UNIVERSITY OF LEEDS

This is a repository copy of *The Art of Measuring the Strength of Theories*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/203713/

Version: Accepted Version

Article:

Rathjen, M. orcid.org/0000-0003-1699-4778 (Cover date: August 2023) The Art of Measuring the Strength of Theories. Notices of the American Mathematical Society, 70 (7). pp. 1071-1079. ISSN 0002-9920

https://doi.org/10.1090/noti2733

© Copyright 2023 by the American Mathematical Society. This is an author produced version of an article published in Notices of the American Mathematical Society. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The Art of Measuring the Strength of Theories

Michael Rathjen *

May 30, 2023

1 What are Theories?

In mathematics one encounters theories of groups, fields, modules, vector spaces, Hilbert spaces, Banach algebras, probability spaces etc. Here *theory* means an assemblage of *axioms* that single out structures sharing certain features of interest to mathematicians in various contexts. There is usually no intention that these axioms ought to characterize a single structure - usually quite the opposite. Obviously, the word "axiom" is not meant to convey that one deals with indisputable self-evident truth as in the original meaning of that word since ancient times and until the 19th century. But by adopting a larger perspective, one can also ask what underpins the whole enterprise of mathematics. Surely, it draws on logic, but in addition one needs a plethora of mathematical objects either given to us from the start or constructed in some way. In other words, what are the axioms of mathematics (in the old sense)? Euclid intended to answer this question for geometry. Axiomatizations sufficient unto the task of undergirding the entire edifice of mathematics came rather late. Rigorous accounts of the laws of logic, especially the logic of the quantifiers, had to await the late 19th century. Frege achieved this for logic in his *Begriffsschrift* (concept script) from 1879 and then attempted an axiomatization of mathematics in *Grundgesetze der Arithmetik* (1893). These first steps were followed by the massive Principia Mathematica (1910–1913) of Whitehead and Russell, and, springing forth from Cantor's set theory, the axiomatization of set theory by Zermelo (1908), culminating in the axiomatic system of *Zermelo–Fraenkel set theory*, ZFC. In this article we shall mean by *theory* an axiomatic system in the sense of the foregoing sentence. The intent of such a theory is to provide an axiomatic foundation for the whole of mathematics or at least substantial chunks of it. To distinguish them from theories of groups, Hilbert spaces etc., one should perhaps call them *metamathematical theories*. But as this article will be concerned exclusively with the latter kind, I shall address them just as *theories*. In the wake of Hilbert's program and Gödel's incompleteness theorems, it became clear that theories can have very different *consistency strengths*, also called *proof strengths*. The purpose of this article is to describe and explain how the strength of such theories can be measured by transfinite ordinals, and how this ordinal-theoretic characterization can be milked to extract information about the provably computable functions of the theory and, moreover, yield unprovability and combinatorial independence results.

*The author is a professor of mathematics at the University of Leeds. His email address is m.rathjen@leeds.ac.uk.

2 The theory of natural numbers

The natural numbers equipped with the usual arithmetic functions arguably constitute the most important structure of mathematics.

The laws that govern this structure where explicated by Dedekind in his famous essay *Was sind und was sollen die Zahlen?*. They gave rise to a system of axioms, the Dedekind–Peano axioms, which is collectively known as *elementary number theory* or *first order arithmetic*, but nowadays mainly called *Peano arithmetic*, PA.

Definition 2.1 A theory designed with the intent of axiomatizing the structure

$$\mathfrak{N} = (\mathbb{N}; 0, 1, +, \cdot, \exp, <)$$

of the naturals is *Peano arithmetic*, PA. The language of PA has relation symbols $=, <$ for the equality and the less-than relation, respectively, the function symbols $+, \cdot, \exp$ (for addition, multiplication, exponentiation) and the constant symbols 0 and 1. The *Axioms of* PA comprise the usual equations and laws for addition, multiplication, exponentiation, and the less-than relation. In addition, PA has the *Induction Scheme*

$$(\mathsf{IND}) \qquad \varphi(0) \land \forall x[\varphi(x) \to \varphi(x+1)] \to \forall x \varphi(x)$$

for all formulae φ of the language of PA.
By the usual laws of we mean the following: $0 \neq 1$, $\forall x\, x+1 \neq 0$, $\forall x \forall y\, [x+1 = y+1 \to x = y]$, $\forall x\, x+0 = x$, $\forall x \forall y\, [x+(y+1) = (x+y)+1]$, $\forall x\, x \cdot 0 = 0$, $\forall x \forall y\, [x \cdot (y+1) = x \cdot y + x]$, $\forall x \forall y\, [x < y+1 \leftrightarrow (x < y \lor x = y)]$, $\forall x\, x^0 = 1$, and $\forall x \forall y\, [x^{y+1} = x^y \cdot x]$, writing x^y for $\exp(x, y)$.

As the axioms of PA enable one to do coding, its language is rather expressive. Several famous conjectures such as the Twin Prime Conjecture and those of Goldbach and Riemann[1] can be expressed in it. Moreover, many theorems about numbers such as the Prime Number Theorem (and, conjecturally, Fermat's Last Theorem, i.e., Wiles' Theorem) can be proved in PA.

Gerhard Gentzen gave two different consistency proofs for PA in 1936 and 1938, that is, that no contradiction (inconsistency), such as $0 = 1$, can be inferred from the axioms of PA. He had developed a theory of proofs as suggested by Hilbert in 1917 [Hil18].[2]

> To conquer this field [concerning the foundations of mathematics] we must turn the concept of a specifically mathematical proof itself into an object of investigation, just as the astronomer considers the movement of his position, the physicist studies the theory of his apparatus, and the philosopher criticizes reason itself.

Gentzen [Gen38] presented an ingenious procedure $\mathcal{R}$ whereby any alleged proof P of $0 = 1$ in PA gets reduced to another proof $\mathcal{R}(P)$ of $0 = 1$. The proof $\mathcal{R}(P)$ may not be less complex than P in any ordinary sense, where one just counts the number of symbols or lines, however, Gentzen assigned transfinite ordinals to proofs to the effect that $\mathcal{R}(P)$ receives a smaller ordinal than P. As a result, an inconsistency will lead to an infinite descending sequence of ordinals. However, as an infinite descent is impossible in the ordinals, no contradiction can be deduced from PA. It is of note that all the steps in his argument concerning the manipulation of proofs are very concrete and elementary. For instance, $\mathcal{R}$ and the ordinal assignment are given by basic functions.[3] It is only the invocation of the principle of *no-infinite-descent* that transcends elementary means.

Gentzen's result was also optimal in that he used ordinals below the first ordinal $\rho > 0$ such that $\omega^\rho = \rho$ and showed that no smaller ordinal segment sufficed. What is actually meant by *ordinal* will be explained next.

3 Ordinals, Wellorderings, Ordinal Representation Systems

A set A is *transitive* if $y \in A$ and $x \in y$ entails $x \in A$. In set theory, ordinals are introduced as rather abstract, mostly transfinite objects, namely transitive sets A whose elements happen to be linearly ordered by the elementhood relation $\in$, that is,

[1] This was first observed by Turing.

[2] Gentzen published a beautiful proof system in 1935, known as the *sequent calculus*, in which any proof can be transformed into one without detours. The latter result is called *Gentzen's Hauptsatz*.

[3] Those basic functions can be taken to be the *primitive recursive functions*. *Primitive recursion* is given by the equations $f(\vec{y}, 0) = g(\vec{y})$ and $f(\vec{y}, x+1) = h(\vec{y}, f(\vec{y}, x), x)$ where $\vec{y} = y_1, \ldots, y_n$ and g, h are previously defined primitive recursive functions. The collection of primitive recursive functions is obtained from the constant 0 function, $S(x) = x + 1$ and the projection functions $P^i_n(x_1, \ldots, x_n) = x_i$ by closure under composition and the above recursion schema.

for all $x, y \in A$, $x \in y$ or $y \in x$ or $x = y$. However, many interesting ordinals α have concrete representations as term systems and can be defined as orderings of subsets of $\mathbb{N}$.

Definition 3.1 A set A equipped with a total ordering $\prec$ (i.e. $\prec$ is transitive, irreflexive, and $\forall x, y \in A\,[x \prec y \vee x = y \vee y \prec x]$) is a *wellordering* if every non-empty subset X of A contains a $\prec$-least element, i.e. $(\exists u \in X)(\forall y \in X)[u \prec y \vee u = y]$.

An *ordinal* is a transitive set wellordered by the elementhood relation $\in$.

Fact 3.2 *Every wellordering $(A, \prec)$ is order isomorphic to an ordinal $(\alpha, \in)$.*

The crucial property of a wellordering $\prec$ can be expressed equivalently in a positive manner via the pertinent *transfinite induction principle*:

($\star$) If $\forall u \prec x\, P(u)$ implies $P(x)$ for every $x \in A$, then $\forall y \in A\, P(y)$,

where P is an arbitrary property. Note that ($\star$) is similar to the strong induction principle for the standard ordering of the naturals.

Ordinals are traditionally denoted by lowercase Greek letters $\alpha, \beta, \gamma, \delta, \ldots$ and the relation $\in$ on ordinals is notated simply by $<$. The operations of addition, multiplication, and exponentiation can be defined on all ordinals, however, addition and multiplication are in general not commutative.

The ordinals that Gentzen used in his consistency proof of PA can be nicely explained in terms of a normal form theorem due to Cantor.

Theorem 3.3 (Cantor, 1897) *For every ordinal $\beta > 0$ there exist unique ordinals $\beta_0 > \beta_1 > \cdots > \beta_n$ and nonzero natural numbers $k_0 \ldots, k_n$ such that*

$$\beta = \omega^{\beta_0} \cdot k_0 + \ldots + \omega^{\beta_n} \cdot k_n. \quad (1)$$

The representation of β in (1) is called the *Cantor normal form.* We shall write

$$\beta =_{CNF} \omega^{\beta_1} \cdot k_0 + \cdots \omega^{\beta_n} \cdot k_n$$

to convey that the right hand side exhibits β's normal form.

ε_0 denotes the least ordinal $\alpha > 0$ such that $(\forall \beta < \alpha)\, \omega^\beta < \alpha$. ε_0 can also be described as the least ordinal α such that $\omega^\alpha = \alpha$.

Ordinals $\beta < \varepsilon_0$ have a Cantor normal form with exponents $\beta_i < \beta$ and these exponents have Cantor normal forms with yet again smaller exponents. As this process must terminate, ordinals $< \varepsilon_0$ can be coded by natural numbers. Indeed, such ordinals need not be conceived as denizens of a lofty set-theoretic realm. They can be identified with syntactic expressions of the form (1) defined simultaneously with an ordering $\prec$ between them. Assume that we have already recognized $\beta := \omega^{\beta_0} \cdot k_0 + \ldots + \omega^{\beta_n} \cdot k_n$ and $\delta := \omega^{\delta_0} \cdot l_0 + \ldots + \omega^{\delta_m} \cdot l_m$ as legitimately built ordinals terms. This entails that $\beta_0 \succ \ldots \succ \beta_n$ and $\delta_0 \succ \ldots \succ \delta_m$. Moreover, the naturals k_i and l_j are all greater than 0. To determine the ordering relation $\prec$ between β and δ one proceeds as follows.

1. If there exists an $i \leq \max(n, m)$ such that $\omega^{\beta_i} \cdot k_i \neq \omega^{\delta_i} \cdot l_i$, let i_0 be the least such. In this case the order is determined according to the following recipe.
   (a) If $\beta_{i_0} \succ \delta_{i_0}$, then $\beta \succ \delta$.
   (b) If $\beta_{i_0} \prec \delta_{i_0}$, then $\beta \prec \delta$.
   (c) If $\beta_{i_0} = \delta_{i_0}$ and $k_{i_0} > l_{i_0}$, then $\beta \succ \delta$.
   (d) If $\beta_{i_0} = \delta_{i_0}$ and $k_{i_0} < l_{i_0}$, then $\beta \prec \delta$.
2. If $\beta_i = \delta_i$ and $k_i = l_i$ hold for all $i \leq m$ and $n > m$, then $\beta \succ \delta$.
3. If $\beta_i = \delta_i$ and $k_i = l_i$ hold for all $i \leq n$ and $n < m$, then $\beta \prec \delta$.
4. If $\beta_i = \delta_i$ and $k_i = l_i$ hold for all $i \leq n$ and $n = m$, then $\beta = \delta$.

The foregoing inductively defined set of ordinal terms, $\mathcal{T}_{\varepsilon_0}$, equipped with their ordering $\prec$ can easily be implemented in any programming language and of course formalized in PA. In point of fact, its computational complexity is very low.[4]

[4]It is often stressed that ordinal representation systems are computable structures, which is true and allows for the treatment of their order-theoretic and algebraic aspects in very weak systems of arithmetic, but it is only one of their distin-

Proof strength, however, is required for proving transfinite induction over $\prec$. Let $\omega_0 := 0$ and $\omega_{n+1} := \omega^{\omega_n}\cdot 1$. The supremum of the ω_n is ε_0. Gentzen (1943) showed that for any n, transfinite induction over the initial segment determined by ω_n is provable in PA and also in 1938 that PA does not prove transfinite induction over all ordinal terms (whose set represents the ordinal ε_0). So the idea was borne that ε_0 characterizes the proof-theoretic strength of PA. Furthermore, this was the second time an independence result for PA had been found; indeed, one of a completely different nature than Gödel's.

4 Combinatorial independence

The unprovability of the principle of transfinite induction up to ε_0 from PA is an important result. However, it has a rather logical or set-theoretic flavor. So one might ask whether a purely number-theoretic statement could be distilled from it, that is, one using solely concepts familiar to any number theorist. Indeed, such a statement was unearthed by Goodstein in 1944. He realized that there exists a similarity between the existence of Cantor normal forms and the fact that any positive integer m has a unique presentation with regards to a base $b \geq 2$, that is, m can be uniquely expressed in the form

$$m \quad = \quad b^{n_1}\cdot k_1 + \cdots + b^{n_r}\cdot k_r, \qquad (2)$$

where $m > n_1 > \cdots > n_r \geq 0$ and $0 < k_1, \cdots, k_r < b$. As each $n_i > 0$ is itself of this form we can repeat this procedure, arriving at what is called the **complete b-representation** of m. In this way we get a unique representation of m over the alphabet $0, 1, \cdots, b, +, \cdot$. For example, with $b = 3$ one has $7\,625\,597\,485\,157 = 3^{27}\cdot 1 + 3^4\cdot 2 + 3^1\cdot 2 + 3^0\cdot 2 = 3^{3^3} + 3^{3+1}\cdot 2 + 3^1\cdot 2 + 2$.

Goodstein [Goo44], then, proceeded to define wondrous sequences of naturals, nowadays of course called *Goodstein sequences.*

guishing features. Overstating the computability aspect tends to give the impression that their study is part of the venerable research area of computable orderings. In actual fact, the two subjects have very little in common.

Definition 4.1 For naturals $m > 0$ and $c \geq b \geq 2$ let $\mathrm{S}^b_c(m)$ be the integer resulting from m by replacing the base b in the complete b-representation of m everywhere by c. For example $\mathrm{S}^3_4(34) = 265$, since $34 = 3^3 + 3\cdot 2 + 1$ and $4^4 + 4\cdot 2 + 1 = 265$.

Given any natural number m and non-decreasing function $f : \mathbb{N} \to \mathbb{N}$ with $f(0) \geq 2$ define

$$m^f_0 = m, \quad \cdots \quad, m^f_{i+1} = \mathrm{S}^{f(i)}_{f(i+1)}(m^f_i) \dot{-} 1$$

where $k \dot{-} 1$ is the predecessor of k if $k > 0$, and $k \dot{-} 1 = 0$ if $k = 0$.

$(m^f_i)_{i\in\mathbb{N}}$ is said to be a **Goodstein sequence**. Note that $(m^f_i)_{i\in\mathbb{N}}$ is uniquely determined by f and its starting point $m = m^f_0$.

Goodstein sequences, even for a modest starting point such as 4, quickly climb up to gigantic numbers, but then miraculously begin their descent after reaching an outlandishly large stage in the sequence.

Theorem 4.2 (Goodstein 1944) *Every Goodstein sequence terminates, i.e., there exists k such that $m^f_i = 0$ for all $i \geq k$.*

Proof: This is seen by assigning ordinals to the numbers m^f_i, effected by replacing the base $f(i)$ in the complete $f(i)$-representation of m^f_i by ω. One then sees that the resulting ordinal sequence decreases until it hits 0. Hence $m^f_i = 0$ for a sufficiently large i as there are no infinite descending sequences of ordinals. $\Box$

Conversely, the principle of termination of Goodstein sequences entails constructively that an infinite decent among the ordinal (representations) below ε_0 is impossible. This insight gives rise to an independence result. The language of PA, however, is not sufficiently rich to talk about arbitrary sequences of numbers, but it is capacious enough for formalizing the notion of Turing computable sequences or primitive recursive sequences (as defined in footnote (3)). Thus, Goodstein's results from 1944 yield an equivalence provable in PA.

Corollary 4.3 *Over* PA *the following are equivalent:*

(i) Every primitive recursive Goodstein sequence terminates.

(ii) There are no infinitely descending primitive recursive sequences of ordinal representations $\varepsilon_0 > \alpha_0 > \alpha_1 > \alpha_2 > \cdots$.

In light of Gentzen's consistency proof for PA, which uses the primitive recursive reduction operator $\mathcal{R}$ and a primitive recursive assignment of ordinals to proofs, the following number-theoretic independence result emerges from Gentzen's and Goodstein's insights.

Theorem 4.4 (Gentzen and Goodstein) *Termination of primitive recursive Goodstein sequences is not provable in* PA

The case when f is just a shift function has received special attention. Given any m we define $m_0 = m$ and $m_{i+1} := \mathrm{S}^{i+2}_{i+3}(m_i) \dot{-} 1$ and call $(m_i)_{i\in\mathbb{N}}$ a **special Goodstein sequence**. Thus $(m_i)_{i\in\mathbb{N}} = (m_i^{\mathrm{id}_2})_{i\in\mathbb{N}}$, where $\mathrm{id}_2(x) = x + 2$. Special Goodstein sequences can differ only with respect to their starting points.

As shown much later by Kirby and Paris in 1982 [KP82], already the termination of special Goodstein sequences is unprovable in PA. They used model-theoretic tools. Another famous Ramsey-type independence result is the Paris-Harrington theorem from 1977.

4.1 The general form of ordinal analysis of a theory

Gentzen's analysis of PA gave rise to the idea that ordinals can measure the strength of theories. Given a theory T, in which at least some basic parts of mathematics can be developed, and an ordinal representation system (ORS) one would like to associate an ordinal in the ORS to T. First one should assume that T contains means to develop some basic arithmetic. In practice this means that T contains the system of primitive recursive arithmetic, PRA. The latter theory has often been equated with Hilbert's finitism. Let F be such a basic theory.

Definition 4.5 Suppose that there is an ordinal ρ in the ORS such that F together with the statement

($\star$) There are no infinitely descending primitive recursive sequences of ordinals below ρ.

proves the consistency of T, and, moreover, ρ is the least such ordinal. Then ρ is said to be the *proof-theoretic ordinal of T*.

We will treat the above as a working definition.[5]

5 The birth of second order proof theory

Although a fair chunk of mathematics can be developed in PA, its expressiveness is rather restricted as it doesn't accomodate a reasonable theory of the reals. Hilbert's work on axiomatic geometry marked the beginning of his livelong interest in the axiomatic method. For geometry, he solved the problem of consistency by furnishing arithmetical-analytical interpretations of the axioms, thereby reducing the question of consistency to the consistency of the theory of the real numbers. The consistency of the latter system of axioms is therefore the ultimate problem for the foundations of mathematics. It became the second problem on Hilbert's famous list of problems from 1900.

The language of second-order arithmetic, $\mathcal{L}_2$, is an augmentation of that of PA in that it has variables $X, Y, Z, \ldots$ ranging over sets of natural numbers which can be quantified over and identified with the real numbers. The full theory Z_2 has a comprehension axiom to the effect that

$$\{n \in \mathbb{N} \mid \varphi(n)\} \tag{3}$$

is a set for any formula $\varphi(n)$ of Z_2. (3) is a relative of the set-theoretic comprehension axiom that is responsible for the paradoxes of naïve set theory (as it

[5] A much more thorough discussion of this notion can be found in [Rat99, Section 2]. In practice, this also entails that $\mathsf{F} + (\star)$ proves all theorems of T of the complexity of the twin prime conjecture.

came to be called) found by Cantor, Russell and others. Z_2 is a rather strong theory. It can accomodate a great deal of ordinary mathematics, especially a theory of the reals. Hermann Weyl was very disturbed by the paradoxes. The expression *foundational crisis* was coined by him [Wey21]. As a consequence, he would only countenance *arithmetical comprehension*, that is, only instances of (3), where the formula $\varphi(n)$ doesn't contain quantifiers over sets of numbers. On his view (and that of other predicativists, including Poincaré), if $\varphi(n)$ contains such quantifiers, then these quantifiers already range over the set asserted to exist by (3), thus constituting a circularity. An instance of the latter is said to be an *impredicative* comprehension axiom.

5.1 Takeuti's fundamental conjecture

Proving the consistency of second order arithmetic Z_2, that is, solving Hilbert's second problem, became the holy grail problem of proof theory. In the late 1940's, Gaisi Takeuti wanted to extend Gentzen's methodology to Z_2. He formulated a sequent calculus GLC (for generalized logical calculus) which encompassed Z_2. The idea was to prove a Hauptsatz for GLC à la Gentzen from which the consistency of analysis, i.e., Z_2, would follow. This came to be known as Takeuti's *fundamental conjecture*. Looking back at that time, Takeuti wrote:

> Having proposed the fundamental conjecture, I concentrated on its proof and spent several years in an anguished struggle trying to resolve the problem day and night. [Tak03, p. 133]

It was only much later, in the 1960s, that Takeuti appreciated that he had made substantial progress. Rather than trying to prove the whole conjecture, he finally concentrated on partial results at the suggestion of Maehara. In his 1967 paper [Tak67], he gave an ordinal analysis of the subsystem of Z_2 with Π^1_1-comprehension (Π^1_1-CA) whose main axiom scheme asserts that

$$\{n \in \mathbb{N} \mid \forall X\, \theta(n, X)\}$$

is a set whenever the formula $\theta(n, X)$ contains no further second order quantifiers; so only quantifiers over natural numbers are allowed therein. For this Takeuti returned to Gentzen's method of assigning ordinals (ordinal diagrams, to be precise) to purported derivations of the empty sequent (inconsistency). A sufficiently strong ordinal representation system that captures its strength will be described later.

The next section will be concerned with investigations as to which set existence axioms are necessary for developing various parts of mathematics, and what ordinal strengths these chunks have. From the latter vantage point, the fragment of Z_2 based on Π^1_1-CA turns out to be rather strong in that most of the theorems of ordinary mathematics can be proved in it.

For the sake of comparison, however, it's perhaps worth pointing out that Π^1_1-CA is very weak when compared to Z_2, while Z_2 is a hugely weaker theory than the axiomatic set theory ZFC.

6 Theories for the development of mathematics

It was already mentioned that Weyl in 1918 took the radical consequence of ditching all the mathematics that relied on impredicative set existence axioms. He accepted the infinite set $\mathbb{N}$ as a basis but all further sets had to be obtained by arithmetical comprehension from previously introduced sets of numbers. The resulting theory from [Wey18] is a conservative extension of PA, that is, it proves the same theorems about numbers as PA. Amazingly, contrary to first expectation, he could salvage or rather resurrect a great deal of analysis in his theory. Weyl's book became a blueprint for further investigations. A long list of mathematical logicians (e.g., Hilbert, Bernays, Lorenzen, Takeuti, Feferman, Friedman, Simpson to name a few) showed that large swathes of ordinary mathematics can be undergirded by theories of fairly modest consistency strength. To some extent, this confirms what Hilbert surmised in his conservativity program, namely that number-theoretic results (i.e., those expressible in the language of number theory) proved in abstract, nonconstructive mathematics can often be proved by much more elementary means. To obtain such results, logicians have developed elaborate theories for the formalization of mathematics, and shown, by a plethora of elaborate techniques from mathematical logic, that they are conservative over various elementary theories.

It is known that virtually all of ordinary mathematics can be formalized in Zermelo-Fraenkel set theory with the axiom of choice, ZFC. Hilbert and Bernays [HB38] verified that large swathes of mathematics can be formalized in second order arithmetic. Owing to these observations, proof theory initially focussed on subsystems of second order arithmetic. Further scrutiny revealed that small fragments are sufficient. Continuing in the wake of Hilbert and Bernays, a research program, dubbed *Reverse Mathematics*, was founded by H. Friedman [Fri75] some fifty years ago and then extensively developed by S. Simpson (see [Sim09]). The idea is to ask whether, given a theorem, one can prove its equivalence to some axiomatic system, with the aim of determining what proof-theoretical resources are necessary for the theorems of mathematics. More precisely, the objective of reverse mathematics is to investigate the role of set existence axioms in ordinary mathematics. The main question can be stated as follows:

> *Given a specific theorem τ of ordinary mathematics, which set existence axioms are needed in order to prove τ?*

Central to the above is the reference to what is called 'ordinary mathematics'. This concept, of course, doesn't have a precise definition. Roughly speaking, by ordinary mathematics we mean main-stream, non-set-theoretic mathematics, i.e., the core areas of mathematics which make no essential use of the concepts and methods of set theory and do not essentially depend on the theory of uncountable cardinal numbers.

6.1 Subsystems of second order arithmetic.

The framework chosen for studying set existence in reverse mathematics, though, is second order arithmetic rather than set theory. Second order arithmetic, Z_2, is a two-sorted formal system with one sort of variables $x, y, z, \ldots$ ranging over natural numbers and the other sort $X, Y, Z, \ldots$ ranging over sets of natural numbers. The language $\mathcal{L}_2$ of second-order arithmetic also contains the symbols of PA, and in addition has a binary relation symbol $\in$ for elementhood. Formulae are built from the prime formulae $s = t$, $s < t$, and $s \in X$ (where s, t are numerical terms, i.e. terms of PA) by closing off under the connectives $\wedge, \vee, \rightarrow, \neg$, numerical quantifiers $\forall x, \exists x$, and set quantifiers $\forall X, \exists X$.

The basic arithmetical axioms in all theories of second-order arithmetic are the defining axioms for $0, 1, +, \cdot, \exp, <$ (as for PA) and the *induction axiom*

$$\forall X(0 \in X \wedge \forall x(x \in X \rightarrow x+1 \in X) \rightarrow \forall x(x \in X)).$$

We consider the axiom schema of *$\mathcal{C}$-comprehension* for formula classes $\mathcal{C}$ which is given by

$$\mathcal{C} - \mathsf{CA} \qquad \exists X \forall u(u \in X \leftrightarrow F(u))$$

for all formulae $F \in \mathcal{C}$ in which X does not occur. Natural formula classes are the *arithmetical formulae*, consisting of all formulae without second order quantifiers $\forall X$ and $\exists X$, and the Π^1_n-*formulae*, where a Π^1_n-formula is a formula of the form $\forall X_1 \ldots QX_n\, A(X_1, \ldots, X_n)$ with $\forall X_1 \ldots QX_n$ being a string of n alternating set quantifiers, commencing with a universal one, followed by an arithmetical formula $A(X_1, \ldots, X_n)$.

For each axiom scheme Ax we denote by $(\mathsf{Ax})_0$ the theory consisting of the basic arithmetical axioms plus the scheme Ax. By contrast, (Ax) stands for the theory $(\mathsf{Ax})_0$ augmented by the scheme of induction for all $\mathcal{L}_2$-formulae.

An example for these notations is the theory $(\mathsf{\Pi}^1_1\text{-}\mathsf{CA})_0$ which has the comprehension schema for $\mathsf{\Pi}^1_1$-formulae.

For many mathematical theorems τ, there is a weakest natural subsystem $S(\tau)$ of Z_2 such that $S(\tau)$ proves τ. Very often, if a theorem of ordinary mathematics is proved from the weakest possible set existence axioms, the statement of that theorem will turn out to reversible in that it implies those axioms over a still weaker base theory, giving rise to the name

Reverse Mathematics for this theme. Moreover, it has turned out that $S(\tau)$ often belongs to a small list of specific subsystems of Z_2 dubbed RCA_0, WKL_0, ACA_0, ACA_0^+, ATR_0 and $(\Pi^1_1\text{−}\mathsf{CA})_0$, respectively. The systems are enumerated in increasing strength. The main set existence axioms of RCA_0, WKL_0, ACA_0, ATR_0, and $(\Pi^1_1\text{−}\mathsf{CA})_0$ are comprehension for computable sets, König's lemma for binary branching trees, arithmetical comprehension, existence of ω-fold Turing jumps, arithmetical transfinite recursion, and Π^1_1-comprehension, respectively. For exact definitions of all these systems and their role in reverse mathematics see [Sim09].

The below list is just meant to give an idea of what kind of mathematics can be developed in the various theories of RM and were they can be located on the theory scale of RM.

RCA_0: "Every countable field has an algebraic closure"; "Every countable ordered field has a real closure".

WKL_0: "Cauchy-Peano existence theorem for solutions of ordinary differential equations"; "Hahn-Banch theorem for separable Banach spaces".

ACA_0: "Bolzano-Weierstrass theorem";
"Every countable commutative ring with a unit has a maximal ideal".

ACA_0^+: Proves the "Auslander/Ellis Theorem" of topological dynamics.

ATR_0: "Every countable reduced abelian p-group has an Ulm resolution".

$(\Pi^1_1\text{−}\mathsf{CA})_0$: "Every uncountable closed set of real numbers is the union of a perfect set and a countable set"; "Every countable abelian group is a direct sum of a divisible group and a reduced group".

The proof-theoretic strength of both, RCA_0 and WKL_0 is the same and considerably weaker than that of PA while ACA_0 has the same strength as PA. In terms of proof-theoretic ordinals, one has $|\mathsf{RCA}_0| = |\mathsf{WKL}_0| = \omega^\omega$ and $|\mathsf{ACA}_0| = \varepsilon_0$. The proof-theoretic ordinals of ACA_0^+, ATR_0, and $(\Pi^1_1\text{−}\mathsf{CA})_0$, however, elude expression in the ORS introduced so far.

7 Ordinal representation systems for the 1960s

In the case of PA, Gentzen could rely on Cantor's normal form for a supply of ordinal representations. For stronger theories, though, segments larger than ε_0 have to be employed. Ordinal representation systems utilized by proof theorists in the 1960s arose in a purely set-theoretic context. This subsection will present some of the underlying ideas as progress in ordinal-theoretic proof theory also hinges on the development of sufficiently strong and transparent ordinal representation systems.

In 1904, Hardy wanted to "construct" a subset of $\mathbb{R}$ of size $\aleph_1$. His method was to represent countable ordinals via increasing sequence of natural numbers and then to correlate a decimal expansion with each such sequence. Hardy used two processes on sequences: (i) Removing the first element to represent the successor; (ii) Diagonalizing at limits.

Hardy's two operations give explicit representations for all ordinals $< \omega^2$. Veblen [Veb08], then, extended the initial segment of the countable for which fundamental sequences can be given effectively. The new tools he devised were the operations of *derivation* and *transfinite iteration* applied to *continuous increasing functions* on ordinals.

Definition 7.1 Let ON be the class of ordinals. A (class) function $f : ON \to ON$ is said to be *increasing* if $\alpha < \beta$ implies $f(\alpha) < f(\beta)$ and *continuous* (in the order topology on ON) if

$$f(\lim_{\xi<\lambda} \alpha_\xi) = \lim_{\xi<\lambda} f(\alpha_\xi)$$

holds for every limit ordinal λ and increasing sequence $(\alpha_\xi)_{\xi<\lambda}$. f is called *normal* if it is increasing and continuous.

The function $\beta \mapsto \omega+\beta$ is normal while $\beta \mapsto \beta+\omega$ is not continuous at ω since $\lim_{\xi<\omega}(\xi+\omega) = \omega$ but $(\lim_{\xi<\omega} \xi) + \omega = \omega + \omega$.

Definition 7.2 The *derivative* f' of a function $f : ON \to ON$ is the function which enumerates in increasing order the solutions of the equation $f(\alpha) = \alpha$, also called the *fixed points* of f.

If f is a normal function, $\{\alpha : f(\alpha) = \alpha\}$ is a proper class and f' will be a normal function, too.

Definition 7.3 *Now, given a normal function $f : ON \to ON$, define a hierarchy of normal functions as follows:*

$$f_0 = f \quad \textit{and} \quad f_{\alpha+1} = f'_\alpha \quad \textit{and for limits } \lambda\textit{:}$$
$$f_\lambda(\xi) \quad = \quad \xi^{th} \textit{ element of } \bigcap_{\alpha<\lambda} (\textit{Range of } f_\alpha)\,.$$

In this way, from the normal function f we get a two-place function, $\varphi_f(\alpha, \beta) := f_\alpha(\beta)$. Veblen then discusses the hierarchy $\varphi_\alpha := \varphi_f$, where $f(\beta) = \omega^\beta$.

Theorem 7.4 (Veblen Normal Form) *For every ordinal $\alpha > 0$ there exist uniquely determined ordinals $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_n$ such that:*

1. $\alpha = \varphi_{\xi_1}(\eta_1) + \ldots + \varphi_{\xi_n}(\eta_n)$
2. $\varphi_{\xi_1}(\eta_1) \geq \ldots \geq \varphi_{\xi_n}(\eta_n)$
3. $\eta_i < \varphi_{\xi_i}(\eta_i)$ *for* $i = 1, \dots, n$.

The least ordinal $\rho > 0$, such that $\varphi_\xi(\eta) < \rho$ whenever $\xi, \eta < \rho$, is traditionally called Γ_0. As the ordering of representations in Veblen normal form can be determined by a recursive procedure, similarly as for the Cantor normal, one arrives at an ordinal representation system for Γ_0. With its help, the proof-theoretic ordinals of some further systems of RM can be exhibited:
$|\mathsf{ACA}_0^+| = \varphi_2(0)$ and $|\mathsf{ATR}_0| = \Gamma_0$.[6] The one for $(\Pi^1_1\text{−}\mathsf{CA})_0$, however, still remains elusive.

8 Collapsing functions beyond Veblen

Veblen extended his approach, first to functions having a finite number of ordinal arguments, but then also to a transfinite number of arguments, with the proviso that in, for example $\Phi_f(\alpha_0, \alpha_1, \dots, \alpha_\eta)$, only a finite number of the arguments α_ν may be non-zero. Finally, Veblen singled out the ordinal $E(0)$, where $E(0)$ is the least ordinal $\delta > 0$ which cannot be named in terms of functions $\Phi_\ell(\alpha_0, \alpha_1, \dots, \alpha_\eta)$ with $\eta < \delta$, and each $\alpha_\gamma < \delta$.

Though the "great Veblen number" (as $E(0)$ is sometimes called) is quite an impressive ordinal it does not furnish an ordinal representation sufficient for the task of analyzing a theory as strong as Π^1_1 comprehension. Of course, it is possible to go beyond $E(0)$ and initiate a new hierarchy based on the function $\xi \mapsto E(\xi)$ or even consider hierarchies utilizing finite type functionals over the ordinals. Still all these further steps amount to rather mundane progress over Veblen's methods. In 1950 Bachmann [Bac50] presented a new kind of operation on ordinals which dwarfs all hierarchies obtained by iterating Veblen's methods. Bachmann built on Veblen's work, but his novel idea was the systematic use of a regular *uncountable cardinal* to keep track of the functions defined by diagonalization.

Let $\aleph_1$ be the first uncountable ordinal. Bachmann defines a set of ordinals $\mathfrak{B}$ closed under successor such that with each limit $\lambda \in \mathfrak{B}$ is associated an increasing sequence $\langle \lambda[\xi] : \xi < \tau_\lambda \rangle$ of ordinals $\lambda[\xi] \in \mathfrak{B}$ of length $\tau_\lambda \leq \aleph_1$ and $\lim_{\xi<\tau_\lambda} \lambda[\xi] = \lambda$. A hierarchy of functions $(\overset{\mathfrak{B}}{\varphi}_\alpha)_{\alpha\in\mathfrak{B}}$ is then obtained as follows:

$$\overset{\mathfrak{B}}{\varphi}_0(\beta) = \omega^\beta, \qquad \overset{\mathfrak{B}}{\varphi}_{\alpha+1} = \left(\overset{\mathfrak{B}}{\varphi}_\alpha\right)',$$
$$\overset{\mathfrak{B}}{\varphi}_\lambda = \mathrm{En}\big(\bigcap_{\xi<\tau_\lambda} (\text{Range of } \overset{\mathfrak{B}}{\varphi}_{\lambda[\xi]})\big) \quad \text{if } \tau_\lambda < \aleph_1,$$
$$\overset{\mathfrak{B}}{\varphi}_\lambda = \mathrm{En}\big(\{\beta < \aleph_1 : \overset{\mathfrak{B}}{\varphi}_{\lambda[\beta]}(0) = \beta\}\big) \text{ if } \tau_\lambda = \aleph_1,$$

where $f = \mathrm{En}(X)$ means that f enumerates the ordinals of X.

Modern approaches are much simpler and more transparent. We will briefly look at this.

8.0.1 The Bachmann-Howard ordinal

Definition 8.1 Let Ω_1 be a sufficiently 'big' ordinal. We define the sets $B_{\Omega_1}(\alpha)$ and ordinals $\psi_{\Omega_1}(\alpha)$ by

[6] The ordinal $\varphi_2(0)$ is also of interest in connection with the the smallest class of number-theoretic functions $\mathbb{N} \to \mathbb{N}$ that contains the constant functions, the identity function, and is closed under addition, multiplication and exponentiation, i.e., if f, g belong to it then so are $f + g$, $f \cdot g$, and f^g. Eventual domination yields a linear ordering on this class. Ehrenfeucht (1973) showed that it is a wellordering. Its order-type is at least ε_0. Levitz (1978) showed that it is no bigger than $\varphi_2(0)$. To this day, the exact order type is not known.

transfinite recursion on α as follows

$$B_{\Omega_1}(\alpha) \quad = \quad \begin{cases} \textit{closure of } \{0, \Omega\} \textit{ under:} \\ +,\ \xi \mapsto \omega^\xi \\ (\xi \longmapsto \psi_{\Omega_1}(\xi))_{\xi<\alpha} \end{cases} \tag{4}$$

$$\psi_{\Omega_1}(\alpha) \quad = \quad \min\{\rho < \Omega_1 :\ \rho \notin B_{\Omega_1}(\alpha)\}. \tag{5}$$

Now, the foregoing definition is vague in some important respect. What does it mean for Ω_1 to be sufficiently big? This could be defined implicitly by requiring that $\psi_{\Omega_1}(\alpha)$ is defined for all α (which is implicitly assumed in (5)), meaning that there always exist an ordinal $\rho < \Omega_1$ with $\rho \notin B_{\Omega_1}(\alpha)$. One can see, via a simple cardinality argument, that equating Ω_1 with the first uncountable ordinal, that is, the cardinal $\aleph_1$, will work. But this is surely overkill; much smaller countable ordinals can be substituted for Ω_1. The smallest for which it works is called the *Bachmann-Howard ordinal*. It is usually denoted by $\psi_{\Omega_1}(\varepsilon_{\Omega_1+1})$, where ε_{Ω_1+1} stands for the least ordinal $\eta > \Omega_1$ such that $\omega^\eta = \eta$. And, miracously, the Bachmann-Howard ordinal can be captured by a primitive recursive ORS over the alphabet $0, \Omega_1, +, \omega^{(\cdot)}, \psi_{\Omega_1}$.

Note that the function ψ_{Ω_1} differs significantly from previous proof-theoretic functions, such as $\beta \mapsto \omega^\beta$ and φ_δ, in that $\psi_{\Omega_1}(\alpha)$ can be (and in the most interesting cases is) a smaller ordinal than α. Such proof-theoretic functions have been called *collapsing functions*.

Now, $\psi_{\Omega_1}(\varepsilon_{\Omega_1+1})$ is still much smaller than the proof-theoretic ordinal of $(\Pi^1_1{-}\mathsf{CA})_0$. It is, however, the proof-theoretic ordinal of an important set theory, called *Kripke–Platek set theory* (see [Jäg86]). To climb up to the strength of $(\Pi^1_1{-}\mathsf{CA})_0$ one needs infinitely many sufficiently large ordinals $\Omega_1 < \Omega_2 < \ldots < \Omega_n < \Omega_{n+1} < \ldots$, each equipped with their own collapsing function ψ_{Ω_n}. Again one can use the infinitely many uncountable $\aleph_n$'s to play the role of the Ω_n's. But that again amounts to an enormous overkill. Countable avatars for the Ω_n's suffice. The proof-theoretic ordinal of $(\Pi^1_1{-}\mathsf{CA})_0$ is usually notated by $\psi_{\Omega_1}(\Omega_\omega)$, where $\Omega_\omega = \sup_{n\in\mathbb{N}} \Omega_n$.

By now we have become acquainted with (an idea of) all proof-theoretic ordinals of theories used in RM. Notwithstanding that $(\Pi^1_1{-}\mathsf{CA})_0$ is rather capacious as a framework for ordinary mathematics, there are still very interesting results from graph theory which it cannot prove and to which we turn next. This will provide another example of an independence result obtained via ordinal analysis.

9 The graph minor theorem

If a graph X is obtained from a graph Y by first deleting some vertices and edges, and then contracting some further edges, X is said to be a *minor* of Y. The following theorem holds.

Theorem 9.1 (Robertson and Seymour 1986-2004) *If $G_0, G_1, G_2, \ldots$ is an infinite sequence of finite graphs, then there exist $i < j$ so that G_i is isomorphic to a minor of G_j.*

As to the importance attributed to the graph minor theorem, I quote from a book on graph theory [Die10], p. 333.

> *Our goal* [...] *is a single theorem, one which dwarfs any other result in graph theory and may doubtless be counted among the deepest theorems that mathematics has to offer:* in every infinite set of graphs there are two such that one is a minor of the other. *This* minor theorem*, inconspicuous though it may look at first glance, has made a fundamental impact both outside graph theory and within. Its proof, due to Neil Robertson and Paul Seymour, takes well over 500 pages.*

Theorem 9.1 (GMT hereafter) has many important consequences. Here are a few of them.

Corollary 9.2 *(i)* (Vázsonyi's conjecture) *If all the G_k are trivalent, then there exist $i < j$ so that G_i is embeddable into G_j.*

(ii) (Wagner's conjecture) *For any 2-manifold M there are only finitely many graphs which are not embeddable in M and are minimal with this property.*

A further important consequence of GMT is that any minor closed class of graphs can be characterized by finitely many forbidden minors (a vast generalization of the case of planar graphs). This has important predictive algorithmic consequences: Membership in any minor closed class of graphs can be decided in polynomial (even cubic) time. A case in point (see [Die10, p. 367]) is the class of *knotless* graphs, that is, finite graphs which can be embedded in $\mathbb{R}^3$ such that none of its cycles forms a non-trivial knot. This class is minor closed, so there is a polynomial algorithm. Currently, such an algorithm is not known, but it exists owing to GMT.

GMT is not provable in the strongest system of RM. This independence is a consequence of the ordinal analysis of $(\Pi^1_1\text{-}\mathsf{CA})_0$ in that GMT proves the wellorderedness of its ordinal.

Theorem 9.3 (Friedman, Robertson, Seymour [FRS87]) GMT *is not provable in* $(\Pi^1_1\text{-}\mathsf{CA})_0$.

The paper [KR20] investigated upper bounds for the proof strength of GMT. If one adds a principle of induction, called Π^1_2 *bar induction*, to $\Pi^1_1\text{-}\mathsf{CA}_0$ one can prove GMT as well as many of its generalizations. The resulting system is well within the scope of proof theory of the 1970s.

10 Beyond Π^1_1–comprehension

Proof theorists have widened the realm of theories for which ordinal analyses have been attained way beyond the level Π^1_1–comprehension, especially through the work of Arai (e.g., [Ara15]) and the author (e.g., [Rat95]). The current state of the art is that subsystems of Z_2 with Π^1_2–comprehension can be handled, that is, comprehension of the form

$$\{n \in \mathbb{N} \mid \forall X \exists Y\, \theta(X, Y, n)\}$$

where $\theta(X, Y, n)$ contains no set quantifiers. The difference between the proof power of Π^1_1–comprehension and Π^1_2–comprehension is almost unimaginably huge. In section 8 we have seen that ideas from uncountable cardinals played a role in devising an ORS capable of encapsulating Π^1_1-comprehension. Viewing the representation system as a miniaturization of some cardinal notion has become an important source of inspiration for proof theorists. Accordingly, ideas from large cardinal notions such as inaccessible, Mahlo, and weakly compact cardinals have entered the design of ORSs. A cardinal notion germane to Π^1_2–comprehension is that of the much larger *shrewd cardinals* defined in [Rat95]. Their existence follows from those of subtle cardinals.

The next barrier is Π^1_3–comprehension. One might hope that this is somewhat the generic case that can be generalized to yield an ordinal analysis of any Π^1_n–comprehension, and thus of Z_2. We will see.

Note: The rules of the AMS Notices *for this type of article do not allow more than 20 references. A longer version with the same title and including many more references is available on* arXiv.

References

[Ara15] T. Arai, *Proof theory of ordinals III: Π_N-reflection*, Gentzen's Centenary, 2015, pp. 357–424.

[Bac50] H. Bachmann, *Die Normalfunktionen und das Problem der ausgezeichneten Folgen von Ordinalzahlen*, Vierteljahresschrift der Naturforschergesellschaft Zürich **95** (1950), 115–147.

[Die10] R. Diestel, *Graph theory, 4th ed.*, Springer, 2010. MR1411445

[Fri75] H. Friedman, *Some systems of second order arithmetic and their use*, Proceedings of the International Congress of Mathematicians, Vancouver 1974, vol. 1, 1975, pp. 235–242. MR0429508

[FRS87] H. Friedman, N. Robertson, and P. Seymour, *The metamathematics of the graph minor theorem*, Contemporary Mathematics **65** (1987), 229–262.

[Gen38] G. Gentzen, *Neue Fassung des Widerspruchsfreiheitbeweises für die reine Zahlentheorie*, Forschungen zur Logik und zur Grundlegung der exacten Wissenschaften, Neue Folge **4** (1938), 19–44.

[Goo44] R.L. Goodstein, *On the restricted ordinal theorem*, J. Symb. Log. **9** (1944), 33–41.

[HB38] D. Hilbert and Bernays, *Grundlagen der Mathematik II*, Springer, 1938.

[Hil18] D. Hilbert, *Axiomatisches Denken*, Mathematische Annalen **78** (1918), 405–415.

[Jäg86] G. Jäger, *Theories for admissible sets: a unifying approach to proof theory*, Bibliopolis, 1986.

[KP82] L. Kirby and J. Paris, *Accessible independence results for Peano arithmetic*, Bull. London Math. Soc. **14** (1982), 285–293.

[KR20] M. Krombholz and M. Rathjen, *Upper bounds on the graph minor theorem*, Well-quasi orders in computation, logic, language and reasoning, 2020, pp. 145–159.

[Rat95] M. Rathjen, *Recent advances in ordinal analysis:* Π^1_2*–CA and related systems*, Bulletin of Symbolic Logic **1** (1995), 468–485.

[Rat99] ———, *The realm of ordinal analysis*, Sets and Proofs, 1999, pp. 219–279.

[Sim09] S.G. Simpson, *Subsystems of second order arithmetic, 2nd ed.*, Cambridge University Press, 2009. MR1723993

[Tak03] G. Takeuti, *Memoirs of a Proof Theorist (translated into English by M. Yasugi and, N. Passell)*, World Scientific, New Jersey, 2003. MR1984952

[Tak67] ———, *Consistency proofs of subsystems of classical analysis*, Ann. Math. **86** (1967), 299–348.

[Veb08] O. Veblen, *Continuous increasing functions of finite and transfinite ordinals*, Trans. Amer. Math. Soc. **9** (1908), 280–292.

[Wey18] H. Weyl, *Das Kontinuum*, Veit, Leipzig, 1918.

[Wey21] ———, *Über die neue Grundlagenkrise der Mathematik*, Mathematische Zeitschrift **10** (1921), 39–79.