyadic expansion for the remainder. He does not do any technical work with this, and indeed says that ordered pairs of integers and dyadic expansions will <u>not</u> be the reals. His final theorem in *Grundgesetze II* is a proof of the commutative law for addition for so called positive classes (which are defined to be classes having several of the properties of the positive reals). His final statement is to say that it remains to find a suitable positive class to develop the reals. So, in the end, Frege has no real numbers to show.

After 500 pages of *Grundgesetze I/II* and 689 propositions one could have hoped for more than the commutative law. Nonetheless Frege made a major contribution to the precision of presentation. Nowadays, when studying formal systems, after seeing how formalization is actually carried out we are usually content to skip over as many of the details as possible, quite the opposite of Frege's approach.

Whitehead & Russell picked up on the work of Frege; they solved the immediate contradictions by *typing* the predicates and variables. Thus one could not use $P(P)$ in their system because P must be of higher type than its argument, nor would one be allowed to use $x \in x$.

CHAPTER 4

The Algebra of Logic: Schröder

The monument to the work initiated by Boole, the algebraization of logic, is the three volumes *Algebra der Logik* by Schröder (1841–1902), which appeared in the years 1890–1910, filling over 2,000 pages. Although the spirit of the subject came from the work of Boole and De Morgan, Schröder's volumes are really a tribute to the work of C.S. Peirce, along with Schröder's contributions. In addition to the substantial job of organizing the literature, the lasting contributions of Schröder's volumes are 1) his emphasis on the Elimination Problem and 2) his fine presentation of the Calculus of Binary Relations.

Volumes I and II are devoted to the Calculus of Classes, with the standard operations of union, intersection and complement, adhering to Boole's arithmetic notation for union (+) and intersection (·). Schröder was very much influenced by Peirce's work, and followed him in making the relation of subclass ($\subseteq$) the primitive notion, whose properties are given axiomatically (what we now call the axioms for a bounded lattice, presented as a partially ordered set), then defining the other operations and equality from it.

One of the historically interesting items in Vol. I is Schröder's discovery that the distributive law does not follow from the assumptions Peirce put on $\subseteq$. Schröder's proof is via a model, and indeed a rather complicated one (based on 990 quasigroup equations). Subsequently Dedekind published his first paper on dualgroups (= lattices) in 1897, giving a much shorter proof using a five element example (to show that a lattice need not be distributive).

The main goal of Schröder's work is stated most clearly in Vol. III, p. 241, where he says that

> getting a handle on the consequences of any premisses, or at least the fastest methods for obtaining these consequences, seems to me to be the noblest, if not the ultimate goal of mathematics and logic.

Schröder is very fond of examples and is only too aware that one can get into computational difficulties with the Calculus of Classes. The examples worked out at the end of Vol. I show how demanding the methods of Jevons and Venn become as the number of variables increases. Such difficulties evidently led him to focus on Elimination. In deriving a conclusion $\Psi(\vec{y})$ from some hypotheses $\Phi(\vec{x},\vec{y})$ about classes $\vec{x},\vec{y}$ it is often the case that some of the classes in the hypotheses do not appear in the conclusion. If one could find a $\Phi_0(\vec{y})$ such that

$$\exists\vec{x}\,\Phi(\vec{x},\vec{y}) \leftrightarrow \Phi_0(\vec{y}),$$

then one could concentrate on the apparently simpler problem of deriving $\Psi(\vec{y})$ from $\Phi_0(\vec{y})$. Finding Φ_0, the *Elimination Problem,* is the recurring theme of Schröder's three volumes.

At the end of his work on this problem for the Calculus of Classes he observed that in some cases of elimination he needed to refer to the number of elements in certain classes, a direction that did not appeal to him. However it was a direction that would later be used by Skolem with success (in the general case considered by Schröder). Indeed, Schröder avoided as much as possible the reference to elements in his formal development of the Calculus of Classes. He had no symbol for membership — that would be first introduced by Peano. When Schröder finally does introduce elements of the domain into his formalism, he uses what we would call singletons, but identifies them with the elements. And he introduces them (Vol. II, §47) not as a primitive concept, but as a defined notion, his definition being equivalent to saying i is an element iff

$$(i \not\approx 0) \wedge \forall x \ \left[i \cap x \approx 0 \vee i \cap x' \approx 0 \right].$$

A convention which can make reading the work of Schröder a bit slow today is his deliberate identification of the notation for the Calculus of Classes and for the Propositional Calculus — an idea clearly due to Peirce. For example he will write (Vol II, p. 10) $(2 \times 2 = 5) = 0$ where we would write $(2 \times 2 = 5) \leftrightarrow F$, where F denotes some canonical false statement. Thus $=$ can mean $\leftrightarrow$, and $\subseteq$ can mean $\rightarrow$. The quantifiers Σ and Π are introduced, following Peirce, and are used also in the Calculus of Classes for $\bigcup$ and $\bigcap$.

Vol. III of Schröder is devoted to the *Calculus of Binary Relations,* pioneered by De Morgan, and largely developed by Peirce. In the Calculus of Classes one works with the subclasses of a domain D, whereas one works with the subclasses of D×D in the Calculus of Binary Relations. One still has the operations $\cup$, $\cap$, $'$, and the constants 0, 1 as in the Calculus of Classes, but there are the additional operations of *converse* ($\breve{}$), *relational product* ($\circ$), and *relational sum* ($\oplus$), as well as a constants for the *diagonal relation* (Δ) and its complement. On p. 16 of Vol. III he discusses relations of n arguments, and says that statements involving such can be rephrased as statements involving binary relations, although the price may be the loss of transparency of meaning.

In contrast to his approach to the Calculus of Classes, Schröder develops the Calculus of Binary Relations making extensive use of the primitive notion of membership (again, following the development of Peirce). Schröder had no symbol for membership ($\in$), as we said above — to say $(i, j) \in A$ he would write $A_{ij} = 1$. He says that the Calculus of Binary Relations is determined by 29 properties (Vol. III, §3), which we give in our language, but we retain his numbering:

1) $a \approx b \leftrightarrow a \subseteq b \wedge b \subseteq a$	
2) $0 \subseteq 0, 0 \subseteq 1, 1 \subseteq 1, 1 \not\subseteq 0$	
3) $0 \cdot 0 \approx 0 \cdot 1 \approx 1 \cdot 0 \approx 0, 1 \cdot 1 \approx 1$	$1 + 1 \approx 1 + 0 \approx 0 + 1 \approx 1, 0 + 0 \approx 0$
4) $1' \approx 0$	$0' \approx 1$
5) $a \approx \bigcup_{ij}\{(i,j) : a(i,j)\}$	
6) $1(i,j)$	$\neg 0(i,j)$
7) $\Delta(i,j) \leftrightarrow i \approx j$	$\Delta'(i,j) \leftrightarrow i \not\approx j$
8) $\rho_1(i)(j,k) \leftrightarrow i \approx j$	
9) $\rho_2(i,j)(h,k) \leftrightarrow i \approx h \wedge j \approx k$	
10) $(a \cap b)(i,j) \leftrightarrow a(i,j) \wedge b(i,j)$	$(a \cup b)(i,j) \leftrightarrow a(i,j) \vee b(i,j)$
11) $a'(i,j) \leftrightarrow \neg a(i,j)$	
12) $(a \circ b)(i,j) \leftrightarrow \exists k\, [a(i,k) \wedge b(k,j)]$	$(a \oplus b)(i,j) \leftrightarrow \exists k\, [a(i,k) \vee b(k,j)]$
13) $a\breve{}(i,j) \leftrightarrow a(j,i)$	
14) $a \subseteq b \leftrightarrow \forall ij\, [a(i,j) \rightarrow b(i,j)]$	
15) $(\bigcap_u \varphi)(i,j) \leftrightarrow \forall_u\, \varphi(i,j)$	$(\bigcup_u \varphi)(i,j) \leftrightarrow \exists_u \varphi(i,j).$

The Calculus of Binary Relations is incredibly more complex than that of classes. A considerable portion of this volume deals with the terms $t(x)$ in a single unknown — he is able to find 256 distinct ones before abandoning the problem. There are actually infinitely many distinct terms in one variable. Schröder also incorporated into his study of binary relations Pierce's notation for union and intersection, Σ and Π, ranging over all the binary relations, notation which, as before, could also be used for quantifiers.

Thanks to the fact that $1 \circ x \circ 1$ is a term which takes the value 0 if x is 0, and 1 otherwise (Vol. III, p. 147), Schröder can reduce any finite set of atomic and negated atomic formulas to a *single equation* $t(\vec{x}) \approx 0$, provided the domain has at least two elements (which he always requires). For the case of a single variable, $t(x) \approx 0$, he shows (Vol. III, p. 165) the general solution can be expressed by the following, given a particular solution a:

$$x \approx [a \cap (1 \circ t(u) \circ 1)] \cup [u \cap (0 \oplus t(u)' \oplus 0)].$$

Unfortunately, as Schröder notes, this is not very useful, for if $t(u) \approx 0$ this gives $x \approx u$, and otherwise it gives $x \approx a$.

He also introduced toward the end of the third volume the use of Σ and Π over the elements of the domain, but in a roundabout way. Namely he would identify an individual i of the domain D with the binary relation $\{i\} \times D$, which we have called $\rho_1(i)$ in item 8), and then let Σ and Π range over such relations. Such individuals are determined by the equation $\Delta' \circ x \circ 1 \approx x'$ (Vol. III, p. 408). Relations which are singletons $\{(i,j)\}$, which we called $\rho_2(i,j)$ in item 9) are determined by the equation $\Delta' \circ x \circ \Delta' \approx (x' \oplus 0) \cap (0 \oplus x')$ (Vol. III, p. 427). Also he embedded the study of classes into the study of binary relations by identifying a class A with $A \times D$. Relations associated with classes are then determined by $x \circ 1 \subseteq x$ (Vol. III, p. 450).

One of the few applications of the Calculus of Binary Relations given by Schröder to other areas of mathematics is the formulation and proof of a slight generalization of Dedekind's proof of the induction theorem (for chains). Also there is a formulation of Cantor's basic ideas on infinite classes. For example (Vol. III, p. 587) a binary relation x is a *one-to-one correspondence* iff it satisfies $x \circ x^{\smile} \cup x^{\smile} \circ x \subseteq \Delta$, a *function* iff it satisfies $1 \approx 1 \circ x \cap (\Delta \oplus x')$.

In Vol. III, p. 278, we see Schröder posing the question as to whether the algebra of binary relations will really provide the foundation for mathematics:

> An important but difficult question is that of the completeness of our algebra of binary relations, in particular the question of whether this discipline with its six operations suffices for all purposes of the pure and applied theories (in particular, for the logic of binary relations).

(By 1915 Löwenheim will have no doubts about the expressive power of binary relations.)

One of his claims towards the end of Vol. III, p. 551, involved what could be interpreted as a general method for passage from an expression involving quantifiers over individuals to one which does not. The possibility of expressing first-order formulas in the language of binary relations with equality as equations in the Calculus of Binary Relations, without resort to the use of ρ_1 above, would be picked up by Korselt, who showed that the statement "there exists four distinct elements" could not be so expressed. Löwenheim would turn to an examination of models of first-order statements in relational logic with equality, using the notation of Schröder, and this focused attention in mathematical logic on what we now call model theory.

The Calculus of Binary Relations is not nearly so widely known as the Calculus of Classes. It forms a substantial part of Vol. I of *Principia Mathematica,* and it has been a source of fundamental research under the name of *Relation Algebras* in the school led by Tarski. In 1964 Monk proved that, unlike the Calculus of Classes, there is no finite equational basis for the Calculus of Binary Relations. The recent book *A Formalization of Set Theory without Variables* (1988) by Tarski and Givant shows that relation algebras are so expressive that one can carry out first-order set theory in their equational logic.

One last note: Schröder's name is best known in connection with the famous Schröder-Bernstein theorem. Actually, the proof that Schröder gave in 1896 was full of holes (according to Fraenkel), and it was Bernstein, a student of Cantor, who produced a correct proof the next year.

CHAPTER 5

New Notation for Logic: Peano

Giuseppe Peano (1858–1932)
1889 - The principles of arithmetic, presented by a new method

A basic education in mathematics will include three references to Peano — his axioms for the natural numbers, his space filling curve, and the solvability of $y' = f(x, y)$ for f continuous. Also his influence on mathematical logic was substantial, largely thanks to his young disciple Bertrand Russell.

Peano's first work on logic (1888) showed that the calculus of classes and the propositional calculus were, up to notation, the same. Next, in *The principles of arithmetic, presented by a new method* (1989), he presented logic and set theoretic notation along with the basic axioms of logic and set theory (including abstraction), and stated his convictions about the possibility of presenting any science in a purely symbolic form. As evidence for this he worked out portions of arithmetic, giving the famous Peano axioms, after stating in the preface:

> In addition the recent work by R. Dedekind *Was sind und was sollen die Zahlen?* (Braunschweig, 1888), in which questions pertaining to the foundations of numbers are acutely examined, was quite useful to me.

As to the nature of his new method we again quote from the preface:

> I have indicated by signs all the ideas which occur in the fundamentals of arithmetic. The signs pertain either to logic or to arithmetic $\cdots$.
>
> I believe, however, that with only these signs of logic the propositions of any science can be expressed, so long as the signs which represent the entities of the science are added.

He starts off with the natural numbers N and the successor function given. His axioms are

- 1 is not the successor of any number Peano's Axioms
- if $m' = n'$ then $m = n$
- (induction) if $X \subseteq N$ is closed under successor, and if $1 \in X$, then $X = N$.

Peano skipped over any attempt to define the natural numbers in logic, thus bypassing certain philosophical issues that mathematicians tend to view as being incapable of precise formulation, and concentrated on the manipulation of symbols, something mathematicians find most agreeable.

Thus we see that Peano's axioms are Dedekind's theorems.[1] This approach would be given its most popular form in Landau's *Grundlagen der Analysis*, 1930, (an excellent book for beginning mathematical German), starting with the set of natural numbers N with a successor function obeying the Peano axioms and proceeding to develop the integers, the rationals, the reals and finally the complex numbers with + and $\times$, and proving the basic laws of these operations in 158 pages and 301 propositions.

[1]The subtle point of first showing that the f_m's exist before defining addition on the natural numbers was overlooked by Peano, and later by Landau who was following Peano (who had defined $\leq$ *after* defining +). Grundjot discovered this flaw in Landau's work, and repairs were made following ideas of Kalmar.

Peano's axioms, with induction cast in first-order form, and with the recursive definitions of $+$ and $\times$, would form Peano Arithmetic (PA), a popular subject of mathematical logic. In particular one could derive all known theorems of number theory[2] which could be written in first-order form from Peano Arithmetic; finally, in the mid 1970's Paris & Harrington found a 'natural' example of a first-order number theoretic statement which was true, but could not be derived from PA.

The ambitious *Formulario* project was announced in 1892, the goal being to translate mathematics into Peano's concise and elegant notation. The first edition of this work appeared in 1895, the fifth in 1908. The latter was nearly 500 pages, covering approximately 4,200 theorems on arithmetic, algebra, geometry, limits, differential calculus, integral calculus, and the theory of curves. One could well imagine the satisfaction Peano would enjoy today as director of a mathematics database project.

In 1896 Frege communicated his criticism of Peano's foundations — in particular the lack of clearly stated rules of inference. He doubted that Peano's system could do more than express mathematical theorems. Peano's response was that the ability to give brief and precise form to mathematical theorems would make the importance of his work clear.

Aside from his catalytic influence on Russell we can see that Peano's main contributions to the foundations of mathematics were

i. An elegant notation, which has strongly influenced the symbols used today (e.g. $\cup$, $\cap$, $\subset$, $\supset$, $\in$, and $\emptyset$),
ii. adopting the axiomatic approach to all mathematics (not getting involved in the origins of numbers, etc.), and
iii. the belief that his formalization of logic would suffice for expressing the theorems of any field of science once the symbols appropriate to that field were added.

It is surprising to realize that Peano was the first to introduce distinct notation for *subset of* and *belongs to.*

[2] Of course Gödel had found a statement in first-order number theory which could not be derived from PA, but it was not the sort of statement one would encounter in traditional number theory.

CHAPTER 6

Set Theory: Zermelo

Ernst Zermelo (1871–1953)
1904 - A proof that every set can be well ordered
1908 - Investigations in the foundations of set theory

In 1900 Hilbert had stated at the International Congress of Mathematicians that the question of whether every set could be well-ordered was one of the important problems of mathematics. Cantor had asserted this was true, and gave several faulty proofs. Then, in 1904, Zermelo published a proof that every set can be well-ordered, using the Axiom of Choice. The proof was regarded with suspicion by many. In 1908 he published a second proof, still using the Axiom of Choice. Shortly thereafter it was noted that the Axiom of Choice was actually equivalent to the Well-Ordering Principle (modulo the other axioms of set theory), and subsequently many equivalents were found, including Zorn's Lemma[1] and the linear ordering of sets (under embedding).

But more important for mathematics was the 1908 paper on general set theory. There he says:

> Set theory is that branch of mathematics whose task is to investigate the fundamental notions *number*, *order*, and *function* $\cdots$ to develop thereby the logical foundations of all of arithmetic and analysis $\cdots$. At present, however, the very existence of this discipline seems to be threatened by certain contradictions $\cdots$. In particular, in view of Russel's *antinomy* $\cdots$ it no longer seems admissible today to assign to an arbitrary logically definable notion a set, or class, as its extension $\cdots$. Under these circumstances there is at this point nothing left for us to do but to proceed in the opposite direction and, starting from set theory as it is historically given, to seek out the principles required for establishing the foundation of this mathematical discipline $\cdots$. Now in the present paper I intend to show how the entire theory created by Cantor and Dedekind can be reduced to a few definitions and seven principles, or axioms, $\cdots$. I have not yet even been able to prove rigorously that my axioms are consistent $\cdots$.

Zermelo starts off with a domain D of objects, among which are the sets. He includes $\approx$ and $\in$ in his language, and defines $\subseteq$. He says that an assertion φ is *definite* if

> the fundamental relations of the domain, by means of the axioms and universally valid laws of logic, determine whether it holds or not.

A formula $\varphi(x)$ is definite if for each x from the domain it is definite. Then the axioms of Zermelo are (slightly rephrased):

i. (Axiom of extension) If two sets have the same elements then they are equal.
ii. (Axiom of elementary sets) There is an empty set $\emptyset$; if a is in D then there is a set whose only member is a; if a, b are in D then there is a set whose only members are a and b.
iii. (Axiom of separation) Given a set A and a definite formula $\varphi(x)$ there is a subset B of A such that $x \in B$ iff $x \in A$ and $\varphi(x)$ holds.

[1] Originally proved by K. Kuratowski (1923) and R.L. Moore (1923); this was rediscovered by M. Zorn in 1935, and credited to him by Bourbaki.

iv. (Axiom of power set) To every set A there corresponds a set $P(A)$ whose members are precisely the subsets of A.
v. (Axiom of union) To every set A there corresponds a set $U(A)$ whose members are precisely those elements belonging to elements of A.
vi. (Axiom of choice) If A is a set of nonempty pairwise disjoint sets then there is a subset $C(A)$ of $U(A)$ which has exactly one member from each member of A.
vii. (Axiom of infinity)[2] There is at least one set I such that $\emptyset \in I$, and for each $a \in I$ we have $\{a\} \in I$

In the next few pages he defines $A \sim B$ (i.e., A and B are of the same cardinality),[3] derives Cantor's theorems that the cardinality of A is less than that of $P(A)$, and that every infinite set has a denumerably infinite subset.

Let us use the traditional notation $\{a\}$ for *singleton*, and $\{a, b\}$ for *doubleton*. Let $\bigcup A$ be the *union* of the elements of A. One can define the basic set operations by

- $A \cap B = \{x \in A : x \in B\}$
- $\bigcap A = \{x \in \bigcup A : x \in a \text{ for all } a \in A\}$
- $A \cup B = \bigcup\{A, B\}$

We can let the set ω of *nonnegative integers*[4] be defined to be the smallest set A which contains $\emptyset$ and is closed under $x \in A \Rightarrow \{x\} \in A$. Then one has a successor operation $x' = \{x\}$ on ω which satisfies Peano's Axioms, so one can develop the number systems. Well, actually you need to have functions to do this. Using Kuratowski's definition of *ordered pair*, namely

$$(a, b) = \{\{a\}, \{a, b\}\},$$

one can prove (from Zermelo's axioms) that

$$(a, b) \approx (c, d) \qquad \Longleftrightarrow \qquad a \approx c \ \&\ b \approx d.$$

Then we can define
the *Cartesian product* of A and B:

$$A \times B = \{u \in P(P(A \cup B)) : x \in u \text{ iff } x \approx (a, b) \text{ for some } a \in A,\ b \in B\};$$

the set of *relations* between A and B:

$$Rel(A, B) = P(A \times B)$$

and the set of *functions* from A to B:

$$Func(A, B) = \{f \in Rel(A, B) : \forall a \in A\, \exists! b \in B \ \ (a, b) \in f\}.$$

With these definitions we can now translate Dedekind's development of the natural numbers, rationals, reals and complex numbers into Zermelo's set theory, and prove the basic properties about the operations $+, \times$; also we can carry out Cantor's study of sets, especially cardinals and ordinals. If one then wants to do analysis, for example integration on $[a, b]$, one takes the definite integral $\int_a^b \, dx$ as a certain element of $Func(\mathcal{F}, \mathbf{R})$, where $\mathcal{F}$ is a suitable subset of $Func([a, b], \mathbf{R})$; and then you can prove the fundamental theorem of calculus, etc., *all from Zermelo's seven axioms*.

However Zermelo's axioms had some obvious shortcomings. Skolem noted in Some remarks on axiomatized set theory, 1922, that improvements were needed, in particular

- A definition of a *definite* property

[2]Later versions would use $a \cup \{a\} \in I$.

[3]Zermelo did not used ordered pairs. He starts with disjoint A and B and considers the set M in $P(A \cup B)$ consisting of doubletons with exactly one element from each of A and B. Then he looks for R in M which provide a 1–1 correspondence.

[4]Zermelo first looked at numbers in later articles.

- An axiom to ensure some reasonably large sets.[5]

For the first he suggested one use first-order formulas, i.e., formulas made up from $\forall$, $\exists$, &, $\vee$, $\neg$, $\in$, $\approx$, and variables. (Skolem used the notation of Schröder.) Regarding the second he noted that if one has a model of set theory then let M_0 be the union of the $P^{(n)}(\{\emptyset\})$, the n-fold iterated power set applied to the set $\{\emptyset\}$, n an integer; and then let M be the union of the $P^{(n)}(M_0)$. This gives a rather small submodel, i.e., a small collection of sets that satisfies all of Zermelo's axioms. In particular the set $\{P^{(n)}(\omega) : n \in N\}$ is not a set in this model. To guarantee the existence of such interesting (not too large) sets he suggested the

- (Axiom of Replacement) if $\varphi(x, y)$ is a definite formula such that for every x there is at most one y making it true, then, for every set A there is a set B such that $y \in B$ holds iff there is an x in A such that $\varphi(x, y)$ holds.

Thus one can replace elements of A in the domain of φ with corresponding elements of B. Replacement is actually stronger than separation, for if one is given a set A and a definite property $\theta(x)$, define $\varphi(x, y)$ to be the definite property $y \approx x \,\&\, \theta(x)$. Then $\varphi(x, y)$ applied to A gives $\{x \in A : \theta(x)\}$.

The resulting set theory is called Zermelo-Fraenkel set theory with Choice (ZFC)[6].

Skolem pointed out that Zermelo's approach to set theory took us away from the natural and intuitive possibilities (like Frege's), and thus, as an artificial construction, carried a loss of status:

> Furthermore, it seems to be clear that, when founded in such an axiomatic way, set theory cannot remain a privileged logical theory; it is then placed on the same level as other axiomatic theories.

In 1917 Miramanoff noted the possibility of models with infinite descending chains $\cdots \in y \in x$. Such possibilities led von Neumann to formulate the Axiom of Regularity, namely if $x \not\approx \emptyset$ then for some $y \in x$ we have $x \cap y \approx \emptyset$. This axiom is not always used — it seems to have no application to mathematics, but it does make some proofs and definitions easier, e.g., that of an ordinal.

The treatment of ordinals evolved from Cantor's *abstraction* from well-ordered sets to *equivalence classes* of well-ordered sets to *representative* well-ordered sets. The last step was initiated by von Neumann in Transfinite numbers, 1923, and reached its modern brief form (assuming regularity) in the work of Raphael Robinson (1937), namely an *ordinal* is a transitive set (under $\in$), all of whose elements are also transitive sets.

ZFC requires an infinite number of first-order axioms. The open question of whether one could develop a set theory with a finite number of axioms was answered in the affirmative by J. von Neumann in 1925. Actually he used functions rather than sets as his primitive notion, and the current first-order version is due to the reworking in the late 1930's by (mainly) Bernays as well as Gödel, and called *von Neumann-Bernays-Gödel* set theory, abbreviated to NBG set theory.

Exercises

0.2. If R is a set of ordered pairs, show (using the axioms of ZFC) that the domain and the range of R are also sets.

0.3. Given ω and $+$ as sets, describe a set A and a first-order property $\varphi(x)$ such that the collection of integers Z is $\{x \in A : \varphi(x)\}$ (and thus it is a set by the axiom of separation). [Think of Z as sets of equivalence classes of ordered pairs of integers, where two pairs of integers are in the same class iff the first coordinate minus the second coordinate is the same in each case.]

0.4. [Kuratowski] Let us define (x, y) to be the set $\{\{x\}, \{x, y\}\}$. Use the axioms of Zermelo to prove that $(x, y) \approx (u, v) \leftrightarrow x \approx u \,\&\, y \approx v$.

Could we use the definition $\{x, \{x, y\}\}$ and achieve the same?

[5] Fraenkel published similar observations the same year.

[6] For a leisurely treatment, i.e., in the spirit of Zermelo's original paper, see Halmos' *Naive Set Theory*.

0.5. Show that $x \notin x$ is a theorem in Z+(R)[7]

0.6. [R. ROBINSON] A set x is said to be *transitive* if $u \in v \in x$ implies $u \in x$. Suppose x and every element in x is a transitive set. Show that x is well-ordered by $\in$ (using Z+(R)).

[7] Zermelo's set theory with the axiom of regularity.

CHAPTER 7

Countable Countermodels: Löwenheim.

Löwenheim (1878–1940)
1915 - Über möglichkeiten im Relativkalkül

Löwenheim's 1915 paper

- opened the door to the serious study of model theory by bringing in the *size* of an infinite model of a *first-order sentence*
- gave a canonical procedure to build a countable countermodel to a *first-order sentence* – this initiated some of the most popular work in automated theorem proving (mainly attributed to Herbrand)
- showed that statements in the *first-order predicate calculus (with equality) for monadic predicates* which were valid in finite domains were valid in all domains
- pointed out that *first-order relational logic with equality* was adequate to express all mathematical problems
- showed that one only needed *binary* relations for the previous item.

This paper was such a turning point in the development of logic that it is worth discussing each section. In particular we will later see how much Skolem was influenced by it.

Section 1
Löwenheim presents his framework for the paper, namely working in the calculus of relations as presented by Schröder. He introduces first-order statements under the name *numerical equations.*[1]

Section 2
Schröder seems to claim that first-order formulas are equivalent to quantifier-free formulas. Löwenheim sketches Korselt's argument that this cannot be the case: a first-order statement asserting the existence of 4 distinct elements cannot be expressed by an equation in the Calculus of Relations.

After thus establishing the usefulness of quantifiers in first-order expressions he turns to a remarkable blending of logic and set theory by showing that

> given an infinite domain D and a first-order statement φ which holds in all finite structures, but not in all structures, then it is impossible for φ to hold in all structures on D.

In 1920 this would be restated and reproved, by Skolem, to give the famous Löwenheim-Skolem theorem.

The method by which Löwenheim proved this theorem is as fascinating as the theorem itself (indeed more fascinating for those in automated theorem proving). Given a first-order statement φ, he first puts $\neg\varphi$ into a *normal form* by treating a universal quantifier as an AND over the domain,

[1]The notion of a first-order property (i.e., quantifiers can only range over the elements of the domain) was introduced, in Volume III of Schröder's Algebra of Logic, as part of his study of the Calculus of Relations. Schröder was more interested in quantifying over the relations than over the domain elements, and he did little with the first-order statements.

and an existential quantifier as an OR over the domain; and then distributes to get a disjunctive form.[2].

From the normal form one takes the universally quantified part ψ and uses the fact that $\neg\varphi$ has a model on a given domain iff ψ has. (ψ will later become the *Skolemized form* of $\neg\varphi$). Thus φ has a countermodel [on a given domain] iff ψ is satisfiable [on that domain].

Next Löwenheim gives a canonical procedure to build a finitely branching tree of finite partial structures, successive nodes being extensions of previous nodes, such that:

1. either every branch has a node which does not "satisfy" ψ, and this will imply that ψ, hence φ, cannot be satisfied; or
2. some branch is such that every node "satisfies" ψ. Then the union of the nodes along the branch gives a countable (perhaps finite) model of ψ, and thus a countable countermodel to φ.

He claims that he can use his results to establish some independence and dependence results for various systems of axioms of the Calculus of Classes. He only shows how one such system (of Müller) can be expressed in first-order form (using a binary predicate for subsumption), and says he will give details of the independence proofs in a later paper (which never appeared).

Löwenheim goes on to show that such a theorem cannot be proved for higher-order logic because one can write down a sentence which says the domain is not finite or countably infinite. Thus for the first time a powerful distinction is made between first-order and higher-order logics.

Section 3

The main theorem of this section is that

> a first-order sentence with only monadic relations (i.e., relations with only one argument) which is true on all finite domains (no matter how the relations are interpreted) must be true on all domains.

At the end of this section he gives, by example, his algorithm to determine if such a sentence is true on all domains.

Section 4

Finally he turns to the expressive strength of the first-order calculus of relations, and says that anyone familiar with the development of logic as in Principia Mathematica can see that all mathematical assertions can be expressed by first-order statements in the calculus of relations. Then he goes on to show in some detail that it suffices to use only *binary* relations. (The last statement had been made, without justification, by Schröder in his Vol. III.)

[2]This procedure is described in Schröder's third volume. By avoiding the digression through what appears to be an infinitary language – yielding a result which depends on the domain – Skolem introduced his presentation of this normal form in 1928. The process of creating this normal form is now called *Skolemizing*

CHAPTER 8

Principia Mathematica: Whitehead and Russell

A.N. Whitehead (1861–1947) and B. Russell (1872–1970)
1910–1913 - Principia Mathematica I–III

Russell met Peano at the 1900 International Congress of Mathematicians in Paris, and was captivated by Peano's work on foundations. And, starting in 1900, he was studying the *Grundgesetze I* of Frege. This led to his discovery of the famous contradiction in Frege's system in June, 1901, while writing his *Principles of Mathematics (1903).* Nonetheless, Russell and Whitehead, who started their joint work on foundations in 1900, would carry out the program of Frege to a significant extent, namely the seamless development of mathematics from a few clearly stated axioms and rules of inference in pure logic. However they opted for the more modern notation of Peano instead of Frege's Begriffsschrift. Their work, *Principia Mathematica,* filled three volumes, almost 2,000 pages, and appeared in the years 1910–1913. Their approach was essentially that of Frege, to define mathematical entities, like numbers, in pure logic and then derive their fundamental properties. Indeed their definition of natural numbers was essentially that of Frege, but unlike him, they opted to avoid the philosophical aspects and justifications. In the preface they say

> We have avoided both controversy and general philosophy, and made our statements dogmatic in form $\cdots$.
>
> The general method which guides our handling of logical symbols is due to Peano. His great merit consists not so much in his definite logical discoveries nor in the details of his notations (excellent as both are), as in the fact that he first showed how symbolic logic was to be freed from its undue obsession with the forms of ordinary algebra, and thereby made it a suitable instrument for research $\cdots$.
>
> In all questions of logical analysis, our chief debt is to Frege.

The main innovation of *Principia Mathematica* was to introduce a stratification of Frege's formulas into types, and to use this to restrict which of Frege's formulas would be permitted in their logic. The key idea was that a formula φ could not be substituted for a variable x in a formula ψ unless the variable x was of the appropriate type. Thus, returning to Frege's troublesome theorem

$$P(y) \leftrightarrow y \in \{x : P(x)\},$$

the types restriction would prevent the substitution of $x \notin x$ for *both* of the variables P and y since P is of a higher type than its argument y. Indeed all the known paradoxes were avoided by using types.

Having salvaged Frege's logic, they proceeded to develop some of the elementary theorems of mathematics, covering far more ground than Frege — however we note that they quickly adopted the convention of leaving out easy steps of proofs — and, at the same time, falling far short of the list of theorems in Peano's *Formulario.* Let us briefly sketch the topics covered in *Principia Mathematica:*

Vol. I: Axioms and rules of inference for their higher order logic; elementary results on classes and binary relations (e.g., the study of union, intersection, domain, range, one-to-one, onto, converse, composition, restriction); the definition of the numbers 1 (p. 347) and 2; a discussion of Zermelo's Well-ordering Theorem and the Axiom of Choice, and choice functions; the Schröder-Bernstein theorem; the transitive closure of a relation.

Vol. II: Cardinal numbers and their arithmetic; finite numbers; the arithmetic of binary relations; linear orderings; Dedekind orderings; limit points; continuous functions.

Vol. III: Well-orderings; equivalence of the Axiom of Choice with the Well-ordering Axiom; the $\aleph$'s; dense orderings; orderings like the rationals; orderings like the reals; the integers, rationals, and reals; measurement; measurement modulo a quantity.

A fourth volume, on geometry, never appeared. Although the above topics may look like a small fragment of mathematics, nonetheless Russell and Whitehead had carried the dream of Frege far enough, and in a transparent enough symbolism, that the possibility of developing all of mathematics from a few axioms and rules was made clear. Future developments would focus on the best way to do this, plus efforts to guarantee that one would not find any contradictions. Most important for future developments was the fact that the great leader of mathematics Hilbert would become heavily involved in mathematical logic.

So by 1931 Gödel could boast of two formal systems which, with a few axioms and rules, could encompass all known mathematics, namely Principia Mathematica and the Zermelo-Fraenkel axiom system.

Nonetheless, if one of these systems is consistent, then Gödel showed it would not be strong enough to prove all first-order truths about the nonnegative integers (using $+$ and $\times$ as the operation symbols). And in 1936 Church would show that Peano Arithmetic, if consistent, is undecidable (Rosser improved this in 1936 to show any consistent extension of Peano Arithmetic is undecidable).

CHAPTER 9

Clarification: Skolem

Thoralf Skolem (1887–1963)
1919 - Untersuchung über die Axiome des Klassenkalküls und über Produktations– und Summationsprobleme, welche gewisse Klassen von Aussagen betreffen
1920 - Logisch–kombinatorische Untersuchungen über die Erfüllbarkeit und Beweisbarkeit mathematischen Sätze nebst einem Theoreme über dichte Mengen
1922 - Einige Bemerkung zur axiomatischen Begrundung der Mengenlehre
1928 - Über die mathematische Logik

The 1919 paper

Skolem, like Löwenheim, adopts the notation of Schröder. The 1919 paper has three important parts:

- He gives a thorough analysis of the dependence/independence of the various axioms for the Calculus of Classes due to Peirce, as presented in Schröder, using simple structures which he can easily sketch.[1]
- Skolem shows that by adding predicates for "has at least n elements" to the language of the Calculus of Classes he is able to eliminate quantifiers. As we mentioned Schröder devoted much effort to the elimination problem for the Calculus of Classes. However it is first in Skolem's paper that we see it clearly formulated as taking a formula of the form $\exists x\, \psi(x, \vec{y})$, where ψ is quantifier-free, and finding an equivalent quantifier-free formula φ. Skolem notes that this means that every first-order formula is then equivalent to a quantifier-free formula. This is of course the modern meaning of the elimination of quantifiers.

  And Skolem notes that the final form of such a quantifier-free formula is equivalent to a Boolean combination of assertions about the sizes of the constituents. Thus he has a precise handle on the expressive power of the Calculus of Classes.[2] Because of the clarity of Skolem's work he is often regarded as the inventor of quantifier elimination. This seems rather unfair to the pioneering work of Boole and Schröder.
- Finally Skolem shows that one can easily translate back and forth between the first-order Calculus of Classes and first-order monadic predicate logic. In particular it follows that a statement can only assert a Boolean combination of statements about the size of the universe. Consequently if a statement in the first-order monadic predicate logic holds for all finite domains, it must hold for all domains. This proves the assertion in Löwenheim's section three.

The 1920 paper

Section 1

In this paper Skolem first introduces what is now called the *Skolem normal form*, namely to each

[1]This reminds one of Löwenheim's claim in section 2 of his paper, that he would analyze the dependence/independence of several axiom systems for the Calculus of Classes.

[2]Schröder had worked out some simple cases involving a couple of negated equations — and sketched a combinatorial procedure for the elimination in general. However, because he wanted to keep precise track of all the combinations involved he failed to note the nature of the final result — instead he dwelt on the incredibly complicated nature of the calculations that needed to be done.

first-order statement φ he associates an $\forall\exists$ sentence ψ which is obtained via a simple combinatorial procedure, and has the essential property that φ is satisfiable on a given domain iff ψ is satisfiable on the same domain. He shows that if an $\forall\exists$ statement is satisfiable on an infinite domain, it must also be satisfiable on a countable subdomain. Thus he has a slick proof of Löwenheim's theorem on countermodels. His proof technique is completely different from that of Löwenheim, making use of the notion of "subuniverse generated by" which he has learned from Dedekind's work. For model theorists it gives more information than Löwenheim's theorem — but it requires stronger methods, namely the Axiom of Choice. Also he generalizes Löwenheim's theorem to cover a countable set of statements. This will later be needed for the Skolem Paradox in set theory.

Section 2

Now Skolem turns to an analysis of the Calculus of Groups as presented in Schröder — in modern terminology this is just lattice theory, whereas the Calculus of Classes is the theory of power sets, as Boolean algebras. He is interested in determining the first-order consequences of the Calculus of Groups — in modern terminology he is studying the (first-order) theory of lattices.[3] His main achievement here is to give an algorithm to decide which *universally* quantified statements are consequences of the lattice axioms.[4]

Section 3

In this section he looks at some consequences of first-order axioms for geometry.

Section 4

Shifting gears he shows that the $\aleph_0$-categoricity of $(Q, <)$, the rationals with the usual ordering (proved by Cantor), could be generalized by adding finitely many dense and cofinal subsets Q_i which partition Q.

The 1922 paper

We have already spoken about the importance of this paper in the section on set theory — the recommendation that first-order properties be used, that a stronger axiom (replacement) be added, and the observation that if Zermelo's set theory has a model, it has a countable model by the Löwenheim-Skolem theorem.

Also in this paper he returns to the proof of Löwenheim's countermodel theorem, noting that his 1920 proof had used the Axiom of Choice; and now, in a paper on set theory, he finds it appropriate to eliminate this usage. He gives a very clean version of Löwenheim's proof for a first-order statement (without equality). Except for the use of his normal form from the 1920 paper, it is essentially Löwenheim's proof, the canonical construction of a countermodel.

The 1928 paper

This paper is based on a talk Skolem gave earlier that year. And in it we see him describe an alternative to the ususal method of "derivation from axioms" that has become common in logic, an alternative that he suggests is superior. Actually, he only gives an example, but the idea is clearly that of Löwenheim, namely to use the countermodel construction. It is surprising that he doesn't mention Löwenheim here.

The technique of replacing the existential quantifiers by appropriate functions symbols to get a universal sentence is clearly explained by example — and becomes known as Skolemization. He goes on to show how one can build up the elements of the potential countermodel using these

[3]The fact that the *Gruppenkalul* is nothing other than lattice theory seems to have escaped everyone's attention.

[4]We now know that the first-order theory of lattices is undecidable, so a general algorithm would be impossible.

Skolem functions — this will become known as the Herbrand universe. Skolem's example does not indicate the full power of Löwenheim's method because he does not deal with equality.

CHAPTER 10

Grundzüge der theoretischen Logik: Hilbert & Ackermann

D. Hilbert (1862–1943) and W. Ackermann (1896–1962) 1928 - Grundzüge der theoretischen Logik

Hilbert gave the following courses on logic and foundations in the period 1917–1922:

Principles of Mathematics	Winter semester	1917/1918
Logic Calculus	Winter semester	1920
Foundations of Mathematics	Winter semester	1921/22

He received considerable help in the preparation and eventual write up of these lectures from Bernays. This material was subsequently reworked by Ackermann into the book *Grundzüge der Theoretischen Logik* (1928) by Hilbert and Ackermann. The book was intended as an introduction to mathematical logic, and to the forthcoming book of Hilbert and Bernays[1] (dedicated essentially to the study of first-order number theory). In 120 pages they cover:

Chap. I: the propositional calculus,
Chap. II: the calculus of classes,
Chap. III: (many-sorted) first-order logic of relations (without equality), and
Chap. IV: the higher order calculus of relations (without equality).

Let us make a few comments about each of these chapters.

Chapter I discusses the basic connectives $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$ (they use the notation &, $\vee$, $\rightarrow$, $\approx$, $^{-}$), the commutative, associative and distributive laws, and reducing the number of connectives needed. They say (p. 9):

> As a curiosity let it be noted that one can get by with a single logical sign, as Sheffer showed.

This is closer to the importance now attached to Sheffer's discovery than Whitehead and Russell's statement in the second edition of *Principia Mathematica* that Sheffer's reduction of the propositional logic to a single binary connective was the most important development in logic since their first edition had appeared. Indeed Whitehead and Russell added the necessary text to *Principia* to reduce their development, based on $\vee$ and $\neg$, to the Sheffer stroke.

Next Hilbert and Ackermann show how to put propositions in conjunctive, or disjunctive, normal form, and show how one can use this to describe all the consequences of a finite set of propositions. Then they discuss how one determines if a statement is a tautology (they say *universally valid*), or satisfiable.

They follow *Principia* in the axiomatization of the propositional calculus, using the work of Bernays (1926) to reduce the list of axioms from 5 to 4. Turning to the questions associated with the axiomatic method (*Grundzüge*, p. 29) they say:

> The most important of the questions which arise are those of consistency, independence, and completeness.

[1]Thanks to the work of Gödel (1930, 1931) this project was delayed, and it expanded into two volumes (1934, 1938).

After noting that one can derive precisely the tautologies in their system, the solutions of these questions for this propositional calculus, due to Bernays (1926), are presented. The completeness result is the strong version, namely that adding any non-tautology to the axioms permits the derivation of a contradiction, and hence the derivation of all propositional formulas.

The brief treatment of the calculus of classes in Chapter II is actually a nonaxiomatic version of the first-order logic of unary predicates, a system which is considerably more expressive than the traditional calculus of classes, and in which one can formulate Aristotle's syllogisms.

Chapter III starts off with a famous quote of Kant:

> It is noteworthy that till now it [logic] has not been able to take a single step forward [beyond Aristotle], and thus to all appearances seems to be closed and compete.

Then they say about the logic of Aristotle:

> It fails everywhere that it comes to giving a symbolic representation to a relation between several objects $\cdots$.

and that such a situation (*Grundzüge*, p. 44):

> exists in almost all complicated judgements.

The first-order system they develop uses relations and constant symbols, but equality is not a part of the logic. Furthermore, their relations are many-sorted, both in their examples and formalized logic (*Grundzüge*, pp. 45, 53, 70). In the second edition (1937) the formal version is one-sorted, but the examples are still many-sorted; and a technique using unary domain predicates for converting from many-sorted to one-sorted is given (*Grundzüge*, p. 105).

As one example of how to express mathematics in their formal system they turn to (one-sorted) natural numbers, using two binary relation symbols $=$ and F (for successor) and the constant symbol 1, and write out three properties:

i. $\forall x \exists y\, [F(x,y) \wedge \forall z\, F(x,z) \rightarrow y = z]$
ii. $\neg \exists x F(x,1)$
iii. $\forall x\, [x \not\approx 1 \rightarrow \exists y\, (F(y,x) \wedge \forall z\, (F(z,x) \rightarrow y \approx x))]$.

Their formal system is the following:

1. Propositional variables $X, Y, \cdots$
2. Object variables $x, y, \cdots$
3. Relation symbols $F(\ ), G(\ ), \cdots$
4. Connectives: $\vee$ and $\neg$. ($X \rightarrow Y$ means $\neg X \vee Y$.)
5. Quantifiers: $\forall$ and $\exists$. (They use (x) for $\forall x$, and (Ex) for $\exists x$.)

6. Axioms:

$X \vee X \rightarrow X$
$X \rightarrow X \vee Y$
$X \vee Y \rightarrow Y \vee X$
$(X \rightarrow Y) \rightarrow [Z \vee X \rightarrow Z \vee Y]$
$\forall x\, F(x) \rightarrow F(x)$
$F(x) \rightarrow \exists x\, F(x)$.

7. Rules of Inference:

- substitution rule
- modus ponens
- $\dfrac{\varphi \rightarrow \psi(x)}{\varphi \rightarrow \forall x\, \psi(x)}$ (provided x is not free in φ)

- $\dfrac{\psi(x) \to \varphi}{\exists x\, \psi(x) \to \varphi}$

The simple form of the last two axioms and the last rule of inference was due to Bernays. Using this system they show elementary facts, such as every formula can be put in prenex form.

Then they turn to the questions that they have designated as important. Using a one-element universe they show the above axiom system is consistent (recall Frege's system was not consistent). Then they say (*Grundzüge*, p. 65):

> With this we have absolutely no guarantee that the introduction of assumptions, in symbolic form, to whose interpretation there is no objection, keeps the system of derivable formulas consistent. For example, there is the unanswered question of whether the addition of the axioms of mathematics to our calculus leads to the provability of every formula. The difficulty of this problem, whose solution has a central significance for mathematics, is in no way comparable to that of the problem just handled by us $\cdots$. To successfully mount an attack on this problem, D. Hilbert has developed a special theory.

Ackermann's proof that the above system is not complete in the stronger sense is given (*Grundzüge*, pp. 66-68), namely they show $\exists x\, F(x) \to \forall\, xF(x)$ and its negation are not derivable. Then the following is said (*Grundzüge*, p. 68):

> It is still an unsolved problem as to whether the axiom system is complete in the sense that all logical formulas which are valid in every domain can be derived. It can only be stated on empirical grounds that this axiom system has always been adequate in the applications. The independence of the axioms has not been investigated.

After some examples of using this system, they turn to the decision problem in §11 of Chapter III. We quote (*Grundzüge*, p. 72):

> According to the methods characterized by the last examples one can apply the first-order calculus in particular to the axiomatic treatment of theories $\cdots$.
>
> Once logical formalism is established one can expect that a systematic, so-to-say computational treatment of logical formulas is possible, which would somewhat correspond to the theory of equations in algebra.
>
> We met a well-developed <u>algebra of logic</u> in the propositional calculus. The most important problems solved there were the <u>universal</u> <u>validity</u> and <u>satisfiability</u> of a logical expression. Both problems are called the <u>decision problem</u> $\cdots$.
>
> (*Grundzüge*, p. 73)
> The two problems are dual to each other. If an expression is not universally valid then its negation is satisfiable, and conversely.
>
> The decision problem for first-order logic now presents itself $\cdots$. One can $\cdots$ restrict oneself to the case where the propositional variables do not appear $\cdots$.
>
> The decision problem is solved if one knows a process which, given a logical expression, permits the determination of its validity resp. satisfiability.
>
> The solution of the decision problem is of fundamental importance for the theory of all subjects whose theorems are capable of being logically derived from finitely many axioms $\cdots$.
>
> (*Grundzüge*, p. 74)
> We want to make it clear that for the solution of the decision problem a process would be given by which nonderivability can, in principle, be determined, even though the difficulties of the process would make practical use illusory $\cdots$.

(*Grundzüge*, p. 77)

⋯ the decision problem be designated as the main problem of mathematical logic.

⋯ in first-order logic the discovery of a general decision procedure is still a difficult unsolved problem ⋯.

In the last section of Chapter III they give Löwenheim's decision procedure for first-order statements involving only unary relation symbols, namely one shows that it suffices to examine all domains of size less than or equal to 2^k, where k is the number of unary predicate symbols in the statement. Consequently (p. 80)

... postulating the validity resp. satisfiability of a logical statement is equivalent to a statement about the size of the domain.

In the most general sense one can say the decision problem is solved if one has a procedure that determines for each logical expression for which domains it is valid resp. satisfiable.

One can also pose the simpler problem of when a given expression

is valid for all domains

and when not. This would suffice, with the help of the decision process, to answer whether a given statement of an axiomatically based subject was provable from the axioms.

Then they point out that one cannot hope for a process based on examining finite domains, as in the unary predicate calculus; but that Löwenheim's theorem gives a strong analog, namely a statement is valid in all domains iff it is valid in a countably infinite domain. Then they attempt to tie the decision problem to the completeness problem (*Grundzüge*, p. 80):

Examples of formulas which are valid in every domain are those derived from the predicate calculus. Since one suspects that this system gives all such [valid] formulas, one would move closer to the solution of the decision problem with a characterization of the formulas provable in the system.

A general solution of the decision problem, whether in the first or second formulation, has not appeared till now. Special cases of the decision problem ⋯ have been attacked and solved by P. Bernays, and M. Schoenfinkel, as well as W. Ackermann.

Finally they note that Löwenheim showed one could restrict ones attention to unary and binary predicates.[2]

Chapter IV starts out by trying to show that one needs to extend first-order logic to handle basic mathematical concepts. The extension takes place by allowing quantification of relation symbols. Then one can express complete induction by (*Grundzüge*, p. 83):

$$[P(1) \wedge \forall x \forall y \, (P(x) \wedge Seq(x,y) \rightarrow P(y))] \rightarrow \forall x \, P(x)$$

or, to be more explicit,

one can put the universal quantifier (P) in front of the formula.

Identity between x and y, $\equiv (x,y)$, is *defined* by $\forall F \, F(x) \leftrightarrow F(y)$.

Then they say (*Grundzüge*, p. 86):

The solution of this general decision problem (for the extended logic) would not only permit us to answer questions about the provability of simple geometric theorems, but it would also, at least in principle, make possible the decision about the provability, resp. nonprovability, of an arbitrary mathematical statement.

The first-order calculus was fine for a few special theories,

[2] One finds Schröder alluding to this fact in Vol. III of *Algebra der Logik*, but he does not try to justify his remarks.

> But as soon as one makes the foundations of theories, especially of mathematical theories, as the object of investigation $\cdots$ the extended calculus is indispensable.

The first important application of the extended calculus is to numbers. Individual numbers are realized as properties of predicates, e.g.,

$$0(F) =: \neg\exists x\, F(x)$$

$$1(F) =: \exists x\, [F(x) \wedge \forall y\, (F(y) \rightarrow \equiv (x, y))].$$

The condition for Φ being a [cardinal] number is

$$\forall F\, \forall G\, [(\Phi(F) \wedge \Phi(G) \rightarrow SC(F, G) \wedge (\Phi(F) \wedge SC(F, G) \rightarrow \Phi(G))]\,,$$

where $SC(F, G)$ says there is a 1-to-1 correspondence between the elements satisfying F and those satisfying G. Having set up the definitions and assuming an infinite domain, they say (footnoting Whitehead and Russell) that:

> It is also of particular interest that the number theoretic axioms become logically provable theorems.

They develop set theory in this extended calculus by saying that sets are the extension of unary predicates; thus two predicates F and G determine the same set iff $\forall x\, F(x) \leftrightarrow G(x)$ holds, which is abbreviated to $Aeq(F, G)$. Then properties of sets correspond to unary predicates P of unary predicates, where P is invariant under Aeq. Sets of ordered pairs correspond to binary predicates, etc. Under this translation a number becomes a set of sets of individuals from the domain, namely a number is the set of all sets equivalent to a given set. Their development of set theory ends with union, intersection, ordered sets and well-ordered sets defined, and they say that all the usual concepts of set theory can be expressed symbolically in their system.

Having shown the expressive power of the extended calculus of relations they turn to the problem of finding axioms and rules of inference. Obvious generalizations from the first-order logic lead to contradictions through the well-known paradoxes. Finally they outline Whitehead and Russell's ramified theory of types (with its questionable axiom of reducibility) which allows one to give axioms and rules of inference generalizing those of the first-order, avoiding the usual encoding of the paradoxes as contradictions, and being strong enough to carry out traditional mathematics. After mentioning that the decision problem also applies to the system of *Principia,* the book closes with Hilbert claiming to have a development of extended logic, which will soon appear, which avoids the difficulties of the axiom of reducibility.

Looking back over this book we see that its purpose is to present formal systems, and give examples. There are almost no real theorems of mathematical logic proved, or stated, after the development of the propositional calculus in Chapter I; the main exceptions being a proof of Löwenheim's decision process for the first-order unary relational logic, and the statement of the Löwenheim-Skolem theorem. First-order set theory is not even mentioned.

CHAPTER 11

A finitistic point of view: Herbrand

Jacques Herbrand (1908–1931)
1930 - Recherches sur la théorie de la démonstration

Herbrand said that his goal was to make the work of Löwenheim and Skolem rigorous [from the finitistic point of view]. In essence he introduced a proof system so that one had a notion of derivation $\vdash$ (essentially that of Hilbert and Ackermann), and he described the countermodel procedure, in his own language, and showed that a first-order statement was derivable iff the attempt to build a countermodel failed at some finite stage. Furthermore there was an effective procedure to go from the knowledge that the countermodel failed at the k^{th} stage to a derivation of the statement. Thus we see that Herbrand has come up with a version of the Löwenheim-Skolem theorem that does not mention infinite models.

CHAPTER 12

The Completeness and Incompleteness Theorems: Gödel

Kurt Gödel (1906–1978)
1930 - Die Vollständigkeit der Axiome des logischen Funktionenkalküls
1931 - Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I

Gödel's first paper proves the completeness of the axioms and rules of first-order logic, essentially as given in Hilbert and Ackermann. There has been much discussion as to why Skolem (or Herbrand) were so close, yet did not think of the question. Gödel makes use of a simple transformation which shows that it is enough to prove that a (derivation) consistent set of sentences has a model. Then he uses the Skolem normal forms of his sentences to aid in constructing the model.

The second paper was far more sophisticated. Recall that in 1930 it was well known that all traditional mathematical proofs could be expressed in powerful systems like Principia Mathematica and ZFC. The open question was whether or not they were powerful enough for all future mathematics as well, i.e., were they complete? Gödel, using only elementary number theory, showed that one could encode the workings of these powerful systems into formulas about numbers. Then he was able to construct true sentences (in these formalisms) about numbers which could not be proved using these systems.

Finally he added a brief remark to the effect that a sentence which expresses the consistency of such a system could not be proved in the system. This has widely been regarded as the end of Hilbert's program to prove the consistency of mathematics by finitary means.

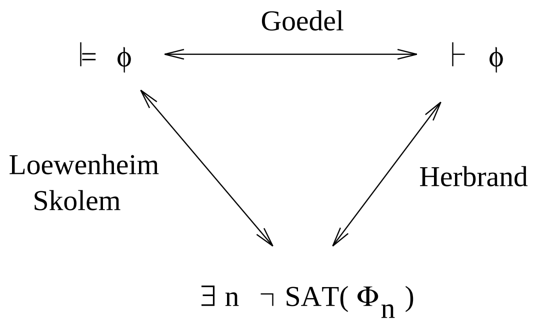

Figure 8: Connecting proof and truth

CHAPTER 13

The Consistency of Arithmetic: Gentzen

Gerhard Gentzen (1909–1945)
1934 - Untersuchungen über das logische Schliessen
1936 - Die Wiederspruchsfreiheit der reinen Zahlentheorie

Gentzen introduced the use of finite sequences of formulas as a basic object, called a sequent. He considered this formalization closer to the way we actually reason.

Then he turned to the question of the consistency of PA. One consequence of Gödel's work is the fact that if one can prove the consistency of PA, then the proof "is not expressible in PA". This is understood to imply that one cannot prove the consistency of PA by finitistic means, as had been hoped by Hilbert.[1] Gentzen did succeed in proving the consistency of PA, but by using a non finitistic framework, namely he used transfinite induction up to the ordinal ε_0, the limit of $\omega, \omega^\omega, \omega^{\omega^\omega}, \cdots$.

[1]In the foreword to Hilbert and Bernays two volumes on logic Hilbert says that there is a widespread and incorrect assumption that Gödel's results have proved his program of proving the consistency of mathematics by finitistic means to be hopeless.