VICTORIA UNIVERSITY OF WELLINGTON
Te Herenga Waka

School of Mathematics and Statistics
Te Kura Matai Tatauranga

PO Box 600
Wellington 6140
New Zealand

Tel: +64 4 463 5341
Fax: +64 4 463 5045
Email: sms-office@vuw.ac.nz

The reverse mathematics of Cousin's lemma

Jordan Mitchell Barrett

Supervisors: Rod Downey, Noam Greenberg

Friday 30th October 2020

Submitted in partial fulfilment of the requirements for the
Bachelor of Science with Honours in Mathematics.

Abstract

Cousin's lemma is a compactness principle that naturally arises when studying the gauge integral, a generalisation of the Lebesgue integral. We study the axiomatic strength of Cousin's lemma for various classes of functions, using Friedman and Simpson's reverse mathematics in second-order arithmetic. We prove that, over RCA_0:

(i) Cousin's lemma for continuous functions is equivalent to the system WKL_0;

(ii) Cousin's lemma for Baire 1 functions is at least as strong as ACA_0;

(iii) Cousin's lemma for Baire 2 functions is at least as strong as ATR_0.

Contents

Figures

Chapter 1

Introduction

Before the 17th century, mathematics essentially comprised arithmetic, geometry and elementary algebra, and generally only dealt with finite objects [Eve69]. Proofs were almost always *constructive*—a statement would be proved by explicitly constructing a witness. Being largely motivated by physics, mathematics was concerned primarily with calculation, and therefore algorithms took centre stage [MN82].

The development of calculus in the 17th century represented the first signs of departure from this. The ideas were present in Archimedes' method of exhaustion and Cavalieri's method of indivisibles [Eve69], but Newton and Leibniz systematised this, manipulating infinite and infinitesimal quantities as if they were numbers. These new methods proved revolutionary throughout mathematics and physics. There was some concern about the rigour of such methods, and this was not fully abated until the 19th century, when Cauchy, Bolzano, and Weierstrass replaced infinitesimals with more rigorous ε-δ definitions [Cau21; Bol17; Sch61].

As calculus flourished into real analysis, the techniques used became gradually less constructive [MN82]. Early analytical proofs would often implicitly appeal to the infinite pigeonhole principle, and (weak forms of) the axiom of choice [Men16], thereby proving the existence of objects without actually constructing them. Despite some backlash, this trend towards nonconstructivism only continued, as analysis was later abstracted to topology and descriptive set theory.

In the 1870s, while studying a problem in topology, Georg Cantor formulated the concept of ordinals [Can83], leading to the creation of set theory. Cantor was the first to systematically study infinity: famously, he showed in 1874 that the real numbers $\mathbb{R}$ cannot be put into bijection with the natural numbers $\mathbb{N}$, thus demonstrating that there are different sizes of infinity [Can74]. As set theory developed, paradoxes arose (most notably Russell's), and the need for a careful and rigorous foundation for mathematics became clear. One such foundation was provided by ZFC in the 1920s [Zer30].

Cantor's work provided new impetus to mathematical logic, a small subfield of mathematics developed by Boole, De Morgan, and Peano in the mid-to-late 1800s [Boo54; DeM47; Pea89]. Around this time, the ideas of computation, mathematical truth and mathematical proof were formalised for the first time. By the 1930s, logic was a thriving area of mathematics—highlights included Gödel's completeness [Göd29] and incompleteness theorems [Göd31], Turing's negative solution to the *Entscheidungsproblem* [Tur37], Tarski's development of model theory [Vau86], and Hilbert's work on geometry [Hil99] and proof theory [HB34].

A later development in logic was reverse mathematics, initiated by Harvey Friedman in the late 1960s [Fri67; Fri69]. Reverse mathematics asks, for a given theorem of mathematics φ, "what axioms are really necessary to prove φ?" More broadly, it studies the logical

implications between foundational principles of mathematics. An early example was the discovery of non-Euclidean geometries, thereby proving the independence of the parallel postulate from Euclid's other axioms [Lob29; Bol32]. Another early result, more in the style of reverse mathematics, was the demonstration that over ZF, the axiom of choice, Zorn's lemma, and the well-ordering principle are all pairwise equivalent [Bir40; FB58; Tra62].

Traditionally, reverse mathematics is done in second-order arithmetic, in which there are two types of objects: natural numbers $n, m, k, \ldots$, and sets of natural numbers $A, B, C, \ldots$, and quantification is allowed over both types of objects. Restricting oneself to natural numbers may seem unnecessary limiting, but this is not so. In fact, most mathematics deals with countable or "essentially countable" objects (such as separable metric spaces), and so can be formalised in second-order arithmetic. This includes virtually all "classical" mathematics, or that taught in undergraduate courses [Sim09, p. xiv].

In practice, reverse mathematics involves attempting to prove a theorem φ of "ordinary" mathematics in a weak subsystem $\mathcal{S}$ of second-order arithmetic. But, supposing we can do this, how do we know we've found the optimal (weakest) system? The empirical phenomenon is thus:

> *"When the theorem is proved from the right axioms, the axioms can be proved from the theorem."*
>
> —*Harvey Friedman* [Fri74]

This is the "reverse" part of reverse mathematics. Having proved φ from $\mathcal{S}$, to show this is optimal, we want to demonstrate a *reversal* of φ: a proof of $\mathcal{S}$ from φ. This means that φ cannot be proved in a weaker system $\mathcal{S}'$, because if it could, then $\mathcal{S}'$ would also prove $\mathcal{S}$ via φ, meaning $\mathcal{S}'$ is not actually a weaker system after all. Practically speaking, reversals are only possible assuming a weak base system $\mathcal{B}$ (i.e. it is really a proof of $\mathcal{S}$ from $\mathcal{B} + \varphi$).

The utility of reverse mathematics is abundant. Apart from its obvious use in finding the "best" proof of a given statement φ, it also gives us a way to quantify how nonconstructive or noncomputable φ is. The idea is that stronger subsystems correspond to more nonconstructive power, so the "constructiveness" of φ is inversely proportional to the strength of the systems $\mathcal{S}$ in which φ can be proved [FSS83]. Similarly, many theorems guarantee a solution to a given problem—reverse mathematics then tells us how complex the solution could be relative to the problem, which can be made precise in terms of computability. For example, in his thesis [Mil04], Mileti proved the Erdős–Rado canonical Ramsey theorem is equivalent to the system ACA_0. From the proof, he extracted new bounds on the complexity of the homogeneous set, improving the classical bounds obtained by Erdős and Rado.

Here is an example of reverse mathematics in ring theory. The usual way to prove that every commutative ring has a prime ideal is to prove that it has a maximal ideal (Krull's theorem), and then prove every maximal ideal is prime. However, Friedman, Simpson and Smith showed that the existence of maximal ideals is equivalent to the system ACA_0, whereas the existence of prime ideals is equivalent to the strictly weaker system WKL_0 [FSS83]. This shows the usual proof strategy is not optimal—there is a "better" way to prove the existence of prime ideals, which doesn't require the stronger assumption that maximal ideals exist.

In this report, we examine the reverse-mathematical content of Cousin's lemma, a particular statement in analysis. Cousin's lemma can be viewed as a kind of compactness principle, asserting that every positive valued function $\delta\colon [0,1] \to \mathbb{R}^+$ has a *partition*—a finite sequence $t_0, \ldots, t_{n-1}$ such that the open balls $B\big(t_i, \delta(t_i)\big)$ cover $[0,1]$. In particular, we establish the following original results over the weak base theory RCA_0:

(i) Cousin's lemma for continuous functions is equivalent to the system WKL_0;

(ii) Cousin's lemma for Baire 1 functions is at least as strong as ACA_0;

(iii) Cousin's lemma for Baire 2 functions is at least as strong as ATR_0.

Notational conventions

Throughout this report, we abide to the following notational conventions.

- We let $\mathbb{N} = \{0, 1, 2, \ldots\}$ be the set of *nonnegative* integers, following the usual practice in logic.
- We will often use $\bar{x}$ to notate a tuple $(x_1, \ldots, x_d)$, where the length should be clear from context.
- We may use $A^{\complement} = \{x : x \notin A\}$ to denote the (absolute) complement of a set A, particularly for sets of natural numbers.
- For mathematical statements φ and ψ, we use $\varphi \vdash \psi$ ("φ proves ψ") to mean there is a proof of ψ from φ. This notation extends to formal systems, e.g. $\mathcal{S} \vdash \varphi$ means there is a proof of φ in the formal system $\mathcal{S}$.
- For a statement φ and a structure $\mathcal{M}$, we use $\mathcal{M} \vDash \varphi$ ("$\mathcal{M}$ models φ") to mean the statement φ is true in $\mathcal{M}$. Similarly, $\mathcal{M} \vDash \mathcal{S}$ means that all axioms of the formal system $\mathcal{S}$ are true in $\mathcal{M}$.

Chapter 2

Integration and Cousin's lemma

The main object of study in this report is *Cousin's lemma*, a compactness principle phrased in terms of positive real-valued functions, rather than open covers. Cousin's lemma arises naturally in the study of the *gauge integral*, a generalisation of the Riemann and Lebesgue integrals due to Kurzweil [Kur57] and Henstock [Hen63]. In this chapter, we review Riemann integration, before generalising to gauge integration and defining Cousin's lemma.

2.1 Riemann integration

The basic idea of Riemann integration is thus: approximate the area under a curve by a series of rectangles, as in Figure 2.1. As we increase the number of rectangles, and decrease their width, we hope that this approximation becomes closer and closer to the true area. Here, we will only consider integration over the unit interval $[0,1]$.

Definition 2.1.1. A *tagged partition of* $[0,1]$ is a finite sequence

$$P = \langle 0 = x_0 < t_0 < x_1 < t_1 < \cdots < t_{n-1} < x_n = 1 \rangle.$$

We call n the *size* of P.

A tagged partition $P = \langle x_i, t_i \rangle$ should be interpreted as follows. The x_i are the *partition points* at which the interval $[0,1]$ is split, and within each subinterval or block $[x_i, x_{i+1}]$, we choose a *tag point* t_i. When using P to approximate the area underneath a function f, each subinterval $[x_i, x_{i+1}]$ will serve as the base of a rectangle of height $f(t_i)$. This is illustrated in Figure 2.1.

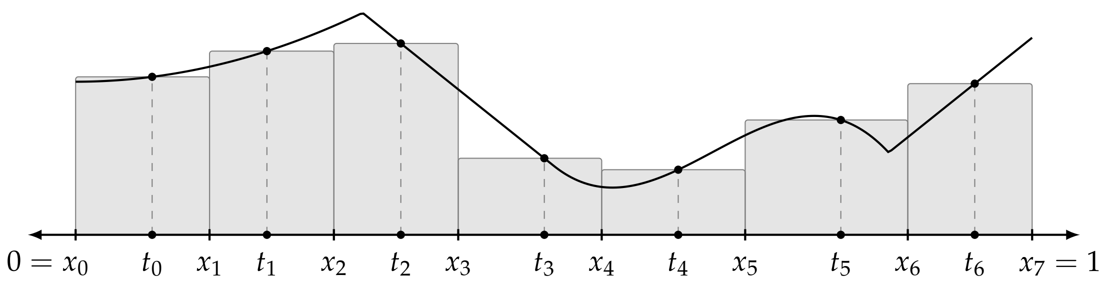

Figure 2.1: A Riemann sum of a continuous function over a partition of size 7.

Definition 2.1.2. Let $f\colon [0,1] \to \mathbb{R}$ be a function, and $P = \langle x_i, t_i \rangle$ a partition of size n. The *Riemann sum of f over P* is

$$\mathrm{RS}(f, P) = \sum_{i=0}^{n-1} f(t_i)[x_{i+1} - x_i]$$

Definition 2.1.3 [Rie54]**.** A function $f\colon [0,1] \to \mathbb{R}$ is *Riemann integrable* if there exists $K \in \mathbb{R}$ such that, for every $\varepsilon > 0$, there exists $\delta > 0$ with $|\mathrm{RS}(f, P) - K| < \varepsilon$ whenever each block of P has size $< \delta$. In this case, we say that K is the *Riemann integral of f*.

The Riemann integral can deal with virtually all functions which one may want to integrate in practice. Riemann (and later Lebesgue) gave a characterisation of exactly which bounded functions are Riemann integrable:

Proposition 2.1.4 [Bro36; Bir73]**.** A bounded function $f\colon [0,1] \to \mathbb{R}$ is Riemann integrable if and only if its set of discontinuities has (Lebesgue) measure zero.

In particular, every continuous function is Riemann integrable. That said, it is not difficult to construct functions which are not Riemann integrable.

Proposition 2.1.5 [Dir29]**.** There are functions which are not Riemann integrable.

Proof. The characteristic function $\chi_\mathbb{Q}$ of $\mathbb{Q}$, also known as Dirichlet's function, provides an example. Concretely, $\chi_\mathbb{Q}\colon [0,1] \to \mathbb{R}$ is defined

$$\chi_\mathbb{Q}(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ 1 & \text{if } x \text{ is rational} \end{cases}$$

We will show $\chi_\mathbb{Q}$ is not Riemann integrable. Pick any $K \in \mathbb{R}$, and let $\varepsilon = 1/3$. Then, for any $\delta > 0$, pick $n > 1/\delta$, and consider the partitions $P = \langle x_i, t_i \rangle$, $P' = \langle x'_i, t'_i \rangle$ of size n defined by:

$$\begin{aligned} x_i = x'_i &= i/n; \\ t_i &= \text{ some irrational point in } [i/n, (i+1)/n]; \\ t'_i &= \text{ some rational point in } [i/n, (i+1)/n]. \end{aligned}$$

Then, each block of P, P' has size $< \delta$, and $\mathrm{RS}(\chi_\mathbb{Q}, P) = 0$ while $\mathrm{RS}(\chi_\mathbb{Q}, P') = 1$. Thus, K cannot be within $\varepsilon = 1/3$ of both. □

2.2 Gauge integration

So, what failed when trying to integrate $\chi_\mathbb{Q}$? Morally, since almost all real numbers in $[0,1]$ are irrational (in the measure-theoretic sense), the integral of $\chi_\mathbb{Q}$ ought to be equal to zero. There were many attempts to solve this, the most famous being Lebesgue's measure theory [Leb02; Leb04]. However, Lebesgue integration is not without its issues—in particular, there are derivatives which are not Lebesgue integrable [Gor96].

Attempting to remedy this, Denjoy defined an integral which could handle all derivatives [Den12]. Shortly after, Luzin [Luz12] and Perron [Per14] gave equivalent characterisations of Denjoy's integral. However, all these definitions were complex and highly nonconstructive, making Denjoy's integral impractical for applications [Gor96].

In 1957, Kurzweil defined the *gauge integral*, a generalisation of Denjoy's integral. He formulated it in elementary terms similar to the Riemann integral [Kur57], thus avoiding the complications of measure theory. Later, Henstock systematically developed the theory of the gauge integral [Hen63]— as a result, it is sometimes known as the Henstock–Kurzweil integral. Kurzweil's ingenious solution was to allow the parameter δ in Definition 2.1.3 to be a *variable*, rather than a constant. In effect, this ensures that some partitions are not allowed, such as P' in the proof of Proposition 2.1.5.

Definition 2.2.1 [Gor94]**.** A *gauge* is a strictly positive-valued function $\delta\colon [0,1] \to \mathbb{R}^+$.

The idea is that δ tells us how *fine* our partition needs to be at any point. At points x where the function is highly discontinuous, or varies greatly, we could make sure $\delta(x)$ is small, so that we only consider partitions which are divided finely enough around x.

Definition 2.2.2 [Gor94]**.** Given a gauge δ, a partition $P = \langle x_i, t_i \rangle$ is *δ-fine* if, for any $i < n$, the open ball $B(t_i, \delta(t_i))$ contains (x_i, x_{i+1}).

Definition 2.2.3 [Kur57]**.** A function $f\colon [0,1] \to \mathbb{R}$ is *gauge integrable* if there exists $K \in \mathbb{R}$ such that, for every $\varepsilon > 0$, there exists a gauge $\delta\colon [0,1] \to \mathbb{R}^+$ with $|\mathrm{RS}(f,P) - K| < \varepsilon$ whenever P is δ-fine. In this case, we say that K is the *gauge integral of f*.

Note that if $\delta(x) = k$ is constant, a partition P is δ-fine if and only if the blocks of P have size $< 2k$. Thus, Riemann integration is a special case of gauge integration, where we only allow constant gauges. It follows that every Riemann integrable function is gauge integrable.[1] However, the converse does not hold, as we now see:

Proposition 2.2.4 [KS04]**.** There are functions which are gauge integrable, but not Riemann integrable.

Proof. Dirichlet's function $\chi_{\mathbb{Q}}$ is again an example. We saw in Proposition 2.1.5 that $\chi_{\mathbb{Q}}$ is not Riemann integrable—we now show that it is gauge integrable, with integral $K = 0$. Pick any $\varepsilon > 0$, and let $\mathbb{Q} = \{q_0, q_1, q_2, \dots\}$ enumerate the rationals. Define $\delta\colon [0,1] \to \mathbb{R}^+$ by $\delta(q_m) = 2^{-m-2}\,\varepsilon$, and $\delta(x) = 1$ for all irrational x. Now suppose $P = \langle x_i, t_i \rangle$ is δ-fine. Then,

$$\begin{aligned}
\mathrm{RS}(\chi_{\mathbb{Q}}, P) &= \sum_{i=0}^{n-1} \chi_{\mathbb{Q}}(t_i)[x_{i+1} - x_i] \\
&= \sum_{\substack{i<n \\ t_i \text{ rational}}} \chi_{\mathbb{Q}}(t_i)[x_{i+1} - x_i] \quad + \sum_{\substack{i<n \\ t_i \text{ irrational}}} \chi_{\mathbb{Q}}(t_i)[x_{i+1} - x_i] \\
&= \sum_{\substack{i<n \\ t_i \text{ rational}}} (x_{i+1} - x_i) \quad + \quad 0 \\
&\leq \sum_{\substack{i<n \\ t_i = q_{m_i}}} 2^{-m_i-1}\,\varepsilon \quad < \quad \varepsilon
\end{aligned}$$

□

2.3 Cousin's lemma

If there were a gauge δ with no δ-fine partition, then Definition 2.2.3 could be vacuously satisfied by choosing this gauge. This would present a problem: every $K \in \mathbb{R}$ would then witness that every f is gauge integrable, so we could not uniquely define the value of the gauge integral. Cousin's lemma states that this situation cannot happen. It is originally due to Cousin [Cou95], who proved the statement in a radically different form.

Lemma 2.3.1 (Cousin's lemma) [Cou95; KS04]**.** Every gauge $\delta\colon [0,1] \to \mathbb{R}^+$ has a δ-fine partition.

[1] Furthemore, the Riemann integral and gauge integral of f will have the same value.

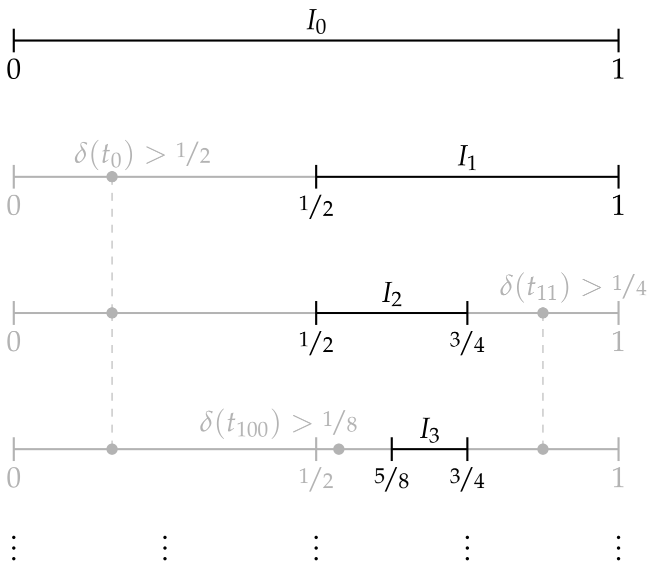

Figure 2.2: The recursive interval-splitting process used in the proof of Cousin's lemma.

Proof [Lee11]. If there is $t \in [0,1]$ with $\delta(t) > 1$, then the partition $\langle 0, t, 1 \rangle$ is δ-fine, so we are done.

Otherwise, split $[0,1]$ into halves $[0, 1/2]$, $[1/2, 1]$, and ask if there are $t_0 \in [0, 1/2]$, $t_1 \in [1/2, 1]$ with $\delta(t_0), \delta(t_1) > 1/2$. If such a t_0 exists, we don't need to split further, since $\delta(t_0)$ covers the subinterval $[0, 1/2]$.

If no such t_0 exists, we split $[0, 1/2]$ into halves $[0, 1/4]$, $[1/4, 1/2]$, and ask if there are $t_{00} \in [0, 1/4]$, $t_{01} \in [1/4, 1/2]$ with $\delta(t_{00}), \delta(t_{01}) > 1/4$. Keep repeating this process, and do similar on the side of t_1.

We claim this procedure must eventually terminate. Suppose it did not — then, there is a nested sequence $[0,1] = I_0 \supsetneq I_1 \supsetneq I_2 \supsetneq \cdots$ of closed intervals, each half the size of the previous. Since $[0,1]$ is compact, we can pick $r \in \bigcap_{n=0}^{\infty} I_n$. But $\delta(r)$ is positive, so $\delta(r) > 2^{-n}$ for sufficiently large n. Then, the procedure would have terminated at stage n, since we would have found $t = r \in I_n$. $\square$

Among other things, Cousin's lemma implies that the value of a gauge integral is unique, if it exists:

Corollary 2.3.2 [KS04]**.** For gauge integrable $f \colon [0,1] \to \mathbb{R}$, there is a unique K witnessing the integrability of f.

Proof. By contradiction, suppose $K_1 \neq K_2$ both witness that f is gauge integrable. Let $\varepsilon = |K_1 - K_2|/3$. By assumption, there are gauges δ_1, δ_2 witnessing K_1, K_2 respectively for this choice of ε. Let $\delta(x) := \min\{\delta_1(x), \delta_2(x)\}$ be the pointwise minimum: this is also a gauge. By Cousin's lemma, there is a δ-fine partition P, which must also be δ_1-fine and δ_2-fine by definition of δ. So $|\mathrm{RS}(f,P) - K_1| < \varepsilon$ and $|\mathrm{RS}(f,P) - K_2| < \varepsilon$, a contradiction. $\square$

We can view Cousin's lemma as a kind of compactness principle. Effectively, it asserts that the open cover $\{B(t, \delta(t)) : t \in [0,1]\}$ has a finite subcover, corresponding to the tag points of a δ-fine partition.

Chapter 3

Logical prerequisites

Before we delve into reverse mathematics, it is necessary to have some background in computability and model theory, insofar as they apply to our setting of second-order arithmetic. First, we review the basic concepts of computability, including computable functions, c.e. sets and the universal function. We then develop the model theory of second-order arithmetic, and define the arithmetical and analytical hierarchies.

3.1 Computability

Reverse mathematics is best understood with a background in computability theory. This is because we generally work over the base system RCA_0, which can be thought of as the "computable world". A reversal of a statement φ in a formal system $\mathcal{S}$ is then a proof of $\mathcal{S}$ from $\mathsf{RCA}_0 + \varphi$. In practice, this involves a computable reduction between $\mathcal{S}$ and φ, hence the importance of computability.

The key notion in computability is that of an *algorithm*, an exact method by which something can be computed. Algorithms can be formalised in many ways—Turing machines, the λ-calculus, μ-recursive functions. All of these formalisations are provably equivalent, and the widely-accepted *Church–Turing thesis* posits that each faithfully captures the idea of something being calculable. Therefore, to avoid unnecessary formality,[1] the following intuitive "definition" is sufficient for us:

"Definition" 3.1.1. An *elementary instruction* is one that can be performed mechanically. An *algorithm* is a finite collection of unambiguous elementary instructions, to be carried out in a specified order. Instructions may be repeated.

We will allow our algorithms to take natural numbers as input, and act on this input during the computation. Furthermore, we will expect our algorithm to produce a natural number as output if the computation terminates, or *halts*.

"Definition" 3.1.1 is intentionally very broad—almost all processes arising in mathematics and elsewhere qualify as algorithms. However, there are processes which don't. The archetypal example in computability is the *halting problem*—determining whether a given program halts on a given input. A non-algorithmic process from classical mathematics arises in the proof of the Bolzano–Weierstrass theorem. Given a sequence $(x_n)_{n=1}^{\infty}$ of real numbers bounded in the interval I, we split I into halves, take a half I' which contains infinitely many of the x_n, and repeat the splitting on I', ad infinitum.

[1] Rigorous definitions of computability are available in any introductory textbook on the subject. See [Rog67; Soa87; Soa16].

Algorithm 3.1.2.

```
1    let x := 0
2    increment x by 1
3    goto line 2
4    halt
```

Figure 3.1: An algorithm which never halts.

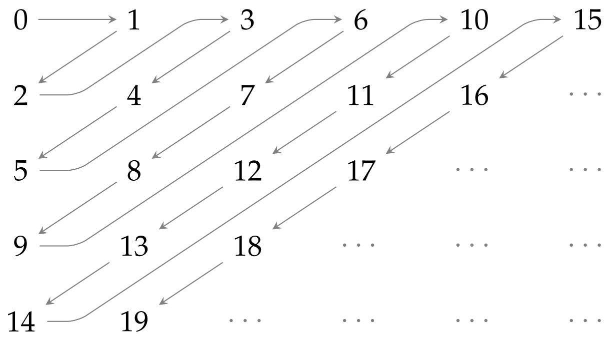

Figure 3.2: Cantor's pairing function.

The problem is that the instruction "take a half I' containing infinitely many x_n" is *not* elementary, in the sense that it is not possible to carry it out given an arbitrary sequence (x_n). We could start counting along the sequence, noting which half x_1 is in, then x_2, and so on, but we will never know which one contains *infinitely* many x_n. Indeed, a computable version of the Bolzano–Weierstrass theorem fails for this reason [Spe49].

There is a key difference between "Definition" 3.1.1 and our intuitive understanding of an algorithm—for us, algorithms *do not* have to halt. Algorithm 3.1.2 in Figure 3.1 is a simple example. This algorithm indeed satisfies "Definition" 3.1.1—each instruction is elementary and unambiguous, and the order in which they are to be executed (1, 2, 3, 2, 3, 2, 3, ...) is hopefully evident. However, Algorithm 3.1.2 never halts, as it will continually increment the variable x, never reaching line 4. Notice that while algorithms may not halt, if they do, this must happen in finite time.

To model this idea, we instead consider *partial functions* $f\colon \mathbb{N} \to \mathbb{N}$, i.e. functions $f\colon A \to \mathbb{N}$ for some subset $A \subseteq \mathbb{N}$, called the *domain* of f.

Definition 3.1.3. A partial function $f\colon \mathbb{N} \to \mathbb{N}$ is *computable* if there exists an algorithm which, on input $n \in \mathbb{N}$:

(i) If $n \in \operatorname{dom}(f)$: halts and outputs $f(n)$;

(ii) If $n \notin \operatorname{dom}(f)$: doesn't halt.

Definition 3.1.4. A set $A \subseteq \mathbb{N}$ is *computable* if its characteristic function $\chi_A\colon \mathbb{N} \to \{0,1\}$ is computable, in the sense of Definition 3.1.3. Concretely, $A \subseteq \mathbb{N}$ is computable if there exists an algorithm which, given input n, always halts, returning 1 if $n \in A$, and 0 if $n \notin A$.

Informally, a set is computable if there is an algorithm which tells us whether or not any given element is in the set. We can extend Definition 3.1.3 to functions $f\colon \mathbb{N}^d \to \mathbb{N}^k$, and Definition 3.1.4 to sets $A \subseteq \mathbb{N}^d$, via the *pairing function*:

Definition 3.1.5 [Can77]**.** The *2-pairing function* $\pi_2\colon \mathbb{N}^2 \to \mathbb{N}$ is the bijection

$$\pi_2(m,n) = \frac{(m+n)(m+n+1)}{2} + m.$$

Further, we define the *3-pairing function* $\pi_3\colon \mathbb{N}^3 \to \mathbb{N}$, $(m,n,k) \mapsto \pi_2\big(m, \pi_2(n,k)\big)$, the *4-pairing function* $\pi_4\colon \mathbb{N}^4 \to \mathbb{N}$, $(m,n,k,\ell) \mapsto \pi_2\big(m, \pi_3(n,k,\ell)\big)$, etc. These are all bijections.

The pairing functions allow us to treat tuples of natural numbers as single natural numbers, and therefore define computability for tuples of natural numbers. Often, we will implicitly use the pairing functions to think of a d-tuple $\bar{x}$ as just a natural number.

Definition 3.1.6. A set $B \subseteq \mathbb{N}^d$ is *computable* if $A = \pi_d(B)$ is computable, in the sense of Definition 3.1.4.

Definition 3.1.7. A partial function $f\colon \mathbb{N}^d \to \mathbb{N}^k$ is *computable* if $\tilde{f} = \pi_k^{-1} \circ f \circ \pi_d$ is computable, in the sense of Definition 3.1.3.

In computability, it is often useful to consider sets which are almost computable, but not quite. The computably enumerable sets provide examples of such things.

Definition 3.1.8. Given a set $B \subseteq \mathbb{N}^2$ of pairs, the *projection of* B is the set

$$\mathrm{proj}(B) = \{m \in \mathbb{N} : \exists n\ (m, n) \in B\}.$$

A set $A \subseteq \mathbb{N}$ is *computably enumerable* (c.e.) if there is a computable set $B \subseteq \mathbb{N}^2$ such that $A = \mathrm{proj}(B)$.

Every computable set A is c.e., since it is the projection of $A \times \{0\} = \{(n, 0) : n \in A\}$. The converse does not hold, and there are many noncomputable c.e. sets; the archetypal example is the *halting problem*.

Proposition 3.1.9. *$A \subseteq \mathbb{N}$ is computable if and only if both A and $A^{\complement}$ are c.e..*

Proof. In the forward direction, $A^{\complement}$ is also computable, and we have noted that every computable set is c.e.. Conversely, suppose $A = \mathrm{proj}(B)$, $A^{\complement} = \mathrm{proj}(C)$ for computable $B, C \subseteq \mathbb{N}^2$. Given $n \in \mathbb{N}$, we decide if $n \in A$ as follows: first check if $(n, 0) \in B$, then check if $(n, 0) \in C$, then if $(n, 1) \in B$, then if $(n, 1) \in C$, and so on. Since it is true that either $n \in A$ or $n \in A^{\complement}$, eventually this algorithm will halt. $\square$

We close this section with a fundamental result of computability, originally due to Turing [Tur37]. By definition, every algorithm admits a finite description. Therefore, we can code algorithms by natural numbers, using a suitable coding scheme. For example, we could code each algorithm in a fixed programming language, and interpret its ASCII code as a natural number written in binary. For a nice coding scheme such as this, we can computably decode these numbers back into functions, and thus compute a function from its code. More formally:

Theorem 3.1.10 [Tur37]**.** *There is a partial computable function $U\colon \mathbb{N}^2 \to \mathbb{N}$ with the following property: for any partial computable function $f\colon \mathbb{N} \to \mathbb{N}$, there is $e \in \mathbb{N}$ such that $U(e, n) = f(n)$ for all n.*

Proof. Compute U as follows: given input (e, n), interpret e as a code for a computable function f, decode it, and compute $f(n)$. $\square$

We call U a *universal computable function*. Essentially, U can be interpreted as a compiler: it takes in the code e of a function and returns the function itself. A corollary of Proposition 3.1.10 is that the sequence $\varphi_0, \varphi_1, \varphi_2, \ldots$, where $\varphi_e(n) = U(e, n)$, lists all the partial computable functions. Furthermore, this is a *uniformly computable listing*, meaning there is an algorithm taking (e, n) to $\varphi_e(n)$ (namely, the algorithm for U). The existence of such a sequence will be useful later.

3.2 Second-order arithmetic

Now, we develop the necessary model-theoretic tools within the setting of arithmetic. The reader may have heard of first-order arithmetic, more commonly known as Peano arithmetic (PA). The reason PA is *first-order* is that quantification is only allowed over natural numbers. For example, a number p being prime is expressible in PA (for all *natural numbers* $m < p$, m divides p iff $m = 1$ or $m = p$), but not the well-foundedness of $\mathbb{N}$ (for every *subset* $A \subseteq \mathbb{N}$, A has a least element).

Here, we work in the stronger setting of *second-order arithmetic*, where quantification over subsets is allowed. We review basic model theory in this setting [CK90; Mar02; Sim09]; in short, we consider structures in the *language of second-order arithmetic* $\mathcal{L}_2 = \{0, 1, +, \cdot, <, \in\}$. This is a two-sorted language, meaning we have two kinds of objects: numbers (denoted by lowercase letters $n, m, k, \ldots$), and sets (denoted in uppercase $A, B, C, \ldots$). The symbols in $\mathcal{L}_2$ are *typed*, e.g. 0 is a constant symbol of number type, $+$ is a binary operation between two object of number type, $\in$ is a binary relation between an object of number type and one of set type, etc.

We can build terms from symbols in $\mathcal{L}_2$, and we have two kinds of terms: numerical terms and set terms. As is usual, terms may include variable symbols, of number type x or set type X. In fact, the only terms of set type are the set variable symbols $X, Y, Z, \ldots$, but there are a wealth of numerical terms:

Definition 3.2.1. The collection of *numerical $\mathcal{L}_2$-terms* is defined as follows:

(i) 0, 1, and any numerical variable symbol x are numerical terms.

(ii) If s, t are numerical terms, then $(s + t)$ and $(s \cdot t)$ are numerical terms.

Intuitively, a numerical term represents a natural number. For example, $(1+1)$, $(1+(0+1))$ and $((1 \cdot (1+1)) + 0)$ are all numerical terms, all representing the number 2. However, these are all different *terms*, since they do not contain the same arrangement of symbols. Frequently, we will omit brackets where there is no ambiguity—the above terms might be written more concisely as $1+1$, $1+0+1$ and $1 \cdot (1+1) + 0$. We will use k to abbreviate the numerical term $\underbrace{1 + 1 + \cdots + 1}_{k \text{ times}}$.

Definition 3.2.2. The collection of *$\mathcal{L}_2$-formulae* is defined as follows:

(i) If s, t are numerical terms, and X is a set variable symbol, then $(s = t)$, $(s < t)$ and $(s \in X)$ are formulae.

(ii) If φ, ψ are formulae, then $(\neg\varphi)$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$ and $(\varphi \leftrightarrow \psi)$ are formulae.

(iii) If φ is a formula, then $(\forall x\ \varphi)$ and $(\exists x\ \varphi)$ are formulae.

(iv) If φ is a formula, then $(\forall X\ \varphi)$ and $(\exists X\ \varphi)$ are formulae.

Intuitively, formulae are statements that may be true or false in a particular situation. Again, unnecessary brackets will often be omitted. We distinguish two types of variables in formulae: *bound variables*, which are preceded by a quantifier over that variable, and *free variables*, which are not. For example, in the formula $\forall x\ (x + y = 1)$, the variable x is bound by the quantifier $\forall x$, while y is free.

Definition 3.2.3. An *$\mathcal{L}_2$-sentence* is an $\mathcal{L}_2$-formula in which all variables are bound.

If φ contained any free variables, then the truth or falsity of φ could conceivably depend on what values were assigned to those free variables. Thus, an $\mathcal{L}_2$-sentence is a formula that can be assigned an unconditional truth value.

Definition 3.2.4. Suppose $\varphi(x_1, \ldots, x_n, X_1, \ldots, X_m)$ is an $\mathcal{L}_2$-formula in free number variables $x_1, \ldots, x_n$, and free set variables $X_1, \ldots, X_m$. Then, the *universal closure* of φ is the $\mathcal{L}_2$-sentence

$$\forall x_1 \cdots \forall x_n \, \forall X_1 \cdots \forall X_m \;\; \varphi(x_1, \ldots, x_n, X_1, \ldots, X_m)$$

Now, what does it mean for an $\mathcal{L}_2$-sentence φ to be true, or false? As is usual in model theory, truth of φ is defined relative to a *model*, consisting of a universe $\mathcal{M}$ of elements and interpretations in $\mathcal{M}$ for all symbols in our language. Since we are working with two sorts (numbers x and sets X), we need to provide *both* a universe of *numbers* $\mathcal{A}$ and a universe of *sets* $\mathcal{B}$, and interpret the symbols in $\mathcal{L}_2$ appropriately. For example, we would interpret $\in$ as a relation between elements of $\mathcal{A}$ and elements of $\mathcal{B}$.

In theory, we could pick any sets $\mathcal{A}$, $\mathcal{B}$ to serve as the universe for our model, and interpret the symbols in $\mathcal{L}_2$ any way we like. However, we will only be interested in the so-called *ω-models*, where $\mathcal{A} = \mathbb{N}$ is the natural numbers, and the symbols $0, 1, +, \cdot, <$ are given their usual interpretations in $\mathbb{N}$.

Definition 3.2.5. An *ω-model of second-order arithmetic* is a subset $\mathcal{B} \subseteq \mathcal{P}(\mathbb{N})$.

A priori, there is no reason that our universe of sets $\mathcal{B}$ actually must consist of subsets of $\mathbb{N}$. We could theoretically pick *any* universe $\mathcal{B}$, and interpret the relation $x \in X$ in *any* way we like. However, we can always identify $\mathcal{B}$ with a subset of $\mathcal{P}(\mathbb{N})$ by identifying each $X \in \mathcal{B}$ with the set $\overline{X} = \{n \in \mathbb{N} : \text{the formula “}n \in X\text{” holds}\}$. So, no generality is lost in only considering subsets of $\mathcal{P}(\mathbb{N})$ in Definition 3.2.5.

Definition 3.2.6. Given an ω-model $\mathcal{M}$, *truth of a sentence φ in $\mathcal{M}$* (notated $\mathcal{M} \vDash \varphi$) is defined in the evident way:

- Numerical terms s, t are given their standard interpretations $s^{\mathcal{M}}$ and $t^{\mathcal{M}}$ in $\mathbb{N}$;
- $\mathcal{M} \vDash (s = t)$ if $s^{\mathcal{M}}$ and $t^{\mathcal{M}}$ are the same natural number;
- $\mathcal{M} \vDash (s < t)$ if $s^{\mathcal{M}}$ is a smaller natural number than $t^{\mathcal{M}}$;
- The rules for Boolean connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$ are as usual;
- $\mathcal{M} \vDash \big(\exists x \; \varphi(x)\big)$ if there is some $n \in \mathbb{N}$ such that $\mathcal{M} \vDash \varphi(n)$;
- $\mathcal{M} \vDash \big(\forall x \; \varphi(x)\big)$ if $\mathcal{M} \vDash \varphi(n)$ for any choice of $n \in \mathbb{N}$;
- $\mathcal{M} \vDash \big(\exists X \; \varphi(X)\big)$ if there is some set $A \in \mathcal{M}$ such that $\mathcal{M} \vDash \varphi(A)$;
- $\mathcal{M} \vDash \big(\forall X \; \varphi(X)\big)$ if $\mathcal{M} \vDash \varphi(A)$ for any set $A \in \mathcal{M}$.

The key part of Definition 3.2.6 is that set quantifiers $\forall X$ and $\exists X$ should be interpreted as ranging over exactly the sets *in the model* $\mathcal{M}$. This is the key difference between the different ω-models. As an example, the sentence $\exists X \; (0 = 0)$ is false in the ω-model $\varnothing$, but is true in any other ω-model.

3.3 The arithmetical and analytical hierarchies

The collection $\mathcal{F}$ of all $\mathcal{L}_2$-formulae, as in Definition 3.2.2, is an extremely rich and varied class. We wish to stratify $\mathcal{F}$ based on the complexity of formulae it contains. Our chosen measure of complexity will be based on the quantifiers, their type (numerical or set), and the number of alternations between universal ($\forall$) and existential ($\exists$). This way, we classify $\mathcal{F}$ into structures known as the *arithmetical hierarchy* and the *analytical hierarchy*.

The lowest level of complexity consists of formulae containing only bounded quantifiers: those of the form $\forall x \; (x < k \rightarrow \psi)$ or $\exists x \; (x < k \rightarrow \psi)$ for some constant $k \in \mathbb{N}$. We will

$$
\begin{tikzcd}
\Sigma^0_0 & & \Sigma^0_1 \arrow[dr, phantom, sloped, "\subsetneq"] & & \Sigma^0_2 & \cdots & \Sigma^1_0 & & \Sigma^1_1 \arrow[dr, phantom, sloped, "\subsetneq"] & & \cdots \\
\Delta^0_0 \arrow[u, equal] \arrow[d, equal] \arrow[r, phantom, "\subsetneq"] & \Delta^0_1 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] & & \Delta^0_2 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] & & \cdots & \Delta^1_0 \arrow[u, equal] \arrow[d, equal] \arrow[r, phantom, "\subsetneq"] & \Delta^1_1 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] & & \Delta^1_2 & \cdots \\
\Pi^0_0 & & \Pi^0_1 \arrow[ur, phantom, sloped, "\subsetneq"] & & \Pi^0_2 & \cdots & \Pi^1_0 & & \Pi^1_1 \arrow[ur, phantom, sloped, "\subsetneq"] & & \cdots
\end{tikzcd}
$$

Figure 3.3: The arithmetical (left) and analytical (right) hierarchies.

often abbreviate these to $(\forall x < k)\, \psi$ and $(\exists x < k)\, \psi$ respectively. From there, universal formulae are given Π classifications, and existential formulae given Σ classifications.

Definition 3.3.1 (arithmetical hierarchy for formulae)**.** Let φ be an $\mathcal{L}_2$-formula. We assign classifications to φ as follows:

(i) φ is called Σ^0_0 and Π^0_0 if it only contains bounded quantifiers.

(ii) φ is called Σ^0_{n+1} if it is of the form $\varphi = \exists x_1 \cdots \exists x_n\ \psi$, where ψ is Π^0_n.

(iii) φ is called Π^0_{n+1} if it is of the form $\varphi = \forall x_1 \cdots \forall x_n\ \psi$, where ψ is Σ^0_n.

We say φ is *arithmetical* if it receives any of these classifications.

We also translate the arithmetical hierarchy from formulae to sets defined by those formulae. This gives us a measure of complexity for subsets of $\mathbb{N}$. Here, we obtain additional Δ classifications.

Definition 3.3.2 (arithmetical hierarchy for sets)**.**

(i) A set $A \subseteq \mathbb{N}$ is called Σ^0_n if there is a Σ^0_n formula $\varphi(x)$ in one free variable such that $A = \{n \in \mathbb{N} : \varphi(n) \text{ holds}\}$. Π^0_n sets are defined analogously.

(ii) $A \subseteq \mathbb{N}$ is called Δ^0_n if it is both Σ^0_n and Π^0_n.

We say A is *arithmetical* if it receives any of these classifications, or equivalently, if it is definable by an arithmetical formula.

For sets, we have that $\Sigma^0_n \subseteq \Pi^0_{n+1}$. If $\varphi(x)$ is a Σ^0_n formula defining A, and y is a variable not in φ, then $\psi(x) = \forall y\ \varphi(x)$ is a Π^0_{n+1} formula also defining A (since the truth value of φ does not depend on y). We also get that $\Sigma^0_n \subseteq \Sigma^0_{n+1}$ by placing such "dummy quantifiers" after all others, whence $\Sigma^0_n \subseteq \Delta^0_{n+1}$. By taking complements, $\Pi^0_n \subseteq \Delta^0_{n+1}$, thus $\Sigma^0_n \cup \Pi^0_n \subseteq \Delta^0_{n+1}$. In fact, all of these containments are strict, but we will not prove this here.

There is a close relationship between the arithmetical hierarchy and computability:

Proposition 3.3.3. Every Δ^0_0 set is computable.

Proof. To say $A \subseteq \mathbb{N}$ is Δ^0_0 is to say that there is an $\mathcal{L}_2$-formula $\varphi(x)$ such that $n \in A \iff \varphi(n)$ holds, and where all quantifiers in $\varphi(x)$ are bounded. Since $\varphi(x)$ can only contain finitely many quantifiers (say d-many), there are only finitely many tuples $(a_1, \ldots, a_d) \in \mathbb{N}^d$ that we need to check to verify whether $\varphi(n)$ holds or not. So, the algorithm to compute A is simply checking all such tuples exhaustively. $\square$

Proposition 3.3.4 [Kle43; Pos44]**.** $A \subseteq \mathbb{N}$ is Σ^0_1 if and only if it is computably enumerable.

Proof. We prove the forward direction. If A is Σ_1^0, then it can be defined by a formula $\varphi(x)$ of the form $\exists y_1 \cdots \exists y_n \ \psi(x, y_1, \dots, y_n)$, where ψ is Δ_0^0. By Proposition 3.3.3, the set $B := \{(x, \bar{y}) : \psi(x, \bar{y})\}$ is computable, and $A = \text{proj}(B)$. $\square$

The reverse direction of Proposition 3.3.4 is harder to prove. It requires coding algorithms using $\mathcal{L}_2$-formulae, for which a formal definition of computability is needed. Therefore, we will not complete the proof here, but it can be found in [Dav58; Rog67].

Corollary 3.3.5. $A \subseteq \mathbb{N}$ is Δ_1^0 if and only if it is computable.

Proof. A is Δ_1^0 if and only if A is both Σ_1^0 and Π_1^0. By the negation rules for $\forall$ and $\exists$ quantifiers, A is Π_1^0 if and only if $A^\complement$ is Σ_1^0. By Proposition 3.3.4, this is if and only if A and $A^\complement$ are c.e.. Hence, the result follows from Proposition 3.1.9. $\square$

Clearly, there are nonarithmetical formulae—any formula containing a set quantifier is an example. It is less obvious that there are also nonarithmetical sets, but this follows from a simple counting argument. Every arithmetical set is defined by an $\mathcal{L}_2$-formula, of which there are countably many, while there are continuum-many subsets of $\mathbb{N}$.

We can extend the arithmetical hierarchy to the *analytical hierarchy* in much the same manner:

Definition 3.3.6 (analytical hierarchy for formulae)**.**

(i) An $\mathcal{L}_2$-formula φ is called Σ_0^1 and Π_0^1 if it is arithmetical.

(ii) φ is called Σ_{n+1}^1 if it is of the form $\varphi = \exists X_1 \cdots \exists X_n \ \psi$, where ψ is Π_n^1.

(iii) φ is called Π_{n+1}^1 if it is of the form $\varphi = \forall X_1 \cdots \forall X_n \ \psi$, where ψ is Σ_n^1.

So, the analytical formulae are those which allow some level of quantification over sets. Σ_n^1, Π_n^1 and Δ_n^1 sets of natural numbers are defined in exactly the same way. As before, we have that $\Sigma_n^1 \cup \Pi_n^1 \subseteq \Delta_{n+1}^1$ for all $n \in \mathbb{N}$.

Chapter 4

Subsystems of second-order arithmetic

With the tools of computability and model theory in hand, we can now develop the formalism of reverse mathematics. Here, we will define the subsystems RCA_0, WKL_0, ACA_0, ATR_0 and $\Pi^1_1\text{-}\mathsf{CA}_0$ and their ω-models, and see where famous theorems of mathematics show up in this hierarchy. Most of the material of this chapter can be found in [Sim09].

4.1 Formal systems

Definition 4.1.1. A *formal system* or *subsystem of second-order arithmetic* is a collection $\mathcal{S}$ of $\mathcal{L}_2$-sentences. We refer to the formulae in $\mathcal{S}$ as *axioms of* $\mathcal{S}$.

Definition 4.1.2. Let $\mathcal{M}$ be an ω-model, and $\mathcal{S}$ be a subsystem of second-order arithmetic. We say $\mathcal{M}$ *is a model of* $\mathcal{S}$ if all the axioms of $\mathcal{S}$ are true in $\mathcal{M}$.

There are infinitely many inequivalent subsystems of second-order arithmetic, but there are five major ones which show up consistently in reverse mathematics. In order of increasing logical strength, these systems are called RCA_0, WKL_0, ACA_0, ATR_0, and $\Pi^1_1\text{-}\mathsf{CA}_0$. These subsystems are affectionately known as the "Big Five"—their significance comes from the fact that almost all classical theorems turn out to be equivalent to one of the Big Five.

All these subsystems include the following set of basic axioms:

Axioms 4.1.3. The *basic axioms* of second-order arithmetic are the following $\mathcal{L}_2$-sentences:

(i) $\forall n\ \neg(n+1=0)$

(ii) $\forall n\ \forall m\ \big[(n+1=m+1)\rightarrow(n=m)\big]$

(iii) $\forall n\ (n+0=n)$

(iv) $\forall n\ \forall m\ \big[n+(m+1)=(n+m)+1\big]$

(v) $\forall n\ (n\cdot 0=0)$

(vi) $\forall n\ \forall m\ \big[n\cdot(m+1)=(n\cdot m)+n\big]$

(vii) $\forall n\ \neg(n<0)$

(viii) $\forall n\ \forall m\ \big[(n<m+1)\leftrightarrow(n<m\ \vee\ n=m)\big]$

Note that the basic axioms are entirely first-order—there is no mention of sets. They are closely related to the Peano axioms PA. The basic axioms formalise the essential properties of $\mathbb{N}$, and are sufficient to prove all basic facts of arithmetic—commutativity, associativity, distributivity, etc. It is easily verified that $\mathbb{N}$ satisfies the basic axioms, whence:

Proposition 4.1.4. Any ω-model satisfies the basic axioms.

The key feature distinguishing the different subsystems is the second-order axioms they contain. Most of the additional axioms we consider will have one of two forms. The first type are *induction axioms*, allowing us to induct over certain statements:

Definition 4.1.5. Let $\varphi(n)$ be an $\mathcal{L}_2$-formula in which n appears freely. The *induction axiom for* φ is the universal closure of $\left[\varphi(0) \wedge \forall n \left(\varphi(n) \rightarrow \varphi(n+1)\right)\right] \rightarrow \forall n\, \varphi(n)$.

The induction axiom for φ allows us to perform induction on φ. When defining subsystems $\mathcal{S}$ of second-order arithmetic, we will generally limit the inductive strength to some point in the arithmetical/analytical hierarchy. For example, $\mathcal{S}$ may include the induction axiom for all Π^0_2 formulae φ. We do this because we are trying to find the weakest subsystem in which a theorem φ is provable; therefore, we don't allow induction beyond what is truly necessary.

As we know, induction in $\mathbb{N}$ is valid for *any* $\mathcal{L}_2$-formula φ, whence:

Proposition 4.1.6. Any ω-model satisfies the induction axiom for any $\mathcal{L}_2$-formula φ.

The second type are *comprehension axioms*, guaranteeing that given sets must exist:

Definition 4.1.7. Let $\varphi(n)$ be an $\mathcal{L}_2$-formula in which n appears freely, but X does not appear. The *comprehension axiom for* φ is the universal closure of $\exists X\, \forall n\, \left[n \in X \leftrightarrow \varphi(n)\right]$.

Essentially, the comprehension axiom for φ asserts that the set $A_\varphi = \{n \in \mathbb{N} : \varphi(n)\}$ exists. Again, our subsystems will generally include all comprehension axioms up to some point in the arithmetical/analytical hierarchy.

4.2 RCA_0

Definition 4.2.1. RCA_0 is the subsystem consisting of:

(i) the basic axioms;

(ii) the induction axiom for every Σ^0_1 formula φ;

(iii) the Δ^0_1 *comprehension scheme*: the universal closure of

$$\left[\forall n\, \left(\varphi(n) \leftrightarrow \psi(n)\right)\right] \rightarrow \exists X\, \forall n\, \left(n \in X \leftrightarrow \varphi(n)\right)$$

for every Σ^0_1 formula $\varphi(x)$ and Π^0_1 formula $\varphi(x)$ containing x as a free variable, but not containing n or X.

RCA_0 stands for "recursive comprehension axiom", as it allows comprehension over Δ^0_1 sets (which in $\mathbb{N}$ are the computable sets, as we saw in Proposition 3.3.5). This is the weakest subsystem we will consider, and intuitively, it should be thought of as corresponding to computable mathematics. Generally, a statement holds in RCA_0 if and only if a "computable version" of the statement is true. Some results from ordinary mathematics do hold in RCA_0:

Proposition 4.2.2. The following theorems are provable in RCA_0:

(i) The Baire category theorem [Sim09];

(ii) The intermediate value theorem [PR89];

(iii) The soundness theorem for first order logic [Sim09];

(iv) Every countable, finite-rank matroid has a basis [HM17];

(v) The Weierstrass approximation theorem [PC75].

Proposition 4.2.2 boils down to the fact that each of these theorems is computably true. For example, the intermediate value theorem holds in RCA_0 because the effective IVT is true: if $f\colon [0,1] \rightarrow \mathbb{R}$ is computable and $f(0) < 0 < f(1)$, then there is a computable real number $x \in [0,1]$ such that $f(x) = 0$. But there are many more results that are not computably true, and thus don't hold in RCA_0:

Proposition 4.2.3. The following theorems are *not* provable in RCA_0:

(i) The Bolzano–Weierstrass theorem [Spe49];

(ii) The Heine–Borel theorem for countable covers [Fri76];

(iii) The extreme value theorem [Sim87; Sim09];

(iv) Gödel's completeness theorem for first order logic [Sim09];

(v) Every continuous functions is Riemann integrable [Sim09];

(vi) Every countable vector space has a basis [FSS83].

Again, interpret Proposition 4.2.3 as saying those theorems are not computably true. For example, the Bolzano–Weierstrass theorem does not hold in RCA_0, because the effective Bolzano–Weierstrass theorem fails—there is a computable Cauchy sequence whose limit is not computable [Spe49].

The standard ω-model of RCA_0 is REC, consisting of all recursive, or computable, subsets of $\mathbb{N}$. Given our intuition about RCA_0, it should be no surprise that REC actually *is* a model of RCA_0:

Proposition 4.2.4. REC is a model of RCA_0.

Proof. Since REC is an ω-model, it satisfies the basic axioms by Proposition 4.1.4, and all induction axioms by Proposition 4.1.6. Now, let φ be a Σ^0_1 formula, and ψ be Π^0_1. If φ and ψ are equivalent, then they both define the same set $A_\varphi = A_\psi \subseteq \mathbb{N}$. By Corollary 3.3.5, A_φ is computable, so $A_\varphi \in$ REC. Thus, A_φ witnesses Δ^0_1 comprehension for φ and ψ. $\square$

Recall our discussion of reverse mathematics from the introduction—a key idea was the *reversal*, where, to show that $\mathcal{S}$ is the weakest system in which φ can be proved, we demonstrate a proof of $\mathcal{S}$ from φ. No single theorem is strong enough to axiomatise mathematics; hence, when doing a reversal in practice, we need to supplement φ with a weak base theory $\mathcal{B}$. It is customary to take $\mathcal{B} = \mathsf{RCA}_0$ (though weaker/stronger systems have been used at times).

4.3 WKL_0

Kőnig's lemma is a statement about infinite well-founded trees, and *weak Kőnig's lemma* is the restriction of this to binary trees, i.e. those where each node has at most two children. It is convenient to define a tree as a certain subset of the following:

Definition 4.3.1. *(Finitary) Cantor space* $2^{<\omega}$ is the set of all finite binary sequences $\sigma = \sigma_0\sigma_1\cdots\sigma_{n-1}$, where each $\sigma_i \in \{0,1\}$. *(Infinitary) Cantor space* 2^ω consists of all infinite binary sequences $X = X_0X_1X_2\cdots$, where each $X_i \in \{0,1\}$.

$2^{<\omega}$ is countable, so we can represent its elements in $\mathbb{N}$ as follows: for $n \in \mathbb{N}$, write $n+1$ in binary and remove the leading 1. For example, 22 in binary is 10110, hence 21 represents the string $0110 \in 2^{<\omega}$. This is a bijection between $\mathbb{N}$ and $2^{<\omega}$ (0 represents the empty string). Henceforth, when we talk about elements $\sigma \in 2^{<\omega}$ in second-order arithmetic, they will be understood as natural numbers via this coding.

However, 2^ω is uncountable, so it cannot be represented in $\mathbb{N}$. Instead, we need to represent 2^ω using subsets of $\mathbb{N}$. The obvious way is to represent $X = X_0X_1X_2\cdots$ by the set $A_X = \{n \in \mathbb{N} : X_n = 1\}$, so that X is essentially the characteristic function of A_X. Then, given an ω-model $\mathcal{M}$, an element $X \in 2^\omega$ exists in $\mathcal{M}$ if and only if $A_X \in \mathcal{M}$.

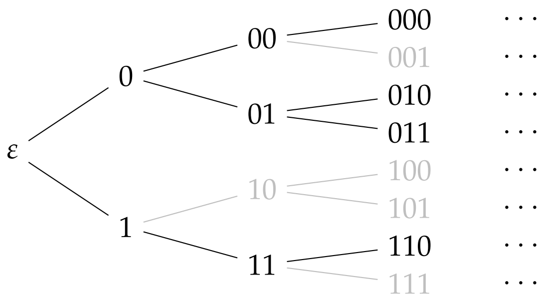

Figure 4.1: A tree $T \subseteq 2^{<\omega}$ as a subgraph of the full binary tree $2^{<\omega}$.

Definition 4.3.2. For an element $\sigma = \sigma_0\sigma_1 \cdots \sigma_{n-1} \in 2^{<\omega}$, its *length* is $|\sigma| = n$. Given $\sigma \in 2^{<\omega}$ and $\tau \in (2^{<\omega} \cup 2^{\omega})$, we say τ *extends* σ (notated $\sigma \preccurlyeq \tau$) if $|\sigma| \leq |\tau|$ and for all $k < |\sigma|$, $\sigma_k = \tau_k$.

Definition 4.3.3. A tree is a subset $T \subseteq 2^{<\omega}$ which is closed under initial segments. That is, if $\sigma = \sigma_0\sigma_1 \cdots \sigma_{n-1} \in T$, then $\sigma\restriction_k := \sigma_0\sigma_1 \cdots \sigma_{k-1} \in T$ for any $k \leq n$.

As in Figure 4.1, we intuitively think of a tree $T \subseteq 2^{\omega}$ as a certain graph, where the vertices are the elements, and $\sigma \in T$ is connected to its direct extensions $\sigma^\frown 0$ and $\sigma^\frown 1$, if they are in T. Here, $^\frown$ denotes concatenation of strings. Every nonempty tree contains the empty sequence ε; this is the "root" of the tree.

Definition 4.3.4. Given a tree $T \subseteq 2^{<\omega}$, a *path through* T is an element $X \in 2^{\omega}$ such that $X\restriction_n := X_0X_1 \cdots X_{n-1} \in T$ for any $n \in \mathbb{N}$.

Definition 4.3.5. *Weak Kőnig's lemma* is the statement that every infinite tree $T \subseteq 2^{<\omega}$ contains a path. WKL_0 is the subsystem consisting of RCA_0 plus weak Kőnig's lemma.

By definition, WKL_0 is at least as strong as RCA_0 in terms of logical strength. We can show WKL_0 is strictly stronger. The idea is to show the standard model REC of RCA_0 is *not* a model of WKL_0, by constructing a computable tree $T \subseteq 2^{<\omega}$ in REC with no computable path $X \in$ REC. This result is originally due to Jockusch and Soare [JS72], though our construction is different to theirs. We use a typical *diagonalisation argument* ubiquitous in computability; we ensure at the eth step that φ_e is not a branch through T.

Proposition 4.3.6 [JS72]**.** There is a computable tree $T \subseteq 2^{<\omega}$ with no computable path.

Proof. Recall $\varphi_0, \varphi_1, \varphi_2, \ldots$ is a uniformly computable listing of all partial computable functions (Proposition 3.1.10). Construct a tree T as follows. To test if $\sigma \in T$, for every $e < |\sigma|$, run the computation of $\varphi_e(e)$ for $|\sigma|$ steps. If any of these computations halt with $\varphi_e(e) = \sigma_e$, then $\sigma \notin T$; otherwise, $\sigma \in T$.

T is a tree: instead of showing T is downwards closed, we (equivalently) show the complement is upwards closed. Suppose that $\sigma \notin T$—then, there is $e < |\sigma|$ such that $\varphi_e(e)$ halts within $|\sigma|$ steps, and $\varphi_e(e) = \sigma_e$. Then, for any extension $\tau \succcurlyeq \sigma$, $\varphi_e(e)$ also halts within $|\tau| \geq |\sigma|$ steps, and $\varphi_e(e) = \sigma_e = \tau_e$. Hence, $\tau \notin T$.

T is infinite, since every level n is nonempty. To see this, for each $e < n$ such that $\varphi_e(e)$ is defined, we can pick $\sigma_e \neq \varphi_e(e)$. If $\varphi_e(e)$ is undefined, just pick σ_e arbitrarily. Then $\sigma \in T$ is on level n.

T is computable, since we gave an algorithm to compute it. Hence, T exists in REC. But we claim T has no path in REC. Suppose T did have a path $X \in$ REC: then, $X = \varphi_e$ for some e, since the sequence (φ_n) lists *all* partial computable functions. Since X is total, the computation of $\varphi_e(e) = X_e$ halts after, say, s steps. But then $X\restriction_s \notin T$ by definition, so X is not a path through T. $\square$

Since we have constructed an infinite tree $T \in$ REC with no path $X \in$ REC, it follows that weak Kőnig's lemma does not hold in REC, whence:

Corollary 4.3.7. REC is not a model of WKL_0.

Weak Kőnig's lemma is closely related to the finite intersection characterisation of compactness; indeed, it can be viewed as asserting that Cantor space 2^ω, the infinite product of the discrete space 2, is compact. Therefore, WKL_0 is generally strong enough to perform compactness arguments. WKL_0 can prove all the results of Proposition 4.2.2, and it is equivalent to the following results (in other words, WKL_0 is the weakest system in which they can be proved):

Proposition 4.3.8. Over RCA_0, WKL_0 is equivalent to:

(i) The Heine–Borel theorem for countable covers [Fri76];

(ii) Continuous functions on $[0, 1]$ are bounded [Sim87; Sim09];

(iii) Continuous functions on $[0, 1]$ are uniformly continuous [Sim87; Sim09];

(iv) Continuous functions on $[0, 1]$ are Riemann integrable [Sim09];

(v) The extreme value theorem [Sim87; Sim09];

(vi) Gödel's completeness theorem for first order logic [Sim09];

(vii) Every countable commutative ring has a prime ideal [FSS83];

(viii) Brouwer's fixed point theorem [ST90];

(ix) The Hahn–Banach theorem for separable Banach spaces [BS86].

That said, WKL_0 is still insufficient to prove many important mathematical results, such as the completeness of $\mathbb{R}$, the Bolzano–Weierstrass theorem, and the existence of bases for vector spaces. We will see below that each of these statements is equivalent to the stronger system ACA_0.

4.4 ACA_0

Now, we move on to stronger subsystems of second-order arithmetic. The first is ACA_0 (arithmetical comprehension), which guarantees the existence of any arithmetical set:

Definition 4.4.1. ACA_0 is the subsystem consisting of:

(i) the basic axioms;

(ii) the induction axiom for every arithmetical formula φ;

(iii) the comprehension axiom for every arithmetical formula φ.

ACA_0 has a standard ω-model ARITH, consisting of all arithmetical subsets of $\mathbb{N}$.

Proposition 4.4.2. ARITH is a model of ACA_0.

Proof. Very similar to Proposition 4.2.4, so omitted. □

Clearly ACA_0 implies RCA_0, and therefore all of the results of Proposition 4.2.2. It is less obvious that ACA_0 implies WKL_0, and thereby the results of Proposition 4.3.8. In fact, ACA_0 is strong enough to prove almost all the results of classical mathematics (algebra, analysis, etc.), and virtually all the theorems taught in undergraduate mathematics.

Proposition 4.4.3. Over RCA_0, ACA_0 is equivalent to:

(i) The sequential completeness of the reals [Fri76];

(ii) The Bolzano–Weierstrass theorem [Fri76];

(iii) Every countable commutative ring has a maximal ideal [FSS83];

(iv) Every countable vector space has a basis [FSS83];

(v) Kőnig's lemma: every infinite, finitely branching tree has an infinite path [Fri74; Fri76];

(vi) Ramsey's theorem for k-tuples, for fixed $k \geq 3$ [Joc72; Sim09].

There are a few mathematical theorems still out of reach for ACA_0; for instance, in areas such as set theory, Ramsey theory and descriptive set theory, where strong set existence axioms are required. We will see some examples in the next two sections.

4.5 ATR_0

ATR_0 (arithmetical transfinite recursion) comprises ACA_0 plus the assertion that any "arithmetical operator" can be iterated along any countable ordinal, starting with any set. Let's try to understand what all of this means.

Definition 4.5.1. A (countable) ordinal is a set $\alpha \subseteq \mathbb{N}$ with a linear order $<_\alpha$ that is *well-founded*: there is no infinite descending sequence $a_0 >_\alpha a_1 >_\alpha a_2 >_\alpha \cdots$ in α.

Ordinals are important in mathematics because we can induct on them: if $\varphi(\alpha)$ is a statement about ordinals, such that $\varphi(0)$ holds, and $\varphi(\beta)$ for all $\beta < \gamma$ implies $\varphi(\gamma)$, then $\varphi(\alpha)$ holds for *all* ordinals α. One classic and important result of set theory is that any two ordinals are *comparable*: either they are isomorphic, or one is isomorphic to a strict initial segment of the other. ACA_0 is not strong enough to prove this result, which is part of the motivation for introducing ATR_0.

Now, let $\theta(n, X)$ be an arithmetical formula, with one free number variable n and one free set variable X. θ defines an "arithmetical operator" $\Theta\colon \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ by

$$\Theta(X) = \{n \in \mathbb{N} : \theta(n, X) \text{ holds}\}$$

For a set $Y \subseteq \mathbb{N} \times \alpha$, for each $\beta \in \alpha$, we let $Y^{[\beta]} = \{n \in \mathbb{N} : (n, \beta) \in Y\}$ be the βth column of Y, and $Y^{[<\beta]} = \{(n, \gamma) \in Y : \gamma <_\alpha \beta\}$ be all the columns up to β.

Definition 4.5.2. For a countable ordinal $\alpha \subseteq \mathbb{N}$ and set $X \subseteq \mathbb{N}$, let $\Theta^\alpha(X)$ be the subset[1] $Y \subseteq \mathbb{N} \times \alpha$ such that $Y^{[0]} = X$ and $Y^{[\beta]} = \Theta\big(Y^{[<\beta]}\big)$. Then, ATR_0 is the subsystem consisting of ACA_0, plus the assertion that $\Theta^\alpha(X)$ exists, for every arithmetical operator Θ, countable ordinal α, and set $X \subseteq \mathbb{N}$.

Of course, ATR_0 implies ACA_0, and thereby all the results of the previous section. Generally, ATR_0 is needed to prove theorems which use ordinals in an essential way. An example is *Ulm's theorem* on countable abelian p-groups: we assign each such group a countable ordinal α, and a sequence in $\mathbb{N} \cup \{\infty\}$ of length α called its *Ulm invariant*. Ulm's theorem states that two countable abelian p-groups are isomorphic if and only if they have the same Ulm invariant; this theorem is equivalent to ATR_0 [FSS83].

Proposition 4.5.3. Over RCA_0, ATR_0 is equivalent to:

[1] ACA_0 is required to prove the set $\Theta^\alpha(X)$ is uniquely defined.

(i) Any two countable ordinals are comparable [Ste77];

(ii) Every uncountable closed set in $\mathbb{R}$ has a perfect subset [Fri74; Fri76];

(iii) Determinacy for open or clopen sets in $\mathbb{N}^{\mathbb{N}}$ [Ste77];

(iv) Ramsey's theorem for open or clopen sets in $\mathbb{N}^{\mathbb{N}}$ [FMS82].

4.6 Π^1_1-CA_0

The strongest system we will discuss is Π^1_1-CA_0, guaranteeing the existence of any Π^1_1 set:

Definition 4.6.1. Π^1_1-CA_0 is the subsystem consisting of:

(i) the basic axioms;

(ii) the induction axiom for every Π^1_1 formula φ;

(iii) the comprehension axiom for every Π^1_1 formula φ.

In terms of subsystems of second-order arithmetic, Π^1_1-CA_0 is "way up in the stratosphere"; it can prove almost any mathematical theorem the reader can imagine. Π^1_1-CA_0 implies ATR_0, and hence all the results of the previous sections. Here are some further results equivalent to Π^1_1-CA_0; thus, they require some level of quantification over sets:

Proposition 4.6.2. Over RCA_0, Π^1_1-CA_0 is equivalent to:

(i) The Cantor–Bendixson theorem [Fri76];

(ii) Every countable abelian group is the direct sum of a divisible group and a reduced group [FSS83];

(iii) Determinacy for sets of the form $U \setminus U'$, where U, U' open in $\mathbb{N}^{\mathbb{N}}$ [Tan91];

(iv) Ramsey's theorem for Δ^0_2 sets in $\mathbb{N}^{\mathbb{N}}$ [Sol78; Sim09];

(v) The minimal bad sequence lemma [Mar96];

(vi) Maltsev's theorem: every countable ordered group has order type $\mathbb{Z}^\alpha$ or $\mathbb{Z}^\alpha\mathbb{Q}$ [Sol01];

(vii) Every countable ring has a prime radical [Con09].

As strong as Π^1_1-CA_0 is, there are still a few results which manage to escape it. These are generally restricted to select theorems in infinitary Ramsey theory, WQO theory, and set theory, where Π^1_2 or Π^1_3 comprehension might be required.

Chapter 5

Analysis in second-order arithmetic

As described in Section 3.2, second-order arithmetic only includes two types of objects: natural numbers $n, m, k, \ldots$ and sets $A, B, C, \ldots$ thereof. Therefore, any other objects which we want to discuss must be *coded* using natural numbers or subsets of $\mathbb{N}$. We've already seen an example in Section 4.3—coding finite binary strings $\sigma \in 2^{<\omega}$ by natural numbers, and infinite binary strings $X \in 2^{\omega}$ using sets. In this section, we code the basic number systems $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ in second-order arithmetic, which then allows us to formalise basic concepts of analysis.

5.1 Number systems

The smallest number system is $\mathbb{N}$, and this is already given in second-order arithmetic, as the collection of all objects of number type. To code the integers $\mathbb{Z}$, we imitate the usual construction of $\mathbb{Z}$ from $\mathbb{N}$, where we use a pair $(a, b) \in \mathbb{N} \times \mathbb{N}$ to represent $a - b \in \mathbb{Z}$, and then quotient $\mathbb{N} \times \mathbb{N}$ by a suitable equivalence relation. To perform this construction, we first need to code pairs of natural numbers.

Definition 5.1.1. For $m, n \in \mathbb{N}$, define the *pair* (m, n) as the natural number $(m+n)^2 + m$.

The reason we use this pairing function, rather than Cantor's pairing function of Definition 3.1.5, is that the definition is more elementary, not requiring division, and thus easier to reason about. It has the disadvantage of not being a bijection.

Definition 5.1.2 [Sim09, §II.4]**.** For pairs (m, n), $(p, q) \in \mathbb{N}^2$, say $(m, n) =_{\mathbb{Z}} (p, q)$ if $m + q = n + p$. Then, an *integer* is a pair (m, n) which is minimal in its $=_{\mathbb{Z}}$-equivalence class.

Again, the pair (m, n) should be interpreted as the integer $m - n$. Instead of taking the equivalence classes as objects, we instead take minimal elements, as this way, integers can be represented by single natural numbers, rather than sets thereof. We can also define the standard arithmetic operations $+_{\mathbb{Z}}$, $-_{\mathbb{Z}}$, $\cdot_{\mathbb{Z}}$ and ordering $<_{\mathbb{Z}}$ on integers in the evident way. For example, $(m, n) -_{\mathbb{Z}} (p, q)$ is the pair $(m + q, n + p)$.

Having defined the integers $\mathbb{Z}$, we can now define the rationals $\mathbb{Q}$ from $\mathbb{Z}$, via the usual field of fractions construction:

Definition 5.1.3. Let $\mathbb{Z}^+ = \{x \in \mathbb{Z} : x >_{\mathbb{Z}} 0_{\mathbb{Z}}\}$. For pairs (a, b), $(c, d) \in \mathbb{Z} \times \mathbb{Z}^+$, say $(a, b) =_{\mathbb{Q}} (c, d)$ if $a \cdot_{\mathbb{Z}} d =_{\mathbb{Z}} b \cdot_{\mathbb{Z}} c$. Then, a *rational number* is a pair (a, b) which is minimal in its $=_{\mathbb{Q}}$-equivalence class.

Here, we interpret the pair (a, b) as the rational number a/b. Again, the standard operations and relations $+_{\mathbb{Q}}$, $-_{\mathbb{Q}}$, $\cdot_{\mathbb{Q}}$, $<_{\mathbb{Q}}$, $|\cdot|_{\mathbb{Q}}$ are defined as expected.

Moving to the real numbers $\mathbb{R}$, we have to change strategy, as we are now moving from countable to uncountable. It will not be possible to define real numbers as pairs of rationals, or even finite sequences of rationals, as there are too many reals. Instead, we define real numbers as certain *infinite* sequences of rationals, mirroring the familiar Cauchy construction of $\mathbb{R}$ from $\mathbb{Q}$, with a small twist.

Definition 5.1.4. Given sets $X, Y \subseteq \mathbb{N}$, a *function* $f\colon X \to Y$ is a set of pairs $(x, y) \in X \times Y$, such that for all $x \in X$, there is a unique $y \in Y$ with $(x, y) \in f$.

Definition 5.1.5. A *sequence of rationals* is a function $f\colon \mathbb{N} \to \mathbb{Q}$. By convention, we will instead denote sequences by (q_i), where $q_i = f(i)$. A *real number* is a sequence of rationals (q_i) such that for all $m \leq n \in \mathbb{N}$, $|q_m - q_n| \leq 2^{-m}$. We say $(q_i) =_{\mathbb{R}} (q'_i)$ if for all k, $|q_k - q'_k| \leq 2^{-k+1}$.

One may wonder why we require $|q_m - q_n| \leq 2^{-m}$, and not just the usual Cauchy condition: for all $\varepsilon \in \mathbb{Q}^+$, there is N such that for all $m, n \geq N$, $|q_m - q_n| \leq \varepsilon$. The reason is that Definition 5.1.5 is modelled on the definition of a *computable* real number—hence, it is the more suitable definition in weak systems such as RCA_0 and WKL_0. ACA_0 is needed to prove the equivalence between Definition 5.1.5 and the usual Cauchy definition.

Also note that we are not picking a representative from each $=_{\mathbb{R}}$-class—this would require strong comprehension/choice axioms which we may not have access to. The standard arithmetic operations $+_{\mathbb{R}}, -_{\mathbb{R}}, \cdot_{\mathbb{R}}, |\cdot|_{\mathbb{R}}$ can be defined "pointwise", and we say $(q_i) \leq_{\mathbb{R}} (q'_i)$ if for all k, $q_k \leq q'_k + 2^{-k+1}$. Furthermore, any rational q can be identified with the real number $r_q = (q, q, q, \ldots)$.

Definition 5.1.6. Given a real number r, we say $r \in [0,1]$ if $0 \leq_{\mathbb{R}} r \leq_{\mathbb{R}} 1$.

5.2 Open sets

A key topological property of $\mathbb{R}$ is that it is *second-countable*, i.e. its topology has a countable basis, consisting of open intervals (p, q) with rational endpoints. This property is essential in allowing us to code open sets of $\mathbb{R}$ in second-order arithmetic. We first use natural numbers to code a basis of rational intervals for $\mathbb{R}$ and $[0,1]$:

Definition 5.2.1.

(i) A pair $(p, q) \in \mathbb{Q} \times \mathbb{Q}$, where $p < q$, codes the open interval $V_{p,q} := (p, q) \subseteq \mathbb{R}$.

(ii) For a real number r, we say $r \in V_{p,q}$ if $p < r < q$.

(iii) We say $V_{p,q} \cap V_{p',q'} \neq \varnothing$ if $(p < q') \wedge (p' < q)$.

(iv) We say $V_{p,q} \subseteq V_{p',q'}$ if $(p \geq p') \wedge (q \leq q')$.

(v) The *length* of $V_{p,q}$ is $\ell(V_{p,q}) := q - p$.

(vi) $\mathcal{B}_{\mathbb{R}} \subseteq \mathbb{Q} \times \mathbb{Q}$ denotes the set of all such intervals.

Definition 5.2.2.

(i) Given $p, q \in \mathbb{Q}$, we define $\overline{p} := \max\{p, 0\}$ and $\overline{q} := \min\{q, 1\}$.

(ii) We also use (p, q), where $\overline{p} < \overline{q}$, to code the open[1] interval $U_{p,q} := (p, q) \cap [0,1]$.

(iii) For a real number r, we say $r \in U_{p,q}$ if $r \in [0,1]$ and $p < r < q$.

[1] In $[0,1]$ with the standard subspace topology.

(iv) We say $U_{p,q} \cap U_{p',q'} \neq \varnothing$ if $(\overline{p} < \overline{q}') \wedge (\overline{p}' < \overline{q})$.

(v) We say $U_{p',q'} \subseteq U_{p,q}$ if $(p < 0 \vee p \leq p') \wedge (q > 1 \vee q \geq q')$.

(vi) The *length* of $U_{p,q}$ is $\ell(U_{p,q}) := \max\{\overline{q} - \overline{p}, 0\}$.

(vii) $\mathcal{B}_{[0,1]} := \mathbb{Q} \times \mathbb{Q}$ denotes the set of all such intervals.

Having coded the basis elements into the model, we can now define arbitrary open sets:

Definition 5.2.3. An *open set* $O \subseteq \mathbb{R}$ is a sequence (U_i) of open intervals $U_i \in \mathcal{B}_{\mathbb{R}}$, i.e. a function $f\colon \mathbb{N} \to \mathcal{B}_{\mathbb{R}}$.

The sequence (U_i) should be interpreted as the open set $O = \bigcup_{i \in \mathbb{N}} U_i$. We will use the same sequence to code the closed set $C = \mathbb{R} \setminus O$. Relatively open and closed sets in $[0,1]$ are defined the same way, starting from $\mathcal{B}_{[0,1]}$.

5.3 The proof of Cousin's lemma, revisited

Let's again look at the proof of Cousin's lemma (Lemma 2.3.1), and attempt to formalise it in second-order arithmetic. We have not yet given a formal definition of a *function* $f\colon [0,1] \to \mathbb{R}$ in second-order arithmetic; for now, let us take it to be a primitive, undefined notion.

Definition 5.3.1. A function $f\colon [0,1] \to \mathbb{R}$ is a *gauge* if for all $x \in [0,1]$, we have $f(x) > 0$. This property will be denoted $f\colon [0,1] \to \mathbb{R}^+$.

Definition 5.3.2. A *tagged partition of* $[0,1]$ is a finite, odd-length sequence of reals

$$P = \langle x_0, t_0, x_1, t_1, \ldots, x_{\ell-1}, t_{\ell-1}, x_\ell \rangle \ \subseteq \ \mathbb{R}^{2\ell+1}$$

such that $x_0 = 0$, $x_\ell = 1$, and for all $j < \ell$, we have $x_j < t_j < x_{j+1}$. The number ℓ is called the *size of* P.

Definition 5.3.3. Let $\delta\colon [0,1] \to \mathbb{R}^+$ be a gauge, and P be a tagged partition of size ℓ. Then, we say P is δ*-fine* if for all $j < \ell$, $t_j - \delta(t_j) \leq x_j$ and $t_j + \delta(t_j) \geq x_{j+1}$.

We can now conduct the proof of Lemma 2.3.1 in second-order arithmetic.

Theorem 5.3.4 (Π^1_1-CA_0)**.** Any gauge $\delta\colon [0,1] \to \mathbb{R}^+$ has a δ-fine partition.

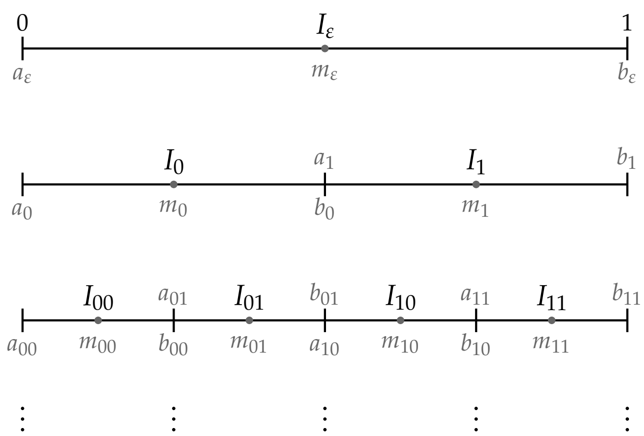

Figure 5.1: The definition of a_σ, b_σ, I_σ, m_σ in the proof of Theorem 5.3.4; ε is the empty string.

"Proof". For each $\sigma \in 2^\omega$, we define rationals $a_\sigma := \sum_{i<n} \sigma_i \cdot 2^{-i-1}$ and $b_\sigma := a_\sigma + 2^{|\sigma|}$, and the interval $I_\sigma = (a_\sigma, b_\sigma)$. For convenience, we also let $m_\sigma = (a_\sigma + b_\sigma)/2$, the midpoint of I_σ. So the strings σ of length n partition $[0,1]$ into 2^n subintervals I_σ of equal length 2^{-n}, as shown in Figure 5.1.

We define a tree T in levels as follows. For each $n \in \mathbb{N}$, inductively define

$$T_n := \left\{ \sigma \in 2^{<\omega} : |\sigma| = n, \forall k < n \, (\sigma\restriction_k \in T_k), \forall r \in I_\sigma \left(\delta(r) \leq 2^{-n} \right) \right\}$$

Each T_n exists by Π^1_1 comprehension. Take $T = \bigcup_{n \in \mathbb{N}} T_n$. Then, $T \subseteq 2^{<\omega}$ is downward closed by construction, hence a tree.

If T is finite, then T defines a δ-fine partition P_T of $[0,1]$ as follows: let $(\sigma^{(0)}, \ldots, \sigma^{(n-1)})$ be a lexicographically sorted list[2] of all the $\sigma \notin T$ such that $\sigma\restriction_k \in T$ for all $k < |\sigma|$. $P_T = \langle x_0 < t_0 < \ldots < x_n \rangle$ is defined by letting $x_i = a_{\sigma^{(i)}} = b_{\sigma^{(i-1)}}$, and $t_i = m_{\sigma^{(i)}}$.

Now, we claim that T must be finite. If not, then WKL_0 proves there is an infinite path $X = X_0 X_1 \cdots$ through T. Define the real $r_X = (q_n)_{n \in \mathbb{N}}$, where each $q_n = m_{X\restriction_n}$. Note that $r_X \in I_{X\restriction_n}$ for every n. Hence, by the definition of T, $\delta(r_X) \leq 2^{-n}$ for every n, whence $\delta(r_X) = 0$, contradicting the fact that δ is a gauge. □

Π^1_1-CA_0 was required when defining the T_n; we used universal quantification over real numbers. The reason this is a "proof", and not a proof, is that second-order arithmetic is unable to talk about *arbitrary* functions $f\colon [0,1] \to \mathbb{R}$; these require uncountably much information to specify (i.e. where f sends each point in $[0,1]$). Furthermore, there are $2^{\mathfrak{c}}$-many functions $[0,1] \to \mathbb{R}$; too many to code even using subsets of $\mathbb{N}$!

However, we will see in Sections 6 and 7 that second-order arithmetic can describe certain types of functions $f\colon [0,1] \to \mathbb{R}$. Essentially, we can formalise any class of functions that can be specified by countable information—examples include continuous functions, functions of a given Baire class, and Borel functions. The above "proof" shows that Π^1_1-CA_0 is an upper bound on the axiomatic strength of Cousin's lemma, for *any* class of functions which can be defined in second-order arithmetic. We will see that this upper bound is often far from tight; in particular, Cousin's lemma for continuous functions can be proven in WKL_0, a much weaker system than Π^1_1-CA_0.

[2] There are finitely many such σ if T is finite.

Chapter 6

Cousin's lemma for continuous functions

Having outlined the main concepts of reverse mathematics in Chapter 4, and begun to formalise analysis in Chapter 5, we are now ready for a formal discussion of Cousin's lemma in second-order arithmetic. First, we will define continuous functions in second-order arithmetic, then determine the axiomatic strength of $\mathsf{CL_c}$, Cousin's lemma for continuous functions. We will construct an explicit example showing that $\mathsf{CL_c}$ fails in $\mathsf{RCA_0}$, and then prove the equivalence between $\mathsf{CL_c}$ and $\mathsf{WKL_0}$ over $\mathsf{RCA_0}$.

6.1 Continuous functions

Since we wish to do real analysis in second-order arithmetic, we would now hope to be able to code *functions* $[0,1] \to \mathbb{R}$. But, as discussed in the previous section, second-order arithmetic cannot describe *arbitrary* functions $f\colon [0,1] \to \mathbb{R}$; for this, we would require *third-order* arithmetic. However, certain types of functions $f\colon \mathbb{R} \to \mathbb{R}$ can be. It is known that any continuous function $f\colon A \to B$, with A separable and B Hausdorff, is uniquely determined by its values on a countable dense subset of A. As $[0,1]$ is separable and $\mathbb{R}$ is Hausdorff, this means continuous functions $f\colon [0,1] \to \mathbb{R}$ can be specified by countable information, and thus coded by subsets of $\mathbb{N}$.

We will only be concerned with continuous functions $f\colon [0,1] \to \mathbb{R}$, but exactly the same idea can be used to code continuous functions $f\colon \mathbb{R} \to \mathbb{R}$. Our method will be to code $f\colon [0,1] \to \mathbb{R}$ by the collection of *pairs* of rational open intervals (U, V) such that $f(U) \subseteq V$.

Definition 6.1.1 [Sim09, Defn II.6.1]**.** A *(partial) continuous function* $f\colon [0,1] \to \mathbb{R}$ is a subset $f \subseteq \mathcal{B}_{[0,1]} \times \mathcal{B}_{\mathbb{R}}$ satisfying the following:

(i) $(U, V) \in f$ and $(U, V') \in f \implies V \cap V' \neq \varnothing$;

(ii) $(U, V) \in f$ and $U' \subseteq U \implies (U', V) \in f$;

(iii) $(U, V) \in f$ and $V \subseteq V' \implies (U, V') \in f$;

Again, we should interpret $(U, V) \in f$ (in the formal sense) to mean "$f(U) \subseteq V$" (in the colloquial sense). Such functions may be partial because they may not give us enough information to define $f(x)$ at a point $x \in [0,1]$. For example, the collection $g = \{(U, V) : U \in \mathcal{B}_{[0,1]},\ V \supseteq (0,1)\}$ meets the conditions of Definition 6.1.1, but for any real $x \in [0,1]$, we only know that $g(x) \in (0,1)$; g does not give us enough information to localise $g(x)$ more than this.

Definition 6.1.2. Let $f\colon [0,1] \to \mathbb{R}$ be a partial continuous function. A real $x \in [0,1]$ is *in the domain of* f if, for all $\varepsilon \in \mathbb{Q}^+$, there is a pair $(U,V) \in f$ such that $x \in U$ and $\ell(V) \leq \varepsilon$. If all $x \in [0,1]$ are in the domain of f, we say f is *total*.

If x is in the domain of f, we define $f(x)$ as the real $(q_n)_{n\in\mathbb{N}}$ obtained as follows. For each n, let $\varepsilon = 2^{-n}$, and for the least $V = V_{p,q}$ witnessing the above, let $q_n = (p+q)/2$. Using the assumptions in Definition 6.1.1, we can verify that (q_n) satisfies Definition 5.1.5.

Let's formalise some basic examples of continuous functions in second-order arithmetic, and check that they satisfy Definition 6.1.1.

Proposition 6.1.3 (RCA_0)**.** For any $m, c \in \mathbb{Q}$, the linear function $f\colon x \mapsto mx + c$ is total continuous.

Proof. If $m = 0$, this is simply the constant function $x \mapsto c$. Then $f = \{(U,V) : U \in \mathcal{B}_{[0,1]},\ c \in V\}$ is continuous, total, and has $f(x) =_{\mathbb{R}} c$ for all $x \in [0,1]$.

Now, suppose $m > 0$. Define $f\colon [0,1] \to \mathbb{R}$ by letting $(U_{p,q}, V_{r,s}) \in f$ if and only if $m\overline{p} + c > r$ and $m\overline{q} + c < s$. We verify Definition 6.1.1:

(i) If $(U_{p,q}, V_{r,s}), (U_{p,q}, V_{r',s'}) \in f$, then $r < m\overline{p} + c < m\overline{q} + c < s'$ since $m > 0$ and $\overline{p} < \overline{q}$. We have $r' < s$ similarly, hence $V_{r,s} \cap V_{r',s'} \neq \varnothing$.

(ii) Suppose $(U_{p,q}, V_{r,s}) \in f$ and $U_{p',q'} \subseteq U_{p,q}$. Then, either $p < 0$, in which case $\overline{p} = 0 \leq \overline{p}'$ or $p \leq p'$, in which case $\overline{p} \leq \overline{p}'$ also. Either way, we get $m\overline{p}' + c \geq m\overline{p} + c > r$. It follows similarly that $m\overline{q}' + c < s$, hence $(U_{p',q'}, V_{r,s}) \in f$ as required.

(iii) If $(U_{p,q}, V_{r,s}) \in f$ and $V_{r,s} \subseteq V_{r',s'}$, then $m\overline{p} + c > r \geq r'$ and $m\overline{q} + c < s \leq s'$. It follows that $(U_{p,q}, V_{r',s'}) \in f$.

(iv) f total: pick a real $x = (x_n)_{n\in\mathbb{N}} \in [0,1]$ and $\varepsilon \in \mathbb{Q}^+$. Let n be least such that $2^{-n} < \varepsilon/2m$. Let $p := x_n - 2^{-n}$, $q := x_n + 2^{-n}$, $r := mx_n + c - \varepsilon/2$ and $s := mx_n + c + \varepsilon/2$. Then, the pair $(U_{p,q}, V_{r,s}) \in f$ is as required.

If $m < 0$, then we let $(U_{p,q}, V_{r,s}) \in f$ if and only if $m\overline{q} + c > r$ and $m\overline{p} + c < s$, and the proof follows similarly. □

Within RCA_0, we also have some methods of constructing new continuous functions from old:

Lemma 6.1.4 (RCA_0)**.** If $\sum_{n=0}^{\infty} \alpha_n$ is a convergent series of nonnegative real numbers, and $(f_n)_{n\in\mathbb{N}}$ is a sequence of continuous functions $[0,1] \to \mathbb{R}$ such that $|f_n(x)| \leq \alpha_n$ for all $x \in [0,1]$, $n \in \mathbb{N}$, then $f = \sum_{n=0}^{\infty} f_n$ is continuous. Furthermore, f is total if all the f_n are total.

Proof. [Sim09, Lemma II.6.5]. □

Lemma 6.1.5 (RCA_0)**.** Let $0 = d_0 < d_1 < d_2 < \cdots < d_k = 1$ be a finite, increasing sequence of rationals, and let $f_1, \ldots, f_k\colon [0,1] \to \mathbb{R}$ be continuous such that $f_i(d_i) = f_{i+1}(d_i)$ whenever $0 < i < k$. Then, the piecewise function f defined $f(x) = f_i(x)$ for $d_{i-1} \leq x \leq d_i$ is also continuous. Furthermore, f is total if all of the f_i are total.

Proof. Formally, we construct the code for f in RCA_0, as per Definition 6.1.1. Let $(U,V) \in f$ if and only if $(U,V) \in f_i$ for all i such that $[d_{i-1}, d_i]$ intersects U. One can easily prove that f is partial continuous.

Now, suppose each f_i is total. To prove f is total, pick some $x \in [0,1]$ and $\varepsilon \in \mathbb{Q}^+$. There are three cases:

Case 1: $x \neq d_i$ for any $i \leq k$. Then, there is a unique i such that $d_{i-1} < x < d_i$. Pick $(U, V) \in f_i$ witnessing that f_i is total for this x and ε, i.e. $\ell(V) \leq \varepsilon$. If $U = U_{p,q}$, then let $p' := \max\{p, d_{i-1}\}$, $q' := \max\{q, d_i\}$, and $U' := U_{p',q'}$. Then $U' \subseteq U$, so by Definition 6.1.1.(ii), $(U', V) \in f_i$ also, and $x \in U'$. Since U' only intersects $[d_{i-1}, d_i]$, it follows that $(U', V) \in f$ is as required.

Case 2: $x = d_i$ for $0 < i < k$. We pick $(U_0, V_0) \in f_i$ and $(U_1, V_1) \in f_{i+1}$ witnessing that f_i (resp. f_{i+1}) is total for this x and $\varepsilon/2$. $U := U_0 \cap U_1$ is a nonempty interval since $x \in U$, and $V := V_0 \cup V_1$ is an interval since $f(x) \in V_0 \cap V_1$. Now let $U' = U \cap (d_{i-1}, d_{i+1})$. We then have $\ell(V) \leq \varepsilon$ and $(U', V) \in f_i \cap f_{i+1} \implies (U', V) \in f$.

Case 3: $x = d_0$ or $x = d_k$. A similar argument to **Case 1** works. □

6.2 $\mathsf{CL_c}$ fails in $\mathsf{RCA_0}$

We are now ready to discuss Cousin's lemma for continuous functions. The usual definitions of gauge, partition, and δ-fine (as in Section 5.3) are still valid here.

Definition 6.2.1. Let $\mathsf{CL_c}$ be the following statement in $\mathsf{RCA_0}$: every total continuous gauge $\delta\colon [0,1] \to \mathbb{R}^+$ has a δ-fine partition.

Remark 6.2.2. When we say that a theorem φ *holds* in a model $\mathcal{M}$ of second-order arithmetic, all the quantification and interpretations should be made relative to *objects in* $\mathcal{M}$. So, to say that $\mathsf{CL_c}$ holds in $\mathcal{M}$ is to say that for every object δ *in* $\mathcal{M}$, which $\mathcal{M}$ *believes* to be total, continuous and a gauge, there is a object P *in* $\mathcal{M}$, which $\mathcal{M}$ *believes* to be a finite, δ-fine sequence of reals.

If we wanted to show $\mathsf{CL_c}$ fails in $\mathcal{M}$, then we would need to construct an object δ *in* $\mathcal{M}$, which $\mathcal{M}$ *believes* to be total, continuous and a gauge, and so that there is no object P *in* $\mathcal{M}$ that $\mathcal{M}$ *thinks* is a δ-fine partition. To show that $\mathsf{CL_c}$ fails in a subsystem $\mathcal{S}$ of second-order arithmetic, we need to demonstrate a model $\mathcal{M}$ of $\mathcal{S}$ where $\mathsf{CL_c}$ fails, in the sense just described.

We want to show that $\mathsf{RCA_0}$ does not prove $\mathsf{CL_c}$; to do this, we exhibit a model of $\mathsf{RCA_0}$ where $\mathsf{CL_c}$ doesn't hold. In fact, this is true in the standard model REC of recursive sets (Proposition 4.2.4). Recall $\varphi_0, \varphi_1, \varphi_2, \ldots$ is a standard enumeration of the partial computable functions (Proposition 3.1.10). Via the coding of $\mathbb{Q}$ into $\mathbb{N}$, we can assume WLOG that the φ_e take values in $\mathbb{Q}$.

To construct our counterexample, we will use the idea of a Π^0_1 *class* from classical computability.

Definition 6.2.3 [JS72]**.** A Π^0_1 *class in* $[0,1]$ is a set of the form

$$K_\psi \;:=\; [0,1] \setminus \bigcup_{n=0}^{\infty} U_{\psi(n)}$$

for some computable function $\psi\colon \mathbb{N} \to \mathbb{Q}^2$.

Recall that in second-order arithmetic, real numbers are defined as fast-converging sequences $(q_i)_{i\in\mathbb{N}}$ of rational numbers. We say a real number is *computable* if this sequence is computable, considered as a function $f\colon \mathbb{N} \to \mathbb{Q}$. These are exactly the real numbers that exist in REC. The following result is closely related to Proposition 4.3.7.

Lemma 6.2.4 [JS72]**.** There exists a nonempty Π^0_1 class which contains no computable reals.

Proof. We define ψ as follows: search over all pairs $(e, s) \in \mathbb{N}^2$ until we find the next one such that $\varphi_e(e+3)$ halts after s steps. When we find such a pair, let $\psi(n)$ be the code for the rational open ball $B_e := B(\varphi_e(e+3), 2^{-e-3})$. We can always find another such pair, so in particular, ψ is total computable.

We claim that K_ψ has the required properties. First, note that each B_e has Lebesgue measure $\lambda(B_e) = 2^{-e-2}$, so their union has measure at most $\frac{1}{2}$. In particular, the complement K_ψ must be nonempty.

Now, suppose $r = (q_n)_{n\in\mathbb{N}}$ is a *computable* real number; then r is computed by some φ_e, i.e. $\varphi_e(n) = q_n$ for all n. By the definition of real number, q_{e+3} is an approximation of r to within 2^{-e-3}. Thus, $r \in B(q_{e+3}, 2^{-e-3}) = B_e$, so $r \notin K_\psi$. ☐

Now, we are ready to construct our counterexample, to show $\mathsf{CL_c}$ fails in $\mathsf{RCA_0}$. The construction was inspired by [Ko91, Thm 3.1], and the idea is as follows. Given a Π^0_1 class K_ψ as in Lemma 6.2.4, we construct a continuous gauge δ which is positive *exactly* on the complement of K_ψ, and furthermore is 1-Lipschitz ($|\delta(x) - \delta(y)| \leq |x-y|$). Then, REC will think that δ is a gauge, since it is positive on all computable reals. However, for any point $b \in K_\psi$ and $t_i \neq b$, it is not possible for $\delta(t_i)$ to cover b, since δ is 1-Lipschitz; therefore, there are no δ-fine partitions in REC. The formal proof follows.

Theorem 6.2.5. $\mathsf{RCA_0}$ does not prove $\mathsf{CL_c}$.

Proof. Let K_ψ be a Π^0_1 class as in Lemma 6.2.4. For each n, define the *nth spike* $\mathrm{sp}_n \colon [0,1] \to \mathbb{R}$ by

$$\mathrm{sp}_n(z) = \begin{cases} 0 & 0 \leq z \leq \overline{p} \\ |z - \overline{p}| & \overline{p} \leq z \leq m \\ |z - \overline{q}| & m \leq z \leq \overline{q} \\ 0 & \overline{q} \leq z \leq 1 \end{cases}$$

where (p, q) is the code of $U_{\psi(n)}$, and $m := (\overline{p} + \overline{q})/2$ is the midpoint of $U_{\psi(n)}$.

As in Figure 6.1, sp_n is graphically a spike whose base is exactly $U_{\psi(n)}$, and whose sides have gradient ± 1. For each fixed n, sp_n is 1-Lipschitz (i.e. $|\mathrm{sp}_n(x) - \mathrm{sp}_n(y)| \leq |x-y|$) and bounded above by $\frac{1}{2}$. By Proposition 6.1.3, each part is total continuous, so sp_n is total continuous by Lemma 6.1.5.

Then, we define $\delta \colon [0,1] \to \mathbb{R}$ by

$$\delta(x) = \sum_{n=0}^{\infty} 2^{-n-2} \cdot \mathrm{sp}_n(x)$$

which is total continuous by Lemma 6.1.4. In REC, δ is a gauge; for any real $x \in \mathrm{REC}$, x is a computable real number, so x is in some $U_\psi(n)$ by definition of K_ψ. Then, $\delta(x) \geq \mathrm{sp}_n(x) > 0$, since $U_{\psi(n)}$ is open.

We claim there is no δ-fine partition in REC. Suppose, by contradiction, that $P = \langle x_i, t_i \rangle$ is such a partition, of size ℓ. By assumption, K_ψ is nonempty, so pick any point $b \in K_\psi$,

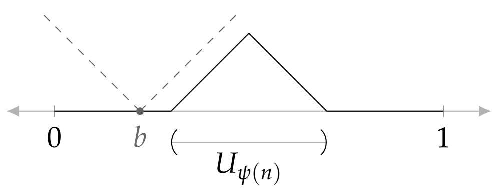

Figure 6.1: The nth spike sp_n, compared to $z \mapsto |z-b|$ for a point $b \notin U_{\psi(n)}$.

which is necessarily noncomputable. There is unique $m < \ell$ such that $b \in (x_m, x_{m+1})$; then, we claim $\delta(t_m) < |t_m - b|$. This would imply that $b \notin B(t_m, \delta(t_m))$ and thus $(x_m, x_{m+1}) \nsubseteq B(t_m, \delta(t_m))$, so P is not δ-fine after all.

Note that for all $n \in \mathbb{N}$, $b \notin U_{\psi(n)}$, and so $\mathrm{sp}_n(t_m) \leq |t_m - b|$ (see Figure 6.1). We compute:

$$\delta(t_m) \;=\; \sum_{n=0}^{\infty} 2^{-n-2} \cdot \mathrm{sp}_n(t_m) \;\leq\; \sum_{n=0}^{\infty} 2^{-n-2}|t_m - b| \;\leq\; \tfrac{1}{2}|t_m - b| \;<\; |t_m - b|.$$

□

6.3 WKL_0 proves $\mathsf{CL_c}$

We saw in the previous section that RCA_0 is not strong enough to prove Cousin's lemma for continuous functions, $\mathsf{CL_c}$. In this section, we show that WKL_0 *is* strong enough to prove $\mathsf{CL_c}$. In the next section, we demonstrate a reversal of $\mathsf{CL_c}$ in WKL_0, thus showing that $\mathsf{CL_c}$ and WKL_0 are equivalent, and that WKL_0 is the weakest subsystem of second-order arithmetic in which $\mathsf{CL_c}$ can be proved.

The idea is a variation on the "proof" of Theorem 5.3.4. We avoid using Π^1_1 comprehension by only considering the *midpoint* of each I_σ, rather than *all* real numbers in I_σ. The proof then proceeds exactly the same. To deduce a contradiction, we will use the fact (provable in WKL_0) that every continuous function $f\colon [0,1] \to \mathbb{R}$ is *uniformly continuous*:

Lemma 6.3.1 (WKL_0)**.** Let $f\colon [0,1] \to \mathbb{R}$ be a total continuous function. Then f has a *modulus of uniform continuity*, i.e. a function $h\colon \mathbb{N} \to \mathbb{N}$ such that for all $x, y \in [0,1]$,

$$|x - y| \leq 2^{-h(n)} \implies |f(x) - f(y)| \leq 2^{-n}$$

Proof. [Sim09, Thm IV.2.2]. □

Theorem 6.3.2. WKL_0 proves $\mathsf{CL_c}$.

Proof. Let $\delta\colon [0,1] \to \mathbb{R}^+$ be a total continuous gauge. For each $x \in [0,1]$, $\delta(x)$ is a real number, hence a sequence of rationals; let $\delta{\restriction}_n(x)$ denote the nth term in this sequence. For each $\sigma \in 2^{<\omega}$, let

$$m_\sigma = 2^{|\sigma|-1} + \sum_{i<n} \sigma_i \cdot 2^{-i-1}$$

be as in the "proof" of Theorem 5.3.4.

We define a tree $T \subseteq 2^{<\omega}$ in levels. For each $n \in \mathbb{N}$, inductively define

$$T_n \;:=\; \left\{\sigma \in 2^{<\omega} : \; |\sigma| = n,\; \forall k < n\; (\sigma|_k \in T_k),\; \delta{\restriction}_n(m_\sigma) \leq 2^{-n+1}\right\}$$

The $\leq$ relation between rationals is computable, so by Δ^0_1 comprehension, each T_n exists. Then, $T = \bigcup_{n=0}^{\infty} T_n$ is a tree by construction; if it is finite, we construct a δ-fine partition as in Theorem 5.3.4 (this can all be done in RCA_0).

Now, we claim T must be finite, so the above construction always works. Suppose by contradiction that T is infinite. By WKL_0, there is an infinite path X through T. Again, we define the real $r_X = (q_n)_{n\in\mathbb{N}}$, where each $q_n = m_{X{\restriction}_n}$. By definition of T, we have $\delta{\restriction}_n(q_n) \leq 2^{-n+1}$, and $\left|\delta(q_n) - \delta{\restriction}_n(q_n)\right| \leq 2^{-n}$ by definition of $\delta{\restriction}_n(x)$, so for each n,

$$\delta(q_n) \;\leq\; 3 \cdot 2^{-n} \tag{6.1}$$

By Lemma 6.3.1, pick $h\colon \mathbb{N} \to \mathbb{N}$ a modulus of uniform continuity for δ. By definition of r_X, for each $n, k \in \mathbb{N}$, we have

$$\left|q_{h(n)+k} - r_X\right| \;\leq\; 2^{-h(n)-k} \;\leq\; 2^{-h(n)}$$

hence by definition of h,

$$\left|\delta(q_{h(n)+k}) - \delta(r_X)\right| \;\leq\; 2^{-n}$$

Combining this with equation (6.1), we get that for all $n, k \in \mathbb{N}$,

$$\delta(r_X) \;\leq\; 2^{-n} \;+\; 3 \cdot 2^{-h(n)-k}$$

We can make this arbitrarily small by picking the right n and k; hence $\delta(r_X) = 0$, contradicting the fact that δ is a gauge. $\square$

6.4 $\mathsf{CL_c}$ is equivalent to $\mathsf{WKL_0}$

In the previous section we showed that $\mathsf{CL_c}$ can be proved in $\mathsf{WKL_0}$. Now, we show that $\mathsf{WKL_0}$ is the weakest system having this property, by demonstrating a reversal of $\mathsf{CL_c}$ in $\mathsf{WKL_0}$. The reversal goes through the Heine–Borel theorem HB, which is known to be equivalent to $\mathsf{WKL_0}$.

Definition 6.4.1. An *open cover of* $[0,1]$ is a (finite or infinite) sequence (U_i) of open intervals in $\mathcal{B}_{[0,1]}$, such that every $x \in [0,1]$ is in some U_i.

Definition 6.4.2 [Sim09, Lem IV.1.1]**.** Let HB be the following statement in $\mathsf{RCA_0}$: for every infinite open cover $(U_i)_{i\in\mathbb{N}}$ of $[0,1]$, there is n such that $(U_i)_{i\leq n}$ is a finite open cover of $[0,1]$.

Proposition 6.4.3 ($\mathsf{RCA_0}$) [Sim09, Lem IV.1.1]**.** $\mathsf{WKL_0}$ is equivalent to HB.

Now, we show that over $\mathsf{RCA_0}$, $\mathsf{CL_c}$ implies HB, and thereby $\mathsf{WKL_0}$. The idea of the proof is similar to Theorem 6.2.5; given an open cover $(U_i)_{i\in\mathbb{N}}$, we define δ on the U_i in the same way. This time, we assume $\mathsf{CL_c}$, so there is a δ-fine partition $P = \langle x_j, t_j \rangle$. Then, for each t_j, we can find some U_i such that $B\big(t_j, \delta(t_j)\big) \subseteq U_i$. Since the balls $\big(t_j, \delta(t_j)\big)$ cover $[0,1]$, it follows that the corresponding U_i also cover $[0,1]$, so we get a finite subcover.

Theorem 6.4.4 ($\mathsf{RCA_0}$)**.** $\mathsf{CL_c}$ implies HB.

Proof. Let $(U_i)_{i\in\mathbb{N}}$ be an open cover of $[0,1]$. We define δ as in the proof of Theorem 6.2.5, where for each i, $\psi(i)$ is the code for U_i. Formally, define $\mathrm{sp}_i\colon [0,1] \to \mathbb{R}$ by

$$\mathrm{sp}_n(z) = \begin{cases} 0 & 0 \leq z \leq \overline{p} \\ |z - \overline{p}| & \overline{p} \leq z \leq m \\ |z - \overline{q}| & m \leq z \leq \overline{q} \\ 0 & \overline{q} \leq z \leq 1 \end{cases}$$

where (p, q) is the code for U_i, and $m := (\overline{p} + \overline{q})/2$ is the midpoint of U_i. Again, each sp_n is total continuous.

Then, we define $\delta\colon [0,1] \to \mathbb{R}$ by

$$\delta(x) = \sum_{n=0}^{\infty} 2^{-n-2} \cdot \mathrm{sp}_n(x)$$

which is total continuous by Lemma 6.1.4. Furthermore, δ is a gauge, since for any $x \in [0,1]$, $x \in U_n$ for some n, so $\delta(x) \geq \mathrm{sp}_n(x) > 0$.

Claim 6.4.4.1. Let $r = (q_n)_{n\in\mathbb{N}}$ be a real number, and for each $e \in \mathbb{N}$,

$$y_e \;=\; \sum_{n=0}^{e} 2^{-n-2} \cdot \mathrm{sp}_n(q_e)$$

Then, for each $e \in \mathbb{N}$, $|y_e - \delta(r)| \leq 2^{-e}$.

Proof of Claim 6.4.4.1. Recall that each sp_n is 1-Lipschitz, i.e. $|\mathrm{sp}_n(x) - \mathrm{sp}_n(y)| \leq |x - y|$ for all $x, y \in [0, 1]$; and $\mathrm{sp}_n(x) \leq 1/2$ for all $x \in [0, 1]$. Using these facts, we compute:

$$\begin{aligned}
|y_e - \delta(r)| \;&=\; \left| \sum_{n=0}^{e} 2^{-n-2}\big(\mathrm{sp}_n(q_e) - \mathrm{sp}_n(r)\big) \;-\; \sum_{n=e+1}^{\infty} 2^{-n-2} \cdot \mathrm{sp}_n(r) \right| \\
&\leq\; \sum_{n=0}^{e} 2^{-n-2} |\mathrm{sp}_n(q_e) - \mathrm{sp}_n(r)| \;+\; \sum_{n=e+1}^{\infty} 2^{-n-2} \cdot \mathrm{sp}_n(r) \\
&\leq\; \sum_{n=0}^{e} 2^{-n-2} |q_e - r| \;+\; \sum_{n=e+1}^{\infty} 2^{-n-2} \cdot \tfrac{1}{2} \\
&\leq\; 2^{-e} \sum_{n=0}^{e} 2^{-n-2} \;+\; 2^{-e-3} \\
&\leq\; 2^{-e} \cdot \tfrac{1}{2} \;+\; 2^{-e-3} \;=\; 2^{-e} \cdot \tfrac{5}{8} \;\leq\; 2^{-e}.
\end{aligned}$$

□

By assumption, there exists a δ-fine partition $P = \langle x_j, t_j \rangle$. Let ℓ be the size of P.

Claim 6.4.4.2. For each $j < \ell$, there exists $m = m_j \in \mathbb{N}$ such that $\delta(t_j) < \mathrm{sp}_m(t_j)$.

Proof of Claim 6.4.4.2. Let y_e be as in the previous claim, for $r = t_j$, and let $e := \min\{k : y_k \geq (3k + 1)2^{-k-1}\}$. This set is nonempty since $\delta(t_j) > 0$, so we can find e in RCA_0 by minimisation [Sim09, Thm II.3.5]. We must also have $e \geq 1$.

We claim there is $m \leq e$ such that $\mathrm{sp}_m(q_e) > (3e + 1)2^{-e}$. If there were not (i.e. $\mathrm{sp}_n(q_e) \leq (3e + 1)2^{-e}$ for all $n \leq e$), then

$$y_e \;=\; \sum_{n=0}^{e} 2^{-n-2} \cdot \mathrm{sp}_n(q_e) \;\leq\; \sum_{n=0}^{e} 2^{-n-2}(3e + 1)2^{-e} \;<\; (3e + 1)2^{-e-1}$$

contradicting the definition of e.

We take m_j to be the least such m, and claim this is as required. Because sp_m is 1-Lipschitz, we have $\left|\mathrm{sp}_m(q_e) - \mathrm{sp}_m(t_j)\right| \leq \left|q_e - t_j\right| \leq 2^{-e}$. We compute:

$$\begin{aligned}
\delta(t_j) \;&\leq\; y_{e-1} + 2^{-e+1} && \text{since } \left|\delta(t_j) - y_{e-1}\right| \leq 2^{-e+1} \\
&<\; (3e - 2)2^{-e} + 2(2^{-e}) && \text{by definition of } e \\
&=\; 3e \cdot 2^{-e} \\
&=\; (3e + 1)2^{-e} - 2^{-e} \\
&<\; \mathrm{sp}_m(q_e) - 2^{-e} && \text{by definition of } m \\
&\leq\; \mathrm{sp}_m(t_j) && \text{since } \left|\mathrm{sp}_m(q_e) - \mathrm{sp}_m(t_j)\right| \leq 2^{-e}
\end{aligned}$$

□

Proof of Theorem 6.4.4, continued. For each j, fix m_j as in the claim. Then, $n := \max\{m_j : j < \ell\}$ gives a finite subcover. Taking $z \in [0, 1]$, there is some $j < \ell$ such that $x_j \leq x \leq x_{j+1}$. Then, $t_j - \delta(t_j) \leq z \leq t_j + \delta(t_j)$ since P is δ-fine, i.e. $\left|z - t_j\right| \leq \delta(t_j) < \mathrm{sp}_{m_j}(t_j)$ by the claim. It follows that $x \in U_{m_j}$. □

Theorem 6.4.5 (RCA_0)**.** CL_c is equivalent to WKL_0.

Proof. The forward direction is Theorem 6.4.4 and Proposition 6.4.3, while the reverse direction is Theorem 6.3.2. $\square$

Chapter 7

Cousin's lemma for Baire functions

In Chapter 6, we completely characterised the axiomatic strength of Cousin's lemma for continuous functions $\mathsf{CL_c}$, showing its equivalence to $\mathsf{WKL_0}$ over $\mathsf{RCA_0}$. In this chapter, we will define the Baire classes of functions, and study the strength of $\mathsf{CL}_{\mathsf{B}n}$, Cousin's lemma for functions of a given Baire class n. In contrast to $\mathsf{CL_c}$, the reverse mathematics of $\mathsf{CL}_{\mathsf{B}n}$ appears much harder, and so far has resisted complete characterisation for any $n \geq 1$.

7.1 Baire classes of functions

It is well-known that, while *uniform* limits of continuous functions remain continuous, general pointwise limits don't have to be. A famous example are the functions

$$f_n(x) = \begin{cases} -1 & x \leq {-1}/n \\ nx & |x| < 1/n \\ 1 & x \geq 1/n \end{cases}$$

which converge pointwise, non-uniformly, to the Heaviside step function (Figure 7.1):

$$f(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases}$$

Taking all pointwise limits of continuous functions gives the *Baire 1 functions*. The Baire 1 functions aren't closed under pointwise limits either, so again taking *their* pointwise limits gives the Baire 2 functions.

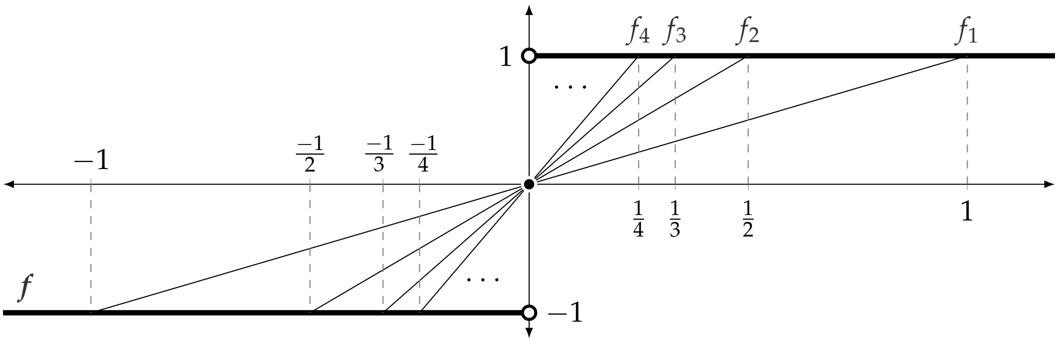

Figure 7.1: A sequence of continuous functions f_n converging to a discontinuous function f.

We can continue this process transfinitely up to ω_1, at which point the Baire hierarchy collapses. In this way, the Baire classes assign a measure of complexity to the Borel functions. Indeed, an equivalent definition of Baire class α is that the preimage of any open set is $\mathbf{\Sigma}^0_{\alpha+1}$ in the Borel hierarchy. We will only be concerned with finite Baire classes here.

The Baire classes were introduced by Baire in his PhD thesis [Bai99], as a natural generalisation of the continuous functions. One motivation for Baire functions is that many functions arising in analysis are not continuous, such as step functions [Hea93], Walsh functions [Wal23], or Dirichlet's function [Dir29]. However, all such "natural" functions generally have low Baire class; for example, the derivative of any differentiable function is Baire 1, as are functions arising from Fourier series [KL90].

The Baire class functions have previously been studied with respect to computability [KT14; PDD17]. In particular, Kuyper and Terwijn showed a real number x is 1-generic ("random") if and only if every *effective* Baire 1 functions is continuous at x [KT14].

Because continuous functions can be specified by countable information, so can Baire 1 functions (countably many continuous functions), and by induction, so can functions of any Baire class. Therefore, we can define Baire functions in second-order arithmetic, coding them using subsets of $\mathbb{N}$.

Definition 7.1.1. The following definitions proceed simultaneously and inductively on n.

(i) The *Baire 0 functions* are exactly the (total) continuous functions of Definition 6.1.1.

(ii) For each $n \in \mathbb{N}$, a *Baire $n+1$ function* $f\colon [0,1] \to \mathbb{R}$ is a countable sequence $(f_n)_{n\in\mathbb{N}}$ of Baire n functions $[0,1] \to \mathbb{R}$ which is pointwise Cauchy. That is, for each $x \in [0,1]$ and $\varepsilon \in \mathbb{Q}^+$, $|f_m(x) - f_n(x)| \leq \varepsilon$ for sufficiently large m, n.

(iii) Given two Baire $n+1$ functions $f = (f_n)_{n\in\mathbb{N}}$, $g = (g_n)_{n\in\mathbb{N}}\colon [0,1] \to \mathbb{R}$, a point $x \in [0,1]$, and $\varepsilon \in \mathbb{Q}^+$, we say $|f(x) - g(x)| \leq \varepsilon$ if for any $\delta \in \mathbb{Q}^+$, we have $|f_n(x) - g_n(x)| \leq \varepsilon + \delta$ for sufficiently large n.

Here, we have only required our Baire functions to be pointwise *Cauchy*. This means, in weak subsystems such as RCA_0 and WKL_0 which can't prove the completeness of the reals, the function value $f(x)$ may not actually *exist*. We could have made the stronger requirement that (f_n) is pointwise *convergent*; however, Definition 7.1.1 proves to be the right one for a reverse-mathematical analysis.

Example 7.1.2. Any Baire n function f can be identified with a Baire $n+1$ function $\tilde{f} = (f)_{n\in\mathbb{N}}$. Therefore, the Baire classes are nested: $\mathsf{B0} \subseteq \mathsf{B1} \subseteq \mathsf{B2} \subseteq \cdots$.

Since the function values $f(x)$ may not actually exist, we must take some care when making definitions concerning Baire functions. We can see this already in Definition 7.1.1, where we had to define what $|f(x) - g(x)| \leq \varepsilon$ means, despite the fact that both values may not exist. These difficulties can generally be overcome with a little caution.

Definition 7.1.3.

(i) The following definition proceeds inductively on n. Given a Baire $n+1$ function $f = (f_n)_{n\in\mathbb{N}}\colon [0,1] \to \mathbb{R}$, a point $x \in [0,1]$, and a point $y \in \mathbb{R}$, we say $f(x) \geq y$ if for all rational $q < y$, $f_m(x) > q$ for sufficiently large m. $f(x) \leq y$ is defined similarly.

(ii) We say $f(x) > y$ if it is not true that $f(x) \leq y$. $f(x) < y$ is defined similarly.

(iii) A Baire n function $\delta\colon [0,1] \to \mathbb{R}$ is a *gauge* if $\delta(x) > 0$ for all $x \in [0,1]$. This property will be denoted $\delta\colon [0,1] \to \mathbb{R}^+$.

Definition 7.1.4. Let $\delta\colon [0,1] \to \mathbb{R}^+$ be a Baire n gauge, and P be a tagged partition of size ℓ. Then, we say P is *δ-fine* if for all $j < \ell$, we have $\delta(t_j) \geq t_j - x_j$ and $\delta(t_j) \geq x_{j+1} - t_j$, in the sense of Definition 7.1.3.

Having defined gauges and δ-fine partitions, we are now ready to define Cousin's lemma for Baire n functions:

Definition 7.1.5. For each $n \in \mathbb{N}$, let $\mathsf{CL}_{\mathsf{B}n}$ be the following statement in RCA_0: every Baire n gauge $\delta\colon [0,1] \to \mathbb{R}^+$ has a δ-fine partition.

Note that $\mathsf{CL}_{\mathsf{c}} = \mathsf{CL}_{\mathsf{B}0}$, and for each $m \geq n$, $\mathsf{CL}_{\mathsf{B}m} \vdash \mathsf{CL}_{\mathsf{B}n}$, since the Baire classes are nested. Combining these with the results of the previous section, immediately we get:

Theorem 7.1.6 (RCA_0)**.** For each $n \in \mathbb{N}$, $\mathsf{CL}_{\mathsf{B}n}$ implies WKL_0.

As we will see, for $n \geq 1$, this is far from optimal; $\mathsf{CL}_{\mathsf{B}n}$ is much stronger than WKL_0.

7.2 $\mathsf{CL}_{\mathsf{B}1}$ proves ACA_0

Having completely classified $\mathsf{CL}_{\mathsf{c}} = \mathsf{CL}_{\mathsf{B}0}$, the natural next step would be to study the reverse-mathematical strength of $\mathsf{CL}_{\mathsf{B}1}$. Our first result about $\mathsf{CL}_{\mathsf{B}1}$ is a reversal; we show that over RCA_0, $\mathsf{CL}_{\mathsf{B}1}$ proves ACA_0. In other words, to prove $\mathsf{CL}_{\mathsf{B}1}$ in second-order arithmetic, we need a system *at least* as strong as ACA_0. The reversal goes through the sequential completeness of $\mathbb{R}$, which is known to be equivalent to ACA_0.

Definition 7.2.1. Let SC be the following statement in RCA_0: every Cauchy sequence of real numbers in $[0,1]$ has a limit.

Theorem 7.2.2 (RCA_0)**.** $\mathsf{CL}_{\mathsf{B}1}$ implies SC.

The idea of the proof is as follows. Supposing we have a Cauchy sequence (z_n) with no limit, we look at the sequence of functions $\delta_n\colon x \mapsto \frac{1}{2}|x - z_n|$. This is pointwise Cauchy, hence Baire 1, and it is a gauge since (z_n) has no limit. But $\delta = (\delta_n)$ can't have a δ-fine partition, since no partition P can cover the gap where $\lim z_n$ should be. Here are the details.

Proof of Theorem 7.2.2. By contradiction: suppose there is a Cauchy sequence $(z_n)_{n\in\mathbb{N}} \subseteq [0,1]$ that has no limit in $[0,1]$. For each $n \in \mathbb{N}$, let $\delta_n\colon x \mapsto \frac{1}{2}|x - z_n|$, which defines a continuous function by earlier lemmas.

The sequence $(\delta_n)_{n\in\mathbb{N}}$ is pointwise Cauchy; fixing $\varepsilon \in \mathbb{Q}^+$, we have $|z_m - z_n| \leq \varepsilon$ for sufficiently large m, n. But $|z_m - z_n| \leq \varepsilon$ implies $|\delta_m(x) - \delta_n(x)| \leq \varepsilon$ by the reverse triangle inequality:

$$|\delta_m(x) - \delta_n(x)| \;=\; \left|\tfrac{1}{2}|x - z_m| - \tfrac{1}{2}|x - z_n|\right| \;\leq\; \tfrac{1}{2}|z_n - z_m| \;\leq\; \varepsilon.$$

It follows that $|\delta_m(x) - \delta_n(x)| \leq \varepsilon$ for sufficiently large m, n, whence $\delta := (\delta_n)_{n\in\mathbb{N}}$ is Baire 1.

Now, because (z_n) does not have a limit, we claim δ is a gauge. For any $x \in [0,1]$, since x is not a limit for (z_n), there is some $\varepsilon \in \mathbb{Q}^+$ such that $\delta_n(x) = |x - z_n| \geq \varepsilon$ eventually. Thus $\delta(x) > 0$. Let $P = \langle x_j, t_j \rangle$ be a partition of $[0,1]$; we will show P is not δ-fine.

Fix $j < \ell$. Since x_j is not the limit of (z_n), we must eventually have $|x_j - z_n| \geq \varepsilon$ for some $\varepsilon \in \mathbb{Q}^+$. But since (z_n) is Cauchy, all the terms are eventually within 2ε of each other; from this point on, we must have $z_n < x_j$ for all n, or $z_n > x_j$ for all n. As $(z_n) \subseteq [0,1]$, we can't have $z_n < x_0 = 0$ or $z_n > x_\ell = 1$. It follows that there is $j < \ell$ such that eventually $x_j < z_n < x_{j+1}$.

For this j, we claim that $\delta(t_j)$ cannot cover (x_j, x_{j+1}). By the same argument to the previous paragraph, either $z_n < t_j$ eventually, or $z_n > t_j$ eventually; let us suppose WLOG that $z_n < t_j$. Then,

$$\delta_n(t_j) \;=\; \frac{1}{2}\left|t_j - z_n\right| \;<\; \left|t_j - z_n\right| \;<\; \left|t_j - x_j\right|$$

for sufficiently large n, so $\delta(t_j) < t_j - x_j$. Thus, P is not δ-fine. $\square$

Having demonstrated the reversal $\mathsf{CL_{B1}} \vdash \mathsf{ACA_0}$, it seems natural to see if we can get a proof of $\mathsf{CL_{B1}}$ in $\mathsf{ACA_0}$. The most natural way to do this is as follows. A Baire 1 function δ is a pointwise limit of continuous gauges δ_n, and we have already seen that $\mathsf{ACA_0}$ (in fact, $\mathsf{WKL_0}$) can construct a δ_n-fine partition for each n. Therefore, one might expect that there would be some inductive way to combine the partitions for δ_n into a partition for δ.

The following proposition suggests that this is not possible. Looking back to the proof of $\mathsf{CL_c}$, for each continuous gauge δ_n, we in fact constructed a *dyadic* δ_n-fine partition—one whose partition points and tag points were all *dyadic rationals*, i.e. of the form $j/2^n$. If there were a way to combine these partitions into one for δ, it would follow that every Baire 1 gauge has a dyadic partition. Now, we present a Baire 1 gauge with no dyadic partition.

Proposition 7.2.3. There is a Baire 1 gauge $\delta\colon [0,1] \to \mathbb{R}$ with no dyadic δ-fine partition.

Proof. The desired gauge is

$$\delta(x) = \begin{cases} \frac{1}{2}\left|x - \frac{1}{3}\right| & x \neq \frac{1}{3} \\ 1 & x = \frac{1}{3} \end{cases}$$

It is the pointwise limit of the following sequence $\delta_n\colon [0,1] \to \mathbb{R}$, where $n \geq 3$:

$$\delta_n(x) = \begin{cases} \frac{1}{2}\left|x - \frac{1}{3}\right| & \left|x - \frac{1}{3}\right| \geq \frac{1}{n} \\ 1 - \left(n - \frac{1}{2}\right)\left|x - \frac{1}{3}\right| & \left|x - \frac{1}{3}\right| < \frac{1}{n} \end{cases}$$

By Lemma 6.1.5, each δ_n is continuous, so δ is Baire 1. Now, any δ-fine partition P must have some $t_i = \frac{1}{3}$, because otherwise $\delta(t_i) = \frac{1}{2}\left|t_i - \frac{1}{3}\right| < \left|t_i - \frac{1}{3}\right|$. Then $\frac{1}{3} \notin B\big(t_i, \delta(t_i)\big)$, so $(x_j, x_{j+1}) \nsubseteq B\big(t_j, \delta(t_j)\big)$ for the $j < \ell$ such that $\frac{1}{3} \in (x_j, x_{j+1})$. It follows that a δ-fine partition P cannot be dyadic. $\square$

7.3 $\mathsf{CL_{B2}}$ fails in $\mathsf{ACA_0}$

In attempting to show the implication in Theorem 7.2.2 is strict, we tried to construct an arithmetical Baire 1 gauge with no arithmetical partition. This would show that $\mathsf{CL_{B1}}$ fails in the standard model ARITH of $\mathsf{ACA_0}$, and thus that $\mathsf{ACA_0} \nvdash \mathsf{CL_{B1}}$.

We have not yet been able to construct such a Baire 1 gauge, but we were able to construct an arithmetical *Baire 2 gauge* with no arithmetical partition. We present the construction in this section, showing that $\mathsf{ACA_0} \nvdash \mathsf{CL_{B2}}$. In the next section, we will slightly generalise the ideas of the proof to show $\mathsf{CL_{B2}} \vdash \mathsf{ATR_0}$.

Instead of working in $[0,1]$ as before, we will actually construct our gauges in Cantor space 2^ω. The definitions are generally analogous to the $[0,1]$ case:

- A *gauge* on 2^ω is a positive real-valued function $\delta\colon 2^\omega \to \mathbb{R}^+$.
- 2^ω is a metric space under the distance function

$$d(X,Y) = \begin{cases} 0 & X = Y \\ 2^{-n} & n \text{ least such that } X_n \neq Y_n \end{cases}$$

- A *δ-fine partition* is a finite set $P \subseteq 2^\omega$ such that $\{B(X, \delta(X)) : X \in P\}$ is an open cover of 2^ω.
- *Cousin's lemma* says that every gauge $\delta \colon 2^\omega \to \mathbb{R}^+$ has a δ-fine partition $P \subseteq 2^\omega$.
- 2^ω has a countable basis of basic open sets $[\sigma] = \{X \in 2^\omega : \sigma \preceq X\}$ for each $\sigma \in 2^{<\omega}$. We can code basic open sets $[\sigma]$ by natural numbers, then give the same definitions of open sets, continuous/Baire functions $f \colon 2^\omega \to \mathbb{R}$, etc... in second-order arithmetic.

There is a well-known embedding $g \colon 2^\omega \to [0,1]$ defined by

$$g(X) \;=\; \sum_{n=0}^{\infty} \frac{2X_n}{3^{n+1}}$$

The range of g is the Cantor middle-thirds set $C \subseteq [0,1]$, and topologically, g is a homeomorphism $2^\omega \to C$. Via the embedding g, we can map any Baire n gauge $\delta \colon [0,1] \to \mathbb{R}^+$ to a Baire n gauge $\bar{\delta} \colon 2^\omega \to \mathbb{R}^+$, and vice versa. Furthermore, we can do this in such a way to preserve covering, i.e. $d(X,Y) < \bar{\delta}(Z)$ if and only if $|g(X) - g(Y)| < \delta(g(Z))$.

There is one difficulty to contend with: when going from 2^ω to $[0,1]$, this correspondence only defines a gauge δ on $C \subseteq [0,1]$. However, since $C \subseteq [0,1]$ is closed, any point $x \notin C$ has positive distance r to C, so we can just choose $\delta(x) < r$. This ensures the aforementioned covering property is preserved, and we can make this choice in a Baire 1 way. It follows that, for $n \geq 1$, $\mathsf{CL}_{\mathsf{B}n}$ for gauges on $[0,1]$ and $\mathsf{CL}_{\mathsf{B}n}$ for gauges on 2^ω are equivalent.

Before we see the proofs, we need to introduce a bit more computability. In Section 3.1, we focused on *absolute* computability—the existence of an algorithm to solve some problem (e.g. membership in a set $A \subseteq \mathbb{N}$). However, many natural problems are not computable in this sense. This leads us to a more general notion of *relative* computability.

The idea is we allow our computations access to an *oracle*—a (noncomputable) set A. While performing our algorithm, we are allowed to query A at any point, and ask if it contains some element or not. We say B is *Turing reducible* to A ($B \leq_{\mathrm{T}} A$) if there is an algorithm which can compute B, with A as an oracle. $\leq_{\mathrm{T}}$ is a preorder on $\mathcal{P}(\mathbb{N})$, forming a hierarchy known as the *Turing degrees*. Intuitively, one should think of $B \leq_{\mathrm{T}} A$ as meaning that A has more computational power than B, or that A is *less computable* than B.

The *Turing jump* is an operation assigning to every set $A \subseteq \mathbb{N}$ a set $A' >_{\mathrm{T}} A$ which is strictly higher in the Turing degrees, i.e. A' (read "A-jump") is less computable than A. We can iterate this operation, getting a sequence $A <_{\mathrm{T}} A' <_{\mathrm{T}} A'' <_{\mathrm{T}} A^{(3)} <_{\mathrm{T}} A^{(4)} <_{\mathrm{T}} \cdots$. This is a countable sequence of countable sets, so we can combine them all into a single countable set $A^{(\omega)} = \{(e,n) : e \in A^{(n)}\}$, called the *arithmetic jump* or *ω-jump* of A. The Turing jumps of the empty set $\varnothing$ provide some useful milestones in the arithmetical hierarchy:

Proposition 7.3.1.

- $\varnothing^{(n)}$ is strictly Σ^0_n in the arithmetical hierarchy.
- $\varnothing^{(\omega)}$ is nonarithmetical.

We are now ready to prove the first result: that $\mathsf{CL}_{\mathsf{B}2}$ fails in ACA_0.

Theorem 7.3.2. ACA_0 does not imply $\mathsf{CL}_{\mathsf{B}2}$.

The idea is similar to Theorem 6.2.5, and proceeds as follows. We work in the standard model ARITH of ACA_0. In Proposition 7.3.1, we saw that $X = \varnothing^{(\omega)}$ is nonarithmetical; hence, it does not exist in ARITH. However, the singleton set $\{X\} \subseteq 2^\omega$ is *effectively G_δ*; i.e. it can be written $\{X\} = \bigcap_{n \in \mathbb{N}} O_n$ for a computable sequence $O_n \subseteq 2^\omega$ of open sets [Odi99, Prop XII.2.19]. Implicitly using this result, we can construct (arithmetically) a Baire 2 function $\delta \colon 2^\omega \to \mathbb{R}^{\geq 0}$ such that:

(i) $\delta(Y) = 0 \iff Y = X$;

(ii) For all $Y \neq X \in 2^\omega$, $\delta(Y) < d(Y, X)$.

By property (i), ARITH believes that δ is a gauge, since it is positive everywhere except X; in particular, at every arithmetical point. However, by property (ii), no point $Y \neq X$ can δ-cover X, and hence there is no arithmetical δ-fine partition.

In what follows, we freely identify an element $Y \in 2^\omega$ of Cantor space with the set $\{n \in \mathbb{N} : Y_n = 1\}$.

Proof of Theorem 7.3.2. Let $X = \varnothing^{(\omega)}$. We define a function $\delta\colon 2^\omega \to \mathbb{R}^{\geq 0}$ as follows. For every $Y \in 2^\omega$, we want to find a position k where $Y_k \neq X_k$, and define $\delta(Y)$ accordingly.

First, we consider the columns $Y^{[n]} = \{e : (e, n) \in Y\}$, and try to find the least column where $Y^{[n]} \neq X^{[n]}$. Ask if $Y^{[0]} = \varnothing$; then if $Y^{[1]} = \left(Y^{[0]}\right)'$, then if $Y^{[2]} = \left(Y^{[1]}\right)'$, etc. The desired column is the first one where the answer is "no". Having found this column $Y^{[n]}$, we simply search along it to find k, the first point of difference from $\left(Y^{[n-1]}\right)'$ or $\varnothing$. Then, let $\delta(Y) = 2^{-k-1}$.

Since $Y^{[n]}$ is the *least* column of difference, by induction we have $Y^{[m]} = \varnothing^{[m]}$ for all $m < n$. Thus, we indeed have $Y_k \neq X_k$. So for any $Y \neq X$, we *will* find a point of difference, whence property (i) above holds. Also, $d(Y, X) \geq 2^{-k} > 2^{-k-1} = \delta(Y)$, giving property (ii).

Now, the question $A = B'$ can be answered by the double-jumps A'' and B'' [Odi99, Prop XII.2.19], so it follows that δ is computable from the double-jump function $X \mapsto X''$. Since the double-jump is Baire 2 [PDD17], and computable reductions are always continuous [PR89], this implies that δ is Baire 2.

So, we have constructed a Baire 2 gauge δ in ARITH; now we claim that it has no δ-fine partition. The argument is as sketched—suppose $P \subseteq 2^\omega$ is a finite subset. Property (ii) implies that for all $Y \in P$, $\delta(Y) < d(Y, X)$, and so $X \notin B\big(Y, \delta(Y)\big)$. Thus, P is not a δ-fine partition. □

Viewing this proof in a different light, it can be construed as a proof from $\mathsf{CL}_{\mathsf{B2}}$ that $\varnothing^{(\omega)}$ exists. The argument is by contradiction: if $\varnothing^{(\omega)}$ doesn't exist, then the function δ constructed in the proof of Theorem 7.3.2 *is* a gauge. As we essentially argued, any δ-fine partition then must include $X = \varnothing^{(\omega)}$, hence this set exists.

The same argument works replacing $\varnothing$ by *any* arithmetical set A. As a corollary, then, we see that $\mathsf{CL}_{\mathsf{B2}}$ implies the stronger system ACA_0^+, consisting of ACA_0 plus the assertion that the ω-jump of any set exists. The system ACA_0^+ has arisen previously in reverse mathematics, first with the work of Blass, Hirst and Simpson in combinatorics and topological dynamics [BHS87]. Later, it surfaced in Shore's work on Boolean algebras [Sho05], and Downey and Kach's work on Euclidean domains [DK11].

7.4 $\mathsf{CL}_{\mathsf{B2}}$ proves ATR_0

Using a similar idea to the proof of Theorem 7.3.2, we can show that $\mathsf{CL}_{\mathsf{B2}}$ implies ATR_0. We don't prove the existence of $\Theta^\alpha(X)$ for *every* arithmetical operator Θ, countable ordinal α, and set $X \subseteq N$ (Definition 4.5.2). Instead, it is enough to show this for $\Theta = \mathrm{TJ}\colon A \mapsto A'$, the Turing jump operator [Sim09, Thm VIII.3.15]. This is because the Turing jump $A \mapsto A'$ is a *universal Σ^0_1 operator* [Sim09, Defn VIII.1.9], and so any arithmetical operator can be expressed using a finite number of Turing jumps. Otherwise, the proof proceeds along similar lines; the details follow.

Theorem 7.4.1. $\mathsf{CL}_{\mathsf{B2}}$ implies ATR_0.

Proof. Let $\mathcal{M}$ be a model of $\mathsf{CL}_{\mathsf{B2}}$, so in particular, $\mathcal{M} \models \mathsf{CL}_{\mathsf{B1}}$ and hence ACA_0 by Theorem 7.2.2. By contradiction, suppose that ATR_0 fails in $\mathcal{M}$. Then, there is a set $A \subseteq \mathbb{N}$ and countable ordinal α such that the αth Turing jump $X = A^{(\alpha)}$ doesn't exist in $\mathcal{M}$.

Using the same construction as the proof of Theorem 7.3.2, for each $Y \in 2^\omega$, we can find the least column $\beta = \beta_Y < \alpha$ such that $Y^{[\beta]} \neq \left(Y^{[<\beta]}\right)'$, then find the first point of difference $k = k_Y$. Furthermore, the function $\delta(Y) = 2^{-k-1}$ is Baire 2, as before.

$\mathsf{CL}_{\mathsf{B2}}$ gives a δ-fine partition P; let $\beta^* = \max\{\beta_Y : Y \in P\}$. Then, $Z = X^{[\beta^*]}$ exists in $\mathcal{M}$ by arithmetical comprehension; we claim Z is not covered by P. For any $Y \in P$, β_Y is defined so that Y disagrees with Z on column $\beta_Y \leq \beta^*$. Hence, k_Y is a point of disagreement between Y and Z, so the *first* point of disagreement is at most k_Y. It follows that $\delta(Y) < d(Y, Z)$, thus P is not δ-fine; contradiction. □

Chapter 8

Conclusion

In this report, we introduced reverse mathematics and Cousin's lemma, and then began a reverse-mathematical analysis of Cousin's lemma for various classes of functions. We have established many original results in this direction: here is the summary of our knowledge so far.

Theorem 8.1 (summary of results)**.** All implications are over RCA_0.

(i) Cousin's lemma for continuous functions is equivalent to WKL_0.

(ii) Cousin's lemma for Baire 1 functions is provable in Π^1_1-CA_0, and it implies ACA_0.

(iii) For $n \geq 2$, Cousin's lemma for Baire n functions is provable in Π^1_1-CA_0, and it implies ATR_0.

Cousin's lemma for continuous functions, CL_c, is the only theorem for which we have been able to *completely* determine the axiomatic strength. Naturally, there is further work to be done on classifying Cousin's lemma for Baire n functions. We are still most interested in the case $n = 1$; therefore, our main open question is:

Question 8.2. Where does Cousin's lemma for Baire 1 functions, $\mathsf{CL}_{\mathsf{B1}}$, fall in the reverse-mathematical hierarchy?

Here is a heuristic reason to believe $\mathsf{CL}_{\mathsf{B1}}$ implies Π^1_1-CA_0, and is thus equivalent to it. Recall an alternative characterisation of Baire 1 functions is that the preimage of any open set is $\boldsymbol{\Sigma}^0_2$ in the Borel hierarchy. Similarly, *effectively* Baire 1 functions can be characterised as those where the preimage of any lightface Σ^0_1 class is lightface Σ^0_2.

$$
\begin{tikzcd}
\mathsf{RCA}_0 \arrow[dr, Rightarrow, "\times" marking] & \mathsf{WKL}_0 \arrow[d, Leftrightarrow] & \mathsf{ACA}_0 \arrow[dr, Rightarrow, "\times" marking] & \mathsf{ATR}_0 & & \\
& \mathsf{CL}_\mathsf{c} & \mathsf{CL}_{\mathsf{B1}} \arrow[l, Rightarrow] \arrow[u, Rightarrow] & \mathsf{CL}_{\mathsf{B2}} \arrow[l, Rightarrow] \arrow[u, Rightarrow] & \mathsf{CL}_{\mathsf{B3}} \arrow[l, Rightarrow] & \cdots \arrow[l, Rightarrow] \\
& & & & \Pi^1_1\text{-}\mathsf{CA}_0 \arrow[ull, Rightarrow] \arrow[ul, Rightarrow] \arrow[u, Rightarrow] \arrow[ur, Rightarrow] &
\end{tikzcd}
$$

Figure 8.1: A graphical summary of our contributions to the reverse mathematics zoo.

Now, the only proof of Cousin's lemma we currently know is the "proof" of Theorem 5.3.4, and variations thereof (Theorem 6.3.2). To make this proof work for a function f, we need to decide if $f^{-1}\big((2^{-n},\infty)\big)$ is empty or not. For continuous functions, this set is Σ^0_1, i.e. open, so it is enough to check if it contains any dyadic rational points (which is what we did in the proof of Theorem 6.3.2).

For Baire 1 functions, this set is Σ^0_2, as mentioned. Unfortunately, there is no easy way to determine whether an arbitrary Σ^0_2 set is empty or not; this problem is Π^1_1-hard in general. This means that Π^1_1 comprehension is *required* to make the proof work for Baire 1 functions. So, this suggests that $\mathsf{CL}_{\mathsf{B1}}$ is equivalent to $\Pi^1_1\text{-}\mathsf{CA}_0$, unless there is a smarter way to prove Cousin's lemma (and we don't believe there is).

If it turns out that $\mathsf{CL}_{\mathsf{B1}}$ and $\Pi^1_1\text{-}\mathsf{CA}_0$ are equivalent, this would also imply the equivalence of $\Pi^1_1\text{-}\mathsf{CA}_0$ to Cousin's lemma for any class of functions containing the Baire 1 functions. Otherwise, there are many more classes of functions $\mathcal{K}$ for which one could explore the strength of Cousin's lemma. The general question is thus:

Question 8.3. For a specified class of functions $\mathcal{K}$ definable in second-order arithmetic, where does Cousin's lemma for functions in $\mathcal{K}$, $\mathsf{CL}_{\mathcal{K}}$, fall in the reverse-mathematical hierarchy?

Theorem 5.3.4 shows that $\Pi^1_1\text{-}\mathsf{CA}_0$ proves $\mathsf{CL}_{\mathcal{K}}$ for any class of functions $\mathcal{K}$ definable in second-order arithmetic. Presumably, for large enough $\mathcal{K}$, $\mathsf{CL}_{\mathcal{K}}$ becomes equivalent to $\Pi^1_1\text{-}\mathsf{CA}_0$; it would be interesting to know where exactly this threshold is. Here are some other classes $\mathcal{K}$ for which one could study the strength of $\mathsf{CL}_{\mathcal{K}}$:

- Of course, the Baire n functions for $n \geq 2$.
- On that note, we mentioned that the Baire hierarchy can be iterated transfinitely, so one could equally study the Baire class α functions, for $\omega \leq \alpha < \omega_1$. Defining these in second-order arithmetic can be quite messy, but it is possible.
- The Borel functions, as the limit of all the Baire classes. There has been some study into Borel sets and functions in reverse mathematics [Sim09].
- One could look at the *strong* Baire classes $\mathsf{B}'\alpha$, where the sequence is required to be pointwise *convergent*, rather than just pointwise Cauchy. We would hope for *stability* here, i.e. $\mathsf{CL}_{\mathsf{B}'\alpha} \equiv \mathsf{CL}_{\mathsf{B}\alpha}$, but this is not immediately clear.
- The Fine continuous functions, which are those continuous with respect to the metric introduced by Fine [Fin49]. These fall strictly between continuous and Baire 1 functions, and have been studied with respect to computability [Mor01; Mor02; Bra02].

Finally, we have only studied a single theorem about gauge integration in this paper. There is a whole theory of gauge integration, with many results waiting to be analysed reverse-mathematically. For those so inclined, the following results would be interesting to study:

- The equivalence between different characterisations of the gauge integral, as given by Denjoy, Perron, Luzin and others;
- Hake's theorem: $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \lim_{c\to b^-}\int_a^c f(x)\,\mathrm{d}x$
- If $f\colon [0,1] \to \mathbb{R}$ is bounded, f is gauge integrable if and only if it is Lebesgue integrable;
- Basic facts about the gauge integral, such as convergence properties.

Bibliography

[Bai99] René-Louis Baire. *Sur les fonctions de variables réelles*. PhD thesis, École Normale Supérieure, Paris, 1899.

[BHS87] Andreas R. Blass, Jeffry L. Hirst, and Stephen G. Simpson. *Logical analysis of some theorems of combinatorics and topological dynamics*. Contemporary Mathematics **65**, pp. 125–156, 1987.

[Bir40] Garrett Birkhoff. *Lattice Theory*. Colloquium Publications **25**, American Mathematical Society, Providence, RI, 1940.

[Bir73] Garrett Birkhoff. *A Source Book in Classical Analysis*. Harvard University Press, 1973.

[Bol17] Bernard Bolzano. *Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege*. Gottlieb Haase, Prague, 1817.

[Bol32] János Bolyai. *Appendix scientiam spatii absolute veram exhibens*. In: Farkas Bolyai, *Tentamen juventutem studiosam in elementa matheseos purae*. Kali, 1832.

[Boo54] George Boole. *An Investigation of the Laws of Thought*. Walton and Maberly, London, 1854.

[Bra02] Vasco Brattka. *Some notes on Fine computability*. Journal of Universal Computer Science **8(3)**, pp. 382–395, 2002.

[Bro36] A. B. Brown. *A proof of the Lebesgue condition for Riemann integrability*. The American Mathematical Monthly **43(7)**, pp. 396–398, 1936.

[BS86] Douglas K. Brown and Stephen G. Simpson. *Which set existence axioms are needed to prove the separable Hahn-Banach theorem?* Annals of Pure and Applied Logic **31**, pp. 123–144, 1986.

[Can74] Georg Cantor. *Ueber eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen*. Journal für die reine und angewandte Mathematik **77**, pp. 258–262, 1874.

[Can77] Georg Cantor. *Ein Beitrag zur Mannigfaltigkeitslehre*. Journal für die reine und angewandte Mathematik **84**, pp. 242–258, 1877.

[Can83] Georg Cantor. *Ueber unendliche, lineare Punktmannichfaltigkeiten*. Mathematische Annalen **21(4)**, pp. 545–591, 1883.

[Cau21] Augustin-Louis Cauchy. *Cours d'Analyse de l'Ecole Royale Polytechnique*. L'Imprimerie Royale, 1821.

[CK90] C. C. Chang and H. J. Keisler. *Model Theory*. 3rd ed. Studies in Logic and the Foundations of Mathematics **73**, North Holland, Amsterdam, 1990.

[Con09] Chris J. Conidis. *On the complexity of radicals in noncommutative rings*. Journal of Algebra **322(10)**, pp. 3670–3680, 2009.

[Cou95] Pierre Cousin. *Sur les fonctions de n variables complexes*. Acta Mathematica **19(1)**, pp. 1–61, 1895.

[Dav58] Martin Davis. *Computability & Unsolvability*. McGraw-Hill, 1958.

[DeM47] Augustus De Morgan. *Formal Logic*. Taylor and Walton, London, 1847.

[Den12] Arnaud Denjoy. *Une extension de l'intégrale de Lebesgue*. Comptes rendus de l'Académie des Sciences **134**, pp. 859–862, 1912.

[Dir29] P. G. L. Dirichlet. *Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données*. Journal für die reine und angewandte Mathematik **4**, pp. 157–169, 1829.

[DK11] Rodney G. Downey and Asher M. Kach. *Euclidean functions of computable Euclidean domains*. Notre Dame Journal of Formal Logic **52(2)**, pp. 163–172, 2011.

[Eve69] Howard Eves. *An Introduction to the History of Mathematics*. 3rd ed. Holt, Rinehart and Winston, 1969.

[FB58] Abraham A. Fraenkel and Yehoshua Bar-Hillel. *Foundations of Set Theory*. North Holland, Amsterdam, 1958.

[Fin49] N. J. Fine. *On the Walsh functions*. Transactions of the American Mathematical Society **65(3)**, pp. 372–372, 1949.

[FMS82] Harvey Friedman, Kenneth McAloon, and Stephen G. Simpson. *A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis*. In: George Metakides (ed.), *Patras Logic Symposion*. Studies in Logic and the Foundations of Mathematics **109**, North Holland, 1982.

[Fri67] Harvey Friedman. *Subsystems of set theory and analysis*. PhD thesis, MIT, 1967.

[Fri69] Harvey Friedman. *Bar induction and* Π^1_1-CA. The Journal of Symbolic Logic **34(3)**, pp. 353–362, 1969.

[Fri74] Harvey Friedman. *Some systems of second order arithmetic and their use*. In: *Proceedings of the International Congress of Mathematicians*. Vancouver, 1974.

[Fri76] Harvey Friedman. *Systems of second order arithmetic with restricted induction, I, II (abstracts)*. The Journal of Symbolic Logic **41**, pp. 557–559, 1976.

[FSS83] Harvey Friedman, Stephen G. Simpson, and Rick L. Smith. *Countable algebra and set existence axioms*. Annals of Pure and Applied Logic **25(2)**, pp. 141–181, 1983.

[Göd29] Kurt Gödel. *Über die Vollständigkeit des Logikkalküls*. PhD thesis, University of Vienna, 1929.

[Göd31] Kurt Gödel. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*. Monatshefte für Mathematik und Physik **38(1)**, pp. 173–198, 1931.

[Gor94] Russell A. Gordon. *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*. Graduate Studies in Mathematics **4**, American Mathematical Society, 1994.

[Gor96] Russell A. Gordon. *Is nonabsolute integration worth doing?* Real Analysis Exchange **22(1)**, pp. 23–33, 1996.

[HB34] David Hilbert and Paul Bernays. *Grundlagen der Mathematik*. Springer, Berlin, 1934.

[Hea93] Oliver Heaviside. *Electromagnetic Theory*. The Electrician Printing and Publishing Company, London, 1893.

[Hen63] R. Henstock. *Theory of Integration*. Butterworths, London, 1963.

[Hil99] David Hilbert. *Grundlagen der Geometrie*. B.G. Teubner, Leipzig, 1899.

[HM17] Jeffry L. Hirst and Carl Mummert. *Reverse mathematics of matroids*. In: Adam Day et al. (eds.), *Computability and Complexity: Essays Dedicated to Rodney G. Downey on the Occasion of His 60th Birthday*. Lecture Notes in Computer Science, Springer International Publishing, Cham, 2017.

[Joc72] Carl G. Jockusch. *Ramsey's theorem and recursion theory*. The Journal of Symbolic Logic **37(2)**, pp. 268–280, 1972.

[JS72] Carl G. Jockusch and Robert I. Soare. *Degrees of members of Π^0_1 classes*. Pacific Journal of Mathematics **40(3)**, pp. 605–616, 1972.

[KL90] A. S. Kechris and A. Louveau. *A classification of Baire class 1 functions*. Transactions of the American Mathematical Society **318(1)**, p. 209, 1990.

[Kle43] S. C. Kleene. *Recursive predicates and quantifiers*. Transactions of the American Mathematical Society **53(1)**, pp. 41–73, 1943.

[Ko91] Ker-I. Ko. *Complexity Theory of Real Functions*. Progress in Theoretical Computer Science, Birkhäuser, Boston, 1991.

[KS04] Douglas S. Kurtz and Charles W. Swartz. *Theories of Integration: The Integrals of Riemann, Lebesgue, Henstock-Kurzweil, and McShane*. World Scientific, 2004.

[KT14] Rutger Kuyper and Sebastiaan Terwijn. *Effective genericity and differentiability*. Journal of Logic and Analysis **6(4)**, pp. 1–14, 2014.

[Kur57] Jaroslav Kurzweil. *Generalized ordinary differential equations and continuous dependence on a parameter*. Czechoslovak Mathematical Journal **7(3)**, pp. 418–449, 1957.

[Leb02] Henri Lebesgue. *Intégrale, longueur, aire*. Annali di Matematica Pura ed Applicata **7(1)**, pp. 231–359, 1902.

[Leb04] Henri Lebesgue. *Leçons sur l'intégration et la recherche des fonctions primitives*. Gauthier-Villars, Paris, 1904.

[Lee11] Tuo Yeong Lee. *Henstock-Kurzweil Integration on Euclidean Spaces*. World Scientific, 2011.

[Lob29] Nikolai Lobachevsky. *A concise outline of the foundations of geometry*. University of Kazan Messenger, 1829.

[Luz12] Nikolai Luzin. *Sur les propriétés de l'intégrale de M. Denjoy*. Comptes rendus de l'Académie des Sciences **155**, pp. 1475–1477, 1912.

[Mar02] David Marker. *Model Theory: An Introduction*. Graduate Texts in Mathematics **217**, Springer-Verlag, New York, 2002.

[Mar96] Alberto Marcone. *On the logical strength of Nash-Williams' theorem on transfinite sequences*. In: Wilfrid Hodges et al. (eds.), *Logic: from foundations to applications*. Clarendon Press, New York, 1996.

[Men16] José Eduardo Mendiola. *Significance of the axiom of choice in mathematics*. 2016. URL: `https://www.researchgate.net/publication/303874180`

[Mil04] Joseph R. Mileti. *Partition Theorems and Computability Theory*. PhD thesis, University of Illinois, Urbana-Champaign, 2004.

[MN82] George Metakides and Anil Nerode. *The introduction of non-recursive methods into mathematics*. In: A. S. Troelstra and D. van Dalen (eds.), *The L. E. J. Brouwer Centenary Symposium, Noordwijkerhout*. Studies in Logic and the Foundations of Mathematics **110**, Elsevier, 1982.

[Mor01] Takakazu Mori. *Computabilities of Fine-continuous functions*. In: Jens Blanck, Vasco Brattka, and Peter Hertling (eds.), *Computability and Complexity in Analysis*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2001.

[Mor02] Takakazu Mori. *On the computability of Walsh functions*. Theoretical Computer Science **284(2)**, pp. 419–436, 2002.

[Odi99] P. Odifreddi. *Classical Recursion Theory, Volume II*. Studies in Logic and the Foundations of Mathematics **143**, North Holland, 1999.

[PC75] Marian B. Pour-El and J. Caldwell. *On a simple definition of computable function of a real variable—with applications to functions of a complex variable*. Zeitschrift für Mathematische Logik und Grundlagen der Mathematik **21**, pp. 1–19, 1975.

[PDD17] Michelle Porter, Adam Day, and Rodney G. Downey. *Notes on computable analysis*. Theory of Computing Systems **60(1)**, pp. 53–111, 2017.

[Pea89] Giuseppe Peano. *Arithmetices principia: nova methodo exposita*. Fratres Bocca, Rome, 1889.

[Per14] Oskar Perron. *Über den Integralbegriff*. Sitzungsberichte, Heidelberger Akademie der Wissenschaften, 1914.

[Pos44] Emil L. Post. *Recursively enumerable sets of positive integers and their decision problems*. Bulletin of the American Mathematical Society **50(5)**, pp. 284–316, 1944.

[PR89] Marian B. Pour-El and J. Ian Richards. *Computability in Analysis and Physics*. Perspectives in Mathematical Logic **1**, Springer-Verlag, Berlin, 1989.

[Rie54] Bernhard Riemann. *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*. Habilitation thesis, University of Göttingen, 1854.

[Rog67] Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, 1967.

[Sch61] H. A. Schwarz. *Lectures of Karl Weierstrass*. Mittag-Leffler Institute, Sweden, 1861.

[Sho05] Richard A. Shore. *Invariants, Boolean algebras and* ACA_0^+. Transactions of the American Mathematical Society **358(03)**, pp. 989–1014, 2005.

[Sim09] Stephen G. Simpson. *Subsystems of Second Order Arithmetic*. 2nd ed. Perspectives in Logic, Cambridge University Press, Cambridge, 2009.

[Sim87] Stephen G. Simpson. *Subsystems of* Z_2 *and reverse mathematics*. In: Gaisi Takeuti, *Proof Theory*. 2nd ed. Studies in Logic and Foundations of Mathematics **81**, North Holland, 1987.

[Soa16] Robert I. Soare. *Turing Computability*. Theory and Applications of Computability, Springer-Verlag, Berlin, Heidelberg, 2016.

[Soa87] Robert I. Soare. *Recursively Enumerable Sets and Degrees*. Perspectives in Mathematical Logic, Springer-Verlag, Berlin, Heidelberg, 1987.

[Sol01] Reed Solomon. Π^1_1-CA_0 *and order types of countable ordered groups*. The Journal of Symbolic Logic **66(1)**, pp. 192–206, 2001.

[Sol78] Robert M. Solovay. *Hyperarithmetically encodable sets*. Transactions of the American Mathematical Society **239**, pp. 99–122, 1978.

[Spe49] Ernst Specker. *Nicht konstruktiv beweisbare Sätze der Analysis*. The Journal of Symbolic Logic **14(3)**, pp. 145–158, 1949.

[ST90] Naoki Shioji and Kazuyuki Tanaka. *Fixed point theory in weak second-order arithmetic*. Annals of Pure and Applied Logic **47(2)**, pp. 167–188, 1990.

[Ste77] John Robert Steel. *Determinateness and subsystems of analysis*. PhD thesis, University of California, Berkeley, 1977.

[Tan91] Kazuyuki Tanaka. *Weak axioms of determinacy and subsystems of analysis II: Δ^0_2 games*. Annals of Pure and Applied Logic **52(1-2)**, pp. 181–193, 1991.

[Tra62] Grace Joy Traylor. *On the equivalence of the axiom of choice, Zorn's lemma, and the well-ordering theorem*. Masters thesis, Atlanta University, 1962.

[Tur37] Alan Turing. *On computable numbers, with an application to the Entscheidungsproblem*. Proceedings of the London Mathematical Society **s2-42(1)**, pp. 230–265, 1937.

[Vau86] Robert L. Vaught. *Alfred Tarski's work in model theory*. The Journal of Symbolic Logic **51(4)**, pp. 869–882, 1986.

[Wal23] J. L. Walsh. *A closed set of normal orthogonal functions*. American Journal of Mathematics **45(1)**, pp. 5–24, 1923.

[Zer30] Ernst Zermelo. *Über Grenzzahlen und Mengenbereiche: neue Untersuchungen über die Grundlagen der Mengenlehre*. Fundamenta Mathematicae **16**, pp. 29–47, 1930.