4. Second Order Arithmetic and Reverse Mathematics

4.1. **The Language of Second Order Arithmetic.**

We've mentioned that Peano arithmetic is sufficient to carry out large portions of ordinary mathematics, but with a qualifier, namely that Peano arithmetic suffices to carry out most proofs *regarding those statements it can express*. Up to now, we've emphasized that this is less of a limitation than it appears at first—while Peano arithmetic nominally talks about numbers, it can also encode other notions, like sequences, finite groups, and proof theory itself, and prove things about those.

But this encoding has a limit. The only objects we can discuss using Peano arithmetic are those which are in some sense finite, or at least finitely describable. This excludes large swaths of mathematics—almost all of topology and analysis, just to start with. We would therefore like to extend our system of logic to describe these notions as well.

For now, we will only move one step up, to include *second order* objects: we will add sets to our langauge.

Definition 4.1. The language of *second order arithmetic* is a two-sorted language: there are two kinds of terms, numeric terms and set terms.

- $\mathbf{0}$ is a numeric term,
- There are infinitely many numeric variables, $x_0, x_1, \ldots$, each of which is a numeric term
- If s is a numeric term then $\mathbf{S}s$ is a numeric term,
- If s, t are numeric terms then $+st$ and $\cdot st$ are numeric terms (abbreviated $s + t$ and $s \cdot t$),
- There are infinitely many set variables, $X_0, X_1, \ldots$, each of which is a set term
- If t is a numeric term and S then $\in tS$ is an atomic formula (abbreviated $t \in S$),
- If s and t are numeric terms then $= st$ and $< st$ are atomic formulas (abbreviated $s < t$).

The formulas are built from the atomic formulas in the usual way.

As the examples in the definition suggest, we use upper case letters for set variables and lower case letters for numeric terms. (Note that the only set terms are the variables.) Later, it will be more convenient to work with functions instead of sets, but within arithmetic, these are equivalent: we can use our pairing operation, and say that X represents a function if for each n there is exactly one m such that the pair (n, m) belongs to X.

We have to consider what we intend the semantics of this language to be. One possibility is the semantics of "true" second order logic: a model consists of a set M, representing the numeric objects, and interpretations of the various functions and relations (probably with the requirement that equality be the genuine equality relation), and a statement

$$\forall X \phi(X)$$

is satisfied by the model if for every possible subset of M, the corresponding statement holds.

"True" second order logic has no corresponding proof system. An easy way to see this is to observe that it has no compactness theorem. For example, the only model (up to isomorphism) of Peano arithmetic together with the second order induction axiom

$$\forall X(\mathbf{0} \in X \wedge \forall x(x \in X \to \mathbf{S}x \in X) \to \forall x\, x \in X).$$

is the standard model $\mathbb{N}$. This is easily seen: any model of Peano arithmetic has an initial segment isomorphic to $\mathbb{N}$; applying the induction axiom to this set, we see that it must be the whole of the model.

Consequently, this is not a very useful approach. An alternative semantics is to treat this as an ordinary two-sorted first order language: a model $\mathfrak{M}$ consists of two sets, M and M', where the elements of M are interpreted as "numbers" and the elements of M' are interpreted as "sets". There is no obligation for M' to actually consist of subsets of M, but we have an interpretation of $\in$, and we may associate each element S of M' with $\{m \in M \mid \mathfrak{M} \vDash m \in S\}$. Importantly, M' need not include all possible subsets of M. In particular, we can obtain nonstandard models of the second order induction axiom by making sure to only include those sets which do satisfy induction—and just leaving out the ones which don't.

In terms of proof theory, this means we simply treat the language of second order arithmetic as an ordinary first order language. In order to get any mileage out of the new content, though, we need axioms which state that there are any sets at all. We will only work in theories which have *comprehension axioms*, which typically have the form

$$\exists X \forall x(x \in X \leftrightarrow \phi(x))$$

where X does not appear free in the formula ϕ.

From here on, this will be our convention. In particular, whenever we use terms like "second order" or "higher order", we should remember that we are still within the confines of first order logic, and are referring only to the intended interpretation.

4.2. $\mathbf{Z}_2$.

Definition 4.2. $\mathbf{Z}_2$, the theory of *full second order arithmetic* consists of:

- The axioms P^-,
- The second order induction axiom

  $$\forall X(\mathbf{0} \in X \wedge \forall x(x \in X \to \mathbf{S}x \in X) \to \forall x\, x \in X),$$

- The *comprehension axiom*

  $$\exists X \forall x(x \in X \leftrightarrow \phi(x))$$

  for each formula ϕ where X does not appear free.

$\mathbf{Z}_2$ is sometimes called the theory of analysis, in contrast to **PA** as the theory of arithmetic, because $\mathbf{Z}_2$ can talk about real numbers in a coherent way.

First, we need to talk about integers and rational numbers. For the moment, let us simply consider coding in $\mathbb{N}$, ignoring any issues of provability or first order logic. We can encode the integers in a standard way: by an element of $\mathbb{Z}$, we mean an element of $\mathbb{N}$ viewed as a pair (n, m), taken to represent $n - m$. Each integer has many representatives under this scheme, so we write $(n, m) \equiv_{\mathbb{Z}} (n', m')$ if $n + m' = n' + m$. We can define the standard operations on the integers in a natural way: $-(n, m) = (m, n)$, $(n, m)+(n', m') = (n+n', m+m')$, $(n, m)\cdot(n', m') = (nn'+mm', nm'+mn')$. All these operations are of course well-defined on $\mathbb{Z}$—they map different representations of the same integer to representations of the same integer.

We can then represent the rationals $\mathbb{Q}$ by pairs of integers (a, b) with $b > 0$: (a, b) represents a/b. Then $(a, b) \equiv_{\mathbb{Q}} (c, d)$ if $ad = bc$. (Of course, equality here means $\equiv_{\mathbb{Z}}$.) Again, addition, multiplication, and so on can be defined on these representations in a natural way.

Now we observe that all the relevant definitions can be written down using first order arithmetic, and their properties can all be proven in **PA** (and even, as will be convenient later, $I\Sigma_1$). Formally, what we mean by this is that ordinary statements about the integers and the rationals can be unfolded into (much more complicated) statements in the language of **PA**. We give one example; fix formulas $\rho_1(z, x)$ and $\rho_2(z, y)$ which will both hold exactly when z is the pair (x, y); formally, we should be able to deduce the formulas:

- $\forall x, y \exists z \rho_1(z, x) \wedge \rho_2(z, y)$,
- $\forall z \forall x, x' (\rho_1(z, x) \wedge \rho_1(z, x') \rightarrow x = x')$,
- $\forall z \forall y, y' (\rho_2(z, y) \wedge \rho_2(z, y') \rightarrow y = y')$.

Then the statement that addition on integers is commutative,

$$\forall s \in \mathbb{Z} \forall t \in \mathbb{Z} (s +_{\mathbb{Z}} t =_{\mathbb{Z}} t +_{\mathbb{Z}} s)$$

becomes

$$\forall s \forall t \forall n, m, m', n' (\rho_1(s, n) \wedge \rho_2(s, m) \wedge \rho_1(t, n') \wedge \rho_2(t, m') \rightarrow (n + n') + (m' + m) = (n' + n) + (m + m')).$$

Here the first part—the conjunction on the ρ—gives us a representation of s, t as pairs, and checks that s, t really do count as integers (that is, are actually pairs of natural numbers); formally, this is the translation of the restriction of the quantification to be over the integers. The term $s +_{\mathbb{Z}} t$ then translates to $(n + n', m + m')$ while $t +_{\mathbb{Z}} s$ translates to $(n' + n, m' + m)$, and then the formula stating that the terms are equal is the right side of the implication.

To talk about real numbers, we'll need to use sets. There are two standard ways of defining the real numbers from the rationals: as Dedekind cuts, and

as Cauchy sequences. For technical reasons, the latter is more convenient for us.

Definition 4.3. In second order arithmetic, we say a set R is a *function* on a set D if for every $x \in D$ there is exactly one y such that $(x, y) \in R$. We say a set R is a *infinite sequence* if for every x there is exactly one y such that $(x, y) \in R$. We often write $R = \langle m_n \rangle$ where for each n, m_n is the number with $(n, m_n) \in R$.

In second order arithmetic, a *real number* is a sequence of rational numbers $\langle q_n \rangle$ such that

$$\forall q \in \mathbb{Q}(q > 0 \to \exists m \forall n(m < n \to |q_m - q_n| < q)),$$

It requires some unfolding, but we emphasize that this is a formula in second order arithmetic. As with integers and rationals, two different sets can encode the same real number, but we can write down formulas indicating when two real numbers are equal, when one is larger than the other, how to add real numbers, and so on.

Note that $\mathbf{Z}_2$ does not prove that "arbitrary" real numbers exist (whatever that means). Rather, if we want to work with a real number, we have to prove, within $\mathbf{Z}_2$, that a corresponding representation exists. It is easy to show that each rational number gives the corresponding real number—take $\langle q \rangle$. Further, all real numbers with "easy" descriptions exist. For instance, to encode $\sqrt{2}$ we can take q_n to be the largest rational of the form $m/10^n$ such that $(m/10^n)^2 < 2$ (in other words, q_n is the first $n + 1$ digits of the decimal expansion of $\sqrt{2}$).

Similar coding efforts suffice to code most countable objects—countable groups, for instance—and further, are sufficient to discuss most *separable* domains (that is, situations where a countable subset is "dense" in some way). Furthermore, the theory $\mathbf{Z}_2$ is very strong; it suffices to prove almost any theorem it can state which has been proven in the mathematics literature. However second order arithmetic cannot state everything—for instance, it still can't express statements about genuinely uncountable mathematics, including things like set theory and portions of abstract analysis—and there are statements (all known examples coming from logic) which can be stated in second order arithmetic but are independent of $\mathbf{Z}_2$.

There is a natural stratification of formulas in second order arithmetic similar to the one in first order arithmetic.

Definition 4.4. The Δ^0_0 formulas are those in which all quantifiers over numeric variables are bounded and there are no quantifiers over set variables. Σ^0_0 and Π^0_0 are alternate names for Δ^0_0.

The Σ^0_{n+1} formulas are formulas of the form

$$\exists x \phi$$

(possibly with a block of several existential quantifiers) where ϕ is Π^0_n.

The Π^0_{n+1} formulas are formulas of the form

$$\forall x \phi$$

(possibly with a block of several universal quantifiers) where ϕ is Σ^0_n.

The *arithmetic formulas*, sometimes written Π^0_∞, Δ^1_0, Σ^1_0, or Π^1_0, are those formulas with no quantifiers over set variables.

The Σ^1_{n+1} formuals are formulas of the form

$$\exists X \phi$$

(possibly with a block of several existential quantifiers) where ϕ is Π^1_n.

The Π^1_{n+1} formulas are formulas of the form

$$\forall X \phi$$

(possibly with a block of several universal quantifiers) where ϕ is Σ^1_n.

The Π^0_n and Σ^0_n formulas are essentially our old Π_n and Σ_n formulas (except that they can contain set *variables*, but not set quantifiers). The Σ^1_n and Π^1_n hierarchies are analagous, but with set quantifiers rather than numeric ones. Note the (largely accurate) implication that even a single set quantifier "trumps" any number of numeric quantifiers.

4.3. $\mathbf{ACA_0}$.

It turns out that $\mathbf{Z}_2$ is *very* strong, and we would like to consider some weaker theories.

Definition 4.5. The axioms of $\mathbf{ACA_0}$ consist of:

- The axioms of P^-,
- The second order induction axiom,
- The comprehension scheme restricted to arithmetic formulas.

$\mathbf{ACA_0}$ stands for "arithmetic comprehension axiom". The subscript 0 means that instead of the induction axiom for arbitrary formulas, we have only the second order induction axiom. In $\mathbf{Z}_2$, this didn't matter—every formula defined a set, so it didn't matter whether induction was over formulas or sets. In $\mathbf{ACA_0}$, the difference is significant—there are plenty of formulas we can write down using set quantifiers which are perfectly sensible formulas, but since the comprehension axiom doesn't apply, they don't define sets, and therefore we do not have a corresponding induction axiom.

Unlike $\mathbf{PA}$, the system $\mathbf{ACA_0}$ can express statements about countable sets, and objects encoded by countable sets. For an elementary example, consider the following basic statement from real analysis:

Theorem 4.6 (Bolzano-Weierstraß Theorem)**.** *If $\langle r_n \mid n \in \mathbb{N} \rangle$ is a bounded infinite sequence of real numbers then there is a subsequence $n_0 < \cdots < n_k < \cdots$ such that $\lim_{k\to\infty} r_{n_k}$ exists.*

This statement is outside the domain of first order arithmetic, but it can be expressed fairly naturally in the language of second order arithmetic. To talk about infinite sequences of real numbers, we can just use pairing: S is

an infinite sequence of real numbers if for each n, $S_n = \{m \mid (n, m) \in S\}$ is a real number.

Theorem 4.7. $\mathbf{ACA_0}$ *proves the Bolzano-Weierstraß Theorem.*

Proof. Let S be a sequence of real numbers, so that for each n, S_n is a real number. Let b be an integer such that for each n, $|S_n| \leq b$. In general a sequence of real numbers could have many limits, and it will be easiest to pick out a natural one, namely the lim sup.

We will define a set T to consist of those pairs (n, k) such that S_n is within 2^{-k} of the lim sup of the sequence. More precisely, $(n, k) \in T$ if there are only finitely many m such that $S_n + 2^{-k} < S_m$; this is an arithmetic formula, so $\mathbf{ACA_0}$ proves that the set T exists.

We claim that for each k, there are infinitely many n such that $(n, k) \in T$. Suppose not, and pick some k with only finitely many n so $(n, k) \in T$. We construct a sequence as follows: take r_1 to be any element so $(r_1, k) \notin T$, and given r_i with $(r_i, k) \notin T$, there are infinitely many m with $S_{r_i} + 2^{-k} < S_m$, only finitely many of which satisfy $(m, k) \in T$, so take r_{i+1} so that $S_{r_i} + 2^{-k} < S_{r_{i+1}}$ and $(r_{i+1}, k) \notin T$. For each i we have

$$-b \leq S_{r_1} < S_{r_2} - 2^{-k} < S_{r_3} - 2 \cdot 2^{-k} < \cdots S_{r_i} - i \cdot 2^{-k},$$

so by choosing $i > b2^{k+1}$, we have $S_{r_i} - i \cdot 2^{-k} > -b$, which means $S_{r_i} > b$, contradicting the assumption that $|S_n| \leq b$ for all b.

We should check that $\mathbf{ACA_0}$ can actually prove that the sequence $r_i, \ldots, r_i$ exists. In $\mathbf{PA}$, we would prove the existence using the induction axiom—we would take the statement "there exists a sequence $\vec{r}$ of length i so that for each $j \leq i$, $(r_j, k) \notin T$ and for each $j < i$, $S_{r_j} + 2^{-k} < S_{r_{j+1}}$", observe that it holds trivially for length 0 (the empty sequence), and that if there is a sequence of length i, there is also a sequence of length $i+1$, which the argument above shows must exist. Similarly, in $\mathbf{ACA_0}$ we can consider the set I of i such that there is a sequence of length i with the desired properties, and argue that $0 \in I$ and whenever $i \in I$, $i + 1 \in I$; therefore every number is in I, so in particular $b2^{k+1} + 1 \in I$.

Having shown that for each k, there are infinitely many n with $(n, k) \in T$, we define a sequence by setting r_k to be least such that $(r_k, k) \in T$ and $r_k > r_{k'}$ for $k' < k$. Again, to formalize this in $\mathbf{ACA_0}$, we consider the sequences $\vec{r}$ of length i such that for each $k \leq i$, r_k is least with $(r_k, k) \in T$ and $r_{k'} < r_k$ for each $k' < k$. The statement that there exists such a sequence is arithmetic, so we take I to be the set of i such that there is such a sequence of length i; then $0 \in I$ since the empty sequence has the desired property, and if there is a sequence of length i, there must also be a sequence of length $i + 1$, so there are such sequences of every length. By a similar inductive argument, we can show that there is a unique such sequence of any given length. Finally, observe that $\mathbf{ACA_0}$ also proves the existence of the set R of (k, r_k) such that r_k is the k-th element in the unique such sequence of length k.

Now that we have our convergent subsequence, constructing the limit is easy: a real number is a convergent sequence of rationals. Recall that each $S_n = \langle S_{n,i} \rangle$ is a sequence of rational numbers converging to S_n, so choose q_k so that $q_k = S_{r_k,i_k}$ where i_k is chosen so that $|S_{r_k,i_k} - S_{r_k}| < 2^{-k}$. Then each q_k is within $2^{-(k-1)}$ of the lim sup of the sequence $\langle S_n \rangle$, so in particular $\langle q_k \rangle$ converges to the lim sup. □

Provability in $\mathbf{ACA_0}$ has been very thoroughly studied. $\mathbf{ACA_0}$ can express the notion of a continuous function (but *not* the notion of an arbitrary function on the real numbers), and therefore much of elementary real analysis. $\mathbf{ACA_0}$ can also describe notions like countable groups, fields, and vector spaces, and prove many standard theorems on these objects.

We only discuss one more family of results in detail, coming from infinitary combinatorics.

Theorem 4.8 (Infinite Pigeonhole Principle)**.** *Let S be an infinite set, let f be a function on S and let T be a finite set such that for every $s \in S$, $f(s) \in T$. Then there is an infinite set $S' \subseteq S$ and a $t \in T$ so that for every $s \in S'$, $f(s) = t$.*

Note that the notion of a set being infinite can be expressed in second order arithmetic by saying that for every n, there is an $m > n$ with $m \in S$. Conversely, to say that a set is finite is to say that there is some n such that every $m \in S$ is $\leq n$.

Proof. Suppose not; in particular, suppose that for each $t \in T$, $S_t = \{s \in S \mid f(s) = t\}$ is finite.

Then for each t, let us define $U_t \subseteq S$ to be the set of $n \in S$ such that for every $m > n$ with $m \in S$, $f(m) \geq t$. Let U be the set of t such that U_t is infinite. Since $U_0 = S$, certainly U_0 is infinite, so $0 \in U$. Suppose $t \in U$, so U_t is infinite. $U_t = S_t \cup U_{t+1}$; since S_t is finite, there is an n so that every element of S is $\leq n$. So for any n', we may choose an $m > \max\{n', n\}$ with $m \in U_t$, and since $m \notin S_t$, we have $m \in U_{t+1}$. So U_{t+1} is infinite, and therefore $t + 1 \in U$.

By the second order induction axiom, every element is in U. In particular, take n large enough that every element of T is $< n$. Then U_n is infinite, so there is an $m \in U_n$; but then $f(m)$ is larger than every element of T. This is a contradiction. □

Definition 4.9. Let $\sigma = \langle s_0, s_1, \ldots, s_n \rangle$ be a finite sequence. An *initial segment* of σ is the sequence $\langle s_0, \ldots, s_m \rangle$ for some $m \leq n$ (including $m = n$ and the empty sequence, where $m = -1$). An *immediate extension* of σ is a sequence $\langle s_0, \ldots, s_n, t \rangle$ for any t.

If $\langle s_n \rangle$ is an infinite sequence, an initial segment is a finite sequence $\langle s_0, \ldots, s_m \rangle$.

A *tree* is a set S of finite sequences with the property that whenever $\sigma \in S$, every initial segment of σ belongs to S.

Theorem 4.10 (König's Lemma)**.** *Suppose that S is a tree of sequences of integers such that S contains infinitely many elements, but each sequence in S has only finitely many immediate extensions in S. Then there is an infinite sequence $\langle s_n \rangle$ such that all its initial segments are in S. Furthermore, this can be proven in* $\mathbf{ACA_0}$.

Proof. First, we consider the proof in general, without worrying about $\mathbf{ACA_0}$. The idea is that we take the infinite sequence $\langle s_n \rangle$ so that each of its initial segments is also an initial segment of infinitely many elements of S. For definiteness, we simply take the s_n least such that this is the case.

Formally, suppose we have defined $s_0, \ldots, s_m$ (possibly with $m = -1$ for the base case) such that $\langle s_0, \ldots, s_m \rangle$ is an initial segment of infinitely many elements of S. (This is true by assumption in the base case, since the empty sequence is an initial segment of all sequences, and so in particular of all the infinitely many elements of S.) By assumption, there are finitely many $t_1, \ldots, t_r$ such that $\langle s_0, \ldots, s_m, t_i \rangle \in S$. Consider those elements τ of S such that $\langle s_0, \ldots, s_m \rangle$ is an initial segment of τ, and which are not equal to $\langle s_0, \ldots, s_m \rangle$. This is an infinite set $S' \subseteq S$. For each $\tau \in S'$, there is exactly one $i \leq r$ such that $\langle s_0, \ldots, s_m, t_i \rangle$ is an initial segment of τ, so we can define a function $f : S' \to [1, r]$. By the infinite pigeonhole principle, there is an $i \leq r$ so that $f(\tau) = i$ for infinitely many $\tau \in S'$; we take $s_{m+1} = t_i$.

To formalize this argument in $\mathbf{ACA_0}$, we first show that we can write down an arithmetic formula (using S as a parameter) defining an example of the desired sequence. It is convenient to do this in two steps. First, we simply define S^* to be the set of sequences $\langle s_0, \ldots, s_m \rangle \in S$ such that there are infinitely many $\tau \in S$ extending $\langle s_0, \ldots, s_m \rangle$; this is easily expressed with a few numeric quantifiers ("for every r there is a $t > r$ coding a sequence extending $\langle s_0, \ldots, s_m \rangle$").

We wish to show that for every $\sigma \in S^*$, there is an immediate extension of $\sigma \in S^*$. To see this, let $\sigma \in S^*$ be given; then $\sigma \in S$ and there are infinitely many extensions of $\sigma \in S$. We carry out the argument above: let $S_\sigma = \{\tau \mid \sigma \text{ is an initial segment of } \tau\}$, let S^0_σ be the set of immediate extensions of σ in S, and define the coloring $f : S_\sigma \to S^0_\sigma$. Since S^0_σ is finite, there is an element of S^0_σ which is the value of f infinitely often, so this value of S^0_σ belongs to S^*.

Now we define our actual infinite sequence, T, to consist of those pairs (m, s_m) such that there is a $\langle s_0, \ldots, s_m \rangle \in S^*$ such that there is no $i \leq m$ and $t < s_i$ such that $\langle s_0, \ldots, s_{i-1}, t \rangle \in S^*$. The definition ensures that for each m, there is at most one s_m such that $(m, s_m) \in T$. To see that there is such an s_m, let $T^* = \{m \mid \exists s_m (m, s_m) \in T\}$; by the induction axiom, it suffices to show that $0 \in T^*$ and whenever $m \in T^*$, also $m + 1 \in T^*$. The arguments are essentially the same; suppose $(0, s_0), \ldots, (m, s_m) \in T$. Then $\langle s_0, \ldots, s_m \rangle = \sigma \in S^*$, so there is some t with $\langle s_0, \ldots, s_m, t \rangle \in S^*$. Let U be the set of t such that for all $t' < t$, $\langle s_0, \ldots, s_m, t' \rangle \notin S^*$; clearly $0 \in U$, but there must be some $t \notin U$—take $t + 1$ where $\langle s_0, \ldots, s_m, t + 1 \rangle \in S^*$, which

exists since every element of S^* has an immediate extension in S^*. Since the induction axiom applies to U, there must be a $t \in U$ such that $t+1 \notin U$. Therefore $\langle s_0, \ldots, s_m, t\rangle \in S^*$, but for every $t' < t$, $\langle s_0, \ldots, s_m, t'\rangle \notin S^*$, so $(m+1, t) \in T$. □

4.4. **The Strength of $\mathbf{ACA_0}$.**

$\mathbf{ACA_0}$ corresponds very closely to **PA**. In fact, the first order consequences of $\mathbf{ACA_0}$ are exactly the theorems provable in **PA**.

Theorem 4.11. *If* $\mathbf{PA} \vdash \phi$ *then* $\mathbf{ACA_0} \vdash \phi$.

Proof. Easily follows from the fact that every axiom of **PA** is deducible in $\mathbf{ACA_0}$. □

Theorem 4.12. *If* $\mathbf{ACA_0} \vdash \phi$ *where* ϕ *is a formula in the language of Peano arithmetic then already* $\mathbf{PA} \vdash \phi$.

In order to prove this, we need some lemmas.

Definition 4.13. Let X be a second order variable, and let $\phi(x)$ be a formula not containing X with a distinguished free variable x. Then we define $\psi[\phi/X]$ recursively by:

- $(t \in X)[\phi/X]$ is $\phi(t)$,
- If ψ is atomic and not $t \in X$ then $\psi[\phi/X]$ is ψ,
- $(\psi \circledast \theta)[\phi/X]$ is $\psi[\phi/X] \circledast \theta[\phi/X]$,
- $(Qy\psi)[\phi/X]$ is $Qy\psi[\phi/X]$,
- $(QX\psi)[\phi/X]$ is $QX\psi$,
- $(QY\psi)[\phi/X]$ is $QY\psi[\phi/X]$ if $Y \neq X$.

Lemma 4.14. *Suppose there is a cut-free deduction of* $\Gamma_{\mathbf{ACA_0}}, \Gamma \Rightarrow \Sigma$ *where* Γ, Σ *are arithmetic. Suppose also that* ϕ *is substitutable for* X *in* Γ, Σ. *Then there is a cut-free deduction of* $\Gamma_{\mathbf{ACA_0}}, \Gamma_{PA}, \Gamma[\phi/X] \Rightarrow \Sigma[\phi/X]$ *of size no larger than the original deduction.*

Proof. By induction on deductions. The only cases which do not follow immediately from the inductive hypothesis are the quantifiers rules $L\exists$, $R\forall$, and $L\forall$.

We first consider the $R\forall$ case. This must be a first order variable, since there are no set quantifiers on the right-hand side of the sequent. If the eigenvariable appears in ϕ we we apply the substitution lemma to replace the eigenvariable with a fresh variable, and apply the same $R\forall$ inference. In the $L\exists$ case, if we are considering a first order variable, the same argument applies. (In both cases, we also need to consider the possibility that the quantifier binds a variable in ϕ; this is prevented by the substitutability assumption.)

If we are substituting for a second order variable, either X is not the eigenvariable, in which case the same argument applies, or X is the eigenvariable, in which case X does not appear in Γ or Σ, so no changes are required.

In the $L\forall$ case, if the main variable is not X, the claim follows immediately from the inductive hypothesis. So consider the difficult case:

$$\dfrac{\dfrac{\dfrac{}{}}{\Gamma_{\mathbf{ACA_0}}, \Gamma, \psi[X/Y] \Rightarrow \Sigma}}{\Gamma_{\mathbf{ACA_0}}, \Gamma, \forall Y \psi \Rightarrow \Sigma}$$

In this case, the inductive hypothesis gives us a deduction of

$$\Gamma_{\mathbf{ACA_0}}, \Gamma[\phi/X], \psi[\phi/Y] \Rightarrow \Sigma[\phi/X].$$

We must have $\forall Y \psi \in \Gamma_{\mathbf{ACA_0}}$, so $\forall Y \psi$ is an induction axiom. But then $\psi[\phi/Y]$ is an axiom in Γ_{PA}, so we have

$$\Gamma_{\mathbf{ACA_0}}, \Gamma_{PA}, \Gamma[\phi/X] \Rightarrow \Sigma[\phi/X].$$

□

Lemma 4.15. *Suppose we have a cut-free deduction of* $\Gamma_{\mathbf{ACA_0}}, \Gamma \Rightarrow \Sigma$ *where* Γ, Σ *are arithmetic. Let* Γ', Σ' *be the result of replacing all free second order variables in* Γ, Σ *with* $\bot$*. Then there is a deduction of* $\Gamma_{PA}, \Gamma' \Rightarrow \Sigma'$ *where all cuts are over arithmetic formulas.*

Proof. By induction on the size of the deduction. The only non-trivial cases are quantifier rules with second order variables, and since Σ is arithmetic, all such formulas appear on the left-hand side. Suppose we are in the second order $L\exists$ case:

$$\dfrac{\dfrac{\dfrac{}{}}{\Gamma_{\mathbf{ACA_0}}, \Gamma, \psi \Rightarrow \Sigma}}{\Gamma_{\mathbf{ACA_0}}, \Gamma', \exists X \psi \Rightarrow \Sigma'}$$

Since the formulas in Γ are arithmetic, $\exists X \psi$ must be an axiom of $\mathbf{ACA_0}$, and so must be a comprehension axiom

$$\exists X \forall x (x \in X \leftrightarrow \phi(x))$$

for some arithmetic formula ϕ. We may apply the previous lemma to substitute ϕ for X, obtaining a deduction of

$$\Gamma_{\mathbf{ACA_0}}, \Gamma_{PA}, \Gamma, \psi[\phi/X] \Rightarrow \Sigma.$$

(Since X is the eigenvariable, it does not appear in $\Gamma\Sigma$.) This deduction is no larger than the original, so we may apply IH, obtaining a deduction of

$$\Gamma_{PA}, \Gamma', \psi[\phi/X]' \Rightarrow \Sigma'.$$

Since $\psi[\phi/X]'$ is derivable, we may apply a cut to obtain a deduction of

$$\Gamma_{PA}, \Gamma' \Rightarrow \Sigma'.$$

Suppose we are in the second order $L\forall$ case:

$$\dfrac{\dfrac{\dfrac{}{}}{\Gamma_{\mathbf{ACA_0}}, \Gamma, \psi \Rightarrow \Sigma}}{\Gamma_{\mathbf{ACA_0}}, \Gamma', \forall X \psi \Rightarrow \Sigma'}$$

Then $\forall X \psi$ must be an induction axiom, so we may substitute X with any formula—say, $\bot$—and apply the previous lemma to obtain a deduction of

$$\Gamma_{\mathbf{ACA_0}}, \Gamma_{PA}, \Gamma[\bot/X], \psi[\bot/X] \Rightarrow \Sigma[\bot/X].$$

SInce $\psi[\bot/X]$ is an instance of the first order induction axiom, this is already

$$\Gamma_{\mathbf{ACA_0}}, \Gamma_{PA}, \Gamma[\bot/X] \Rightarrow \Sigma[\bot/X].$$

We may then apply IH to obtain a deduction of

$$\Gamma_{PA}, \Gamma' \Rightarrow \Sigma'.$$

□

Proof of 4.12. Suppose $\mathbf{ACA_0} \vdash \Gamma \Rightarrow \Sigma$ where Γ, Σ do not contain any second order variables. By the cut-elimination theorem, there must be a cut-free deduction

$$\vdash \Gamma_{\mathbf{ACA_0}}, \Gamma \Rightarrow \Sigma.$$

We simply apply the previous lemma, obtaining a deduction of $\Gamma_{PA}, \Gamma' \Rightarrow \Sigma'$. Since Γ, Σ do not contain any second order variables, $\Gamma' = \Gamma$, $\Sigma' = \Sigma$, so we are done. □

4.5. **Reverse Mathematics.**

Is $\mathbf{ACA_0}$ *necessary* to prove the examples above? Put another way, we know from our work on cut-elimination that there is a natural notion of "strength" of a mathematical theory, namely the proof-theoretic ordinal. Second order arithmetic turns out to be a good place to quantify the strength of mathematical statements.

To make this precise, we will fix a weak "base theory", and ask the what the consequences of some interesting mathematical statement σ are when using the base theory. (Some base theory of arithmetic is needed; for instance, in the absence of enough arithmetic to prove that the rationals and reals have their intended properties, it's hard to argue that the statement even means what it is supposed to mean.)

Definition 4.16. The theory $\mathbf{RCA_0}$ consists of:

- The axioms of P^-,
- $I\Sigma_1$,
- The second order induction axiom

$$\forall X(\mathbf{0} \in X \wedge \forall x(x \in X \to \mathbf{S}x \in X) \to \forall x\, x \in X).$$

- The *recursive comprehension scheme*: if $\phi(x)$ is a Σ_1 formula and $\psi(x)$ is a Π_1 formula then

$$\forall x(\phi(x) \leftrightarrow \psi(x)) \to \exists X \forall x(x \in X \leftrightarrow \phi(x))$$

is an axiom.

Note that $\mathbf{RCA_0}$ has slightly more induction than its comprehension scheme suggests; we have to add induction over Σ_1 formulas. It turns out that without the additional induction axiom, the system is too weak even to use as a base theory for many purposes.

One might wonder why, instead, we don't add comprehension for Σ_1 formulas. The answer is that this theory is equivalent to $\mathbf{ACA_0}$! The reason is that we can use multiple steps: if we want $\{x \mid \forall y \exists z \phi(x, y, z)\}$, we can first define S_1 to be $\{(x, y) \mid \exists z \phi(x, y)\}$, define S_2 to be $\{x \mid \exists y \, (x, y) \notin S_1\}$, and finally our desired set is $\{x \mid x \notin S_2\}$.

In particular, this means that to show that an extension of $\mathbf{RCA_0}$ includes $\mathbf{ACA_0}$, it suffices to show that it has comprehension for Σ_1 formulas with parameters.

In $\mathbf{RCA_0}$, the definition of a real number we gave above does not work. We should define a real number to by a Cauchy sequence which converges *with a specified rate of convergence.*

Definition 4.17. In $\mathbf{RCA_0}$, a real number is a sequence of rational numbers $\langle q_n \rangle$ such that

$$\forall k \in \mathbb{N} \forall m, n (k < m < n \to |q_m - q_n| < 2^{-k}).$$

In $\mathbf{ACA_0}$, this is equivalent to the original definition, in the sense that $\mathbf{ACA_0}$ proves that every number in the original sense is equal to a real number in this stronger sense. On the other hand, $\mathbf{RCA_0}$ cannot prove useful properties of real numbers in the sense of $\mathbf{ACA_0}$.

The next two theorems are typical results in the area known as Reverse Mathematics.

Theorem 4.18. *Over* $\mathbf{RCA_0}$*, the Bolzano-Weierstraß Theorem implies* $\mathbf{ACA_0}$.

We mean the Bolzano-Weierstraß Theorem for real numbers in the sense of $\mathbf{RCA_0}$, of course.

Proof. Fix some Σ_1 formula $\exists x \phi(x, y)$. We wish to show that $\{y \mid \exists x \phi(x, y)\}$ actually exists. The idea is that we'll define a sequence of rational numbers which is already convergent, and whose limit encodes exactly this set. More precisely, we will arrange to converge to

$$\sum_{y | \exists x \phi(x,y)} 2^{-y}.$$

We define $q_n = \sum_{y \leq n, \exists x \leq n \phi(x,y)} 2^{-y}$. Observe that the sequence $\langle q_n \rangle$ *does* exist in $\mathbf{RCA_0}$, since it is given by a formula with only bounded quantifiers. Therefore the corresponding sequence of real numbers $S_n = q_n$ exists as well.

If the Bolzano-Weierstraß Theorem holds, this sequence has a limit $S = \langle q'_n \rangle$. If we want to know whether $\exists x \phi(x, y)$, we can simply look at the binary expansion of q'_{y+1}; q'_{y+1} is within $2^{-(y+1)}$ of the final value of the

sequence, so there is any x such that $\exists x \phi(x, y)$ exactly if the y-th digit in the expansion of q'_{y+1} is 1. □

This means that the Bolzano-Weierstraß Theorem is *equivalent* to $\mathbf{ACA_0}$ (given that we are working in at least $\mathbf{RCA_0}$).

Theorem 4.19. *Over* $\mathbf{RCA_0}$*, König's Lemma implies* $\mathbf{ACA_0}$*.*

Proof. Again, fix some Σ_1 formula $\exists x \phi(x, y)$. We define a tree of sequences by saying that $\sigma = (s_0, \ldots, s_n) \in T$ if for each $i \leq n$, either:

- $s_i = 0$ and there is no $x \leq n$ such that $\phi(x, i)$ holds, or
- $s_i > 0$ and $s_i - 1$ is least such that $\phi(s_i - 1, i)$.

Again, notice that this is given by a formula with only bounded quantifiers.

This is clearly finitely branching; actually, each σ has at most two extensions—one by 0 and one by the least $s_i - 1$ with $\phi(s_i - 1, i)$. To see that this is infinite, it suffices to show that for each n, there is a sequence $\langle s_0, \ldots, s_n \rangle \in T$; we may take this to be the sequence where s_i is $x + 1$ where $x \leq n$ is least such that $\phi(x, i)$, if there is any, and 0 otherwise. This sequence always exists, and always belongs to T.

By König's Lemma, there is an infinite sequence $\langle s_n \rangle$ through this. We claim that $\exists x \phi(x, y)$ iff $s_y \neq 0$. Suppose $\exists x \phi(x, y)$; then there is a least such x. The sequence $\langle s_0, \ldots, s_{y-1}, 0, \ldots, s_x \rangle \notin T$, so we cannot have $s_y = 0$. On the other hand, suppose $\forall x \phi(x, y)$. Then $\langle s_0, \ldots, s_y \rangle \in T$ implies $s_y = 0$ since the second clause can never be satisfied. □

Just as the first order part of $\mathbf{ACA_0}$ is exactly $\mathbf{PA}$, it can be shown that the first order part of $\mathbf{RCA_0}$ is exactly $I\Sigma_1$.

4.6. $\mathbf{WKL_0}$ and other theories weaker than $\mathbf{ACA_0}$.

One of the surprising discoveries in the early investigation of reverse mathematics was that statements from analysis and algebra tended to group into a small number of categories. In fact, almost all the early theorems investigated turned out to be equivalent to one of five theories, the so-called Big 5 theories of reverse mathematics. $\mathbf{RCA_0}$ and $\mathbf{ACA_0}$ are two of them.

Definition 4.20. The theory $\mathbf{WKL_0}$ consists of the axioms of $\mathbf{RCA_0}$ together with the following axiom:

> Suppose S is a tree of sequences of 0's and 1's containing infinitely many elements. Then there is an infinite sequence $\langle s_n \rangle$ such that all its initial segments are in S.

This axiom is called "weak König's lemma". It is clearly a special case of König's lemma, so $\mathbf{WKL_0}$ is between $\mathbf{RCA_0}$ and $\mathbf{ACA_0}$. To see that $\mathbf{WKL_0}$ is actually stronger than $\mathbf{RCA_0}$ is typically an exercise in computability: it is easy to show that there is a model of $\mathbf{RCA_0}$ where the sets are exactly the computable sets, and one then shows that is not a model of $\mathbf{WKL_0}$. On the other hand, it can be shown that not only is $\mathbf{WKL_0}$ weaker than $\mathbf{ACA_0}$, it's actually quite close to $\mathbf{RCA_0}$.

Theorem 4.21. *If* $\mathbf{WKL_0} \vdash \phi$ *where* ϕ *is arithmetic then* $\mathbf{RCA_0} \vdash \phi$.

There are several proofs of this theorem; the simplest one, due to Harrington, uses forcing to show that any model of $\mathbf{RCA_0}$ can be extended to a model of $\mathbf{WKL_0}$ by adding additional sets. An alternate proof, due to Ferreira and Ferreira, uses cut-elimination. It immediately follows that the same holds for Π^1_1 sentences, since deducing $\forall X \phi(X)$ is equivalent to deducing $\phi(X)$ with X free.

One consequence is that the proof-theoretic ordinal of $\mathbf{WKL_0}$ is the same as that of $\mathbf{RCA_0}$ (ω^ω); this is easily seen since stating that an ordinal is well-founded is a Π^1_1 sentence. Nonetheless, $\mathbf{WKL_0}$ proves many theorems about the existence of sets that $\mathbf{RCA_0}$ does not. For example:

Theorem 4.22. *The statement that every consistent theory has a complete consistent extension is equivalent to* $\mathbf{WKL_0}$ *(using* $\mathbf{RCA_0}$ *as a base theory).*

Proof. Recall that we can formalize all our proof theory, interpreting formulas by sequences of symbols, which are in turn coded by numbers, and deductions as sequences of formulas which can again be coded by numbers. Let S be a set of formulas with the property that there is no deduction of $S' \Rightarrow \bot$ where $S' \subseteq S$.

For each sequence σ of 0's and 1's, we will define a finite set of formulas X_σ: if $n < |\sigma|$ and n codes a formula ϕ then $n \in X_\sigma$ if the n-th position of σ is 1, and the number coding $\neg\phi$ is in X_σ if the n-th position of σ is 0. Now we define a tree T consisting of those σ such that there is no deduction of $S' \Rightarrow \bot$ where $S' \subseteq S \cup X_\sigma$ and the code of the deduction is $\leq |\sigma|$.

Note that we really need the bound on the code of the deduction in order to prove in $\mathbf{RCA_0}$ that the tree exists: the statement that $S \cup X_\sigma$ is consistent—that is, that there is no deduction at all—is Σ_1.

We claim that for every n, there is a $\sigma \in T$ with $|\sigma| = n$: if not, for all the finitely many σ of length n we have a deduction of $\bot$ from some subset of $S \cup X_\sigma$. But given a deduction of $\bot$ from $S_0 \cup \{\phi\}$ and a deduction of $\bot$ from $S_1 \cup \{\neg\phi\}$, we may easily derive a deduction of $\bot$ from $S_0 \cup S_1$. Iterating this argument finitely many times, we obtain a deduction of $\bot$ from some subset of S, contradicting our assumption.

By $\mathbf{WKL_0}$, there is an infinite branch Λ through T. Take $X = \bigcup_{\sigma \sqsubset \Lambda} X_\sigma$. Clearly $S \subseteq X$ and X is consistent: if there were any deduction of a contradiction, it would be coded by some number, and use only a finite subset of X, and therefore at some finite length an initial segment would be rejected from T. For any formula ϕ, ϕ is coded by some n, so when $|\sigma| = n + 1$, we have ensured either ϕ is in X_σ or the negation of ϕ is. Therefore X is a complete, consistent extension of S, as desired.

For the converse, suppose every consistent theory has a complete extension, and let T be a tree of 0's and 1's with infinitely many elements. We define a language with countably many zero-ary relation symbols P_σ, one for each sequence σ. Define a theory S to consist of:

- For each n, the formula $\bigvee_{\sigma,|\sigma|=n} P_\sigma$,
- If $|\sigma| = |\tau|$, $P_\sigma \to \neg P_\tau$,
- If $\sigma \notin T$, $\neg P_\sigma$.

Observe that this theory is consistent: a deduction of a contradiction would use finitely many sentences from S, and in particular, only contains symbols P_σ with $|\sigma| \leq n$ for some n. But the formulas with such propositional variables have a model, namely take any $\sigma \in T$ and make P_τ true iff $\tau \sqsubseteq \sigma$. By assumption, this theory has a complete extension X, and we take $\Lambda = \bigcup_{P_\sigma \in X} \sigma$. □

4.7. **Theories stronger than $\mathbf{ACA_0}$.**

The other two theories of the Big 5 are stronger than $\mathbf{ACA_0}$.

The usual definition of the theory $\mathbf{ATR_0}$ is a bit technical, so we use the following equivalent definition:

Definition 4.23. $\mathbf{ATR_0}$ consists of:

- $\mathbf{RCA_0}$,
- Whenever $\prec_\alpha$ and $\prec_\beta$ are well-orders, either there is a function f mapping the domain of $\prec_\alpha$ onto an initial segment of the domain of $\prec_\beta$, or vice-versa.

$\mathbf{ATR_0}$ stands for "arithmetic transfinite recursion". The standard formulation uses an axiom which states that any arithmetic operation can be iterated along any well-ordering.

We briefly describe the proof-theoretic ordinal of $\mathbf{ATR_0}$. We define a two place function on ordinals, $\varphi(\alpha, \beta)$, as follows:

- $\phi(0, \beta) = \omega^\beta$,
- $\phi(\alpha, \beta)$ is the least ordinal γ such that:
  - If $\beta' < \beta$ then $\phi(\alpha + 1, \beta') < \gamma$,
  - If $\delta < \gamma$ and $\alpha' < \alpha$ then $\phi(\alpha', \delta) < \gamma$.

φ is called the *Veblen fucntion*. (Sadly, there are some slight variations in the precise definition, though this version is fairly typical.) Observe that $\phi(1, \beta) = \epsilon_\beta$, the β-th ordinal closed under the operation $\gamma \mapsto \omega^\gamma$. More generally, $\phi(\alpha, \beta)$ is closed under $\gamma \mapsto \phi(\alpha', \gamma)$ for each $\alpha < \alpha'$.

The ordinal Γ_0 is the least ordinal such that whenever $\alpha, \beta < \Gamma_0$, $\phi(\alpha, \beta) < \Gamma_0$. The ordinal Γ_0 has been associated with the philosophical position of *predicativism*: the idea that the well-defined mathematical notions are those which can be constructed from objects which have been constructed at a previous stage.

The final, strongest, theory in the Big 5 is:

Definition 4.24. $\mathbf{\Pi^1_1\text{-}CA_0}$ consists of:

- P^-,
- The second order induction axiom,
- The comprehension scheme for Π^1_1 formulas.

$\mathbf{\Pi^1_1}$-$\mathbf{CA_0}$ obviously includes $\mathbf{ACA_0}$. The proof-theoretic ordinal of $\mathbf{\Pi^1_1}$-$\mathbf{CA_0}$ is too complicated to describe here, but we briefly explain why. Suppose we wanted to replace the Π^1_1 comprehension axiom the same way we replaced the arithmetic comprehension axiom. The main case is where we introduce some formula on the left-hand side:

$$\Gamma_{\mathbf{ACA_0}}, \Gamma, \forall x(x \in X \leftrightarrow \forall Y \phi(x, Y)) \Rightarrow \Sigma.$$

In $\mathbf{ACA_0}$, we could replace the variable X with $\forall Y \phi(x, Y)$ everywhere it appeared in the deduction. Here, if we can try to do the same, we discover that the variable X could have been used to deduce instances of $\exists Y \neg \phi(x, Y)$! In other words, the set X consisting of those x where $\forall Y \phi(x, Y)$ was true might have used the set X itself to demonstrate that $\forall Y \phi(x, y)$ was false for some particular x!

This is not a minor issue. $\mathbf{\Pi^1_1}$-$\mathbf{CA_0}$ introduces a deep new obstacle to cut-elimination: which elements belong to a set defined by a Π^1_1 formula depends on a quantification over all sets, including the set which is in the process of being defined. This is a genuine circularity, and raises genuine new difficulties, both mathematically and philosophically. (We should consider that it is *not* obvious, when stated like this, that Π^1_1 comprehension is actually well-defined at all: how can we be sure that we won't accidentally write down some set X with the property that $x \in X$ iff $x \notin X$?) This theory is *impredicative*: the meaning of a set defined by a Π^1_1 formula depends essentially on a family of objects—the sets of numbers—which has not yet been completely constructed.

For the sake of comparison with the literature, we mention that the proof theoretic ordinal of $\mathbf{\Pi^1_1}$-$\mathbf{CA_0}$ is given as the limit of a sequence of ordinals, with most (though not all) of the technical work needed to produce the first one (the later ones iterate the same idea relative to the first one). The first of these ordinals is known as the *Howard-Bachmann ordinal*, and is usually written $\psi\epsilon_{\Omega+1}$. Here Ω is (a name for) some very large ordinal, larger than any other in the system—for example, $\aleph_1$—$\epsilon_{\Omega+1}$ is the first fixed point of the $\alpha \mapsto \omega^\alpha$ which is larger than Ω, and ψ is a *collapsing function* which takes a very large ordinal and "collapses" it to a smaller one. (Alternatively, one can view $\epsilon_{\Omega+1}$ as representing the function $\alpha \mapsto \epsilon_{\alpha+1}$ and ψ as being an operation which chooses an ordinal which is both a fixed point of this function and also which is larger than the chosen fixed point of all "easily defined" functions which grow more slowly.) Note that this definition is itself impredicative—we define the ordinal in terms of larger ordinals (or the class of functions on ordinals).