

A Hierarchy of Theories of Arithmetic

Michael Hamm

April 30, 2002, and, in revised form, May 9, 2002

Abstract

This paper summarizes work done in establishing a hierarchy of theories of arithmetic.

The main theorems of this paper will summarize the work done in comparing various theories of arithmetic. This work was started by Parsons in [4] and was continued by Kirby in his Ph.D. thesis [2], which he presented, with J. Paris, at Wrocław that same year [3]. Paris, et al., went on to write [1], which also compares the theories we discuss to others with yet weaker axioms.

Certainly more work was done in this area: Some (e.g., [5]) have examined specific theorems of certain theories that are not theorems of others. We will not go so far, but we will express the theorems and their proofs in, we hope, a more readable version than others have before us, and with seemingly slightly weaker hypotheses in two instances.

We will start with some definitions.

Definition 1 P^- denotes the following set of axioms in the language $\mathcal{L} = \{+, \cdot, \langle, \rangle, ()_1, ()_2, <, 0\}$:

$$\begin{aligned}\forall x \forall y & (x + y = y + x \wedge x \cdot y = y \cdot x) \\ \forall x \forall y \forall z & ((x + y) + z = x + (y + z) \wedge x \cdot (y \cdot z) = (x \cdot y) \cdot z) \\ \forall x & (x + 0 = x \wedge x \cdot 0 = 0) \\ \forall x \forall y \forall z & x \cdot (y + z) = (x \cdot y) + (x \cdot z) \\ \forall x & x + 1 \neq 0 \\ \forall x \forall y \forall z & (x + z = y + z \rightarrow x = y) \\ \forall x \forall y & (x < y \leftrightarrow \exists z (x + z) + 1 = y) \\ \forall x \forall y & (x = y \vee x < y \vee y < x) \\ \forall x & x = \langle (x)_1, (x)_2 \rangle \\ \forall x \forall y & \langle x, y \rangle + \langle x, y \rangle = ((x + y) \cdot ((x + y) + 1)) + (x + x)\end{aligned}$$

Definition 2 $x \leq y$ denotes $x < y \vee x = y$. For any k , $\langle x_1, x_2, \dots \rangle$ denotes $x = \langle \langle \dots \langle x_1, x_2 \rangle, x_3 \rangle, \dots \rangle, x_k$, and we denote each x_i by $(x)_i$.¹

¹This introduces some ambiguity. For example, does $(x)_2$ refer to the second element of $x = \langle (x)_1, (x)_2 \rangle$ or, perhaps, to the second element of $x = \langle (x)_1, (x)_2, (x)_3, (x)_4, (x)_5, (x)_6, (x)_7 \rangle$? It should be clear from the context.

Definition 3 The set of terms is the smallest class of strings in \mathcal{L} containing the variables and 0, and such that, for terms s and t , we have $s + t$, $s \cdot t$, $(s)_1$, $(s)_2$, and $\langle s, t \rangle$ also being terms. An atomic formula is of the form $s = t$ or $s < t$ for terms s, t . Δ_0 is the smallest set of formulas containing atomic formulas and closed under binary \vee , binary \wedge , unary \neg , and bounded quantification using \exists or \forall . $\Sigma_0 = \Pi_0 = \Delta_0$. Σ_{n+1} is the set of formulas of the form $\exists x_1 \exists x_2 \dots \exists x_k \phi$, where $\phi \in \Pi_n$. Π_{n+1} is the set of formulas of the form $\forall x_1 \forall x_2 \dots \forall x_k \phi$, where $\phi \in \Sigma_n$.

Lemma 4 Any Σ_n formula is equivalent to one in which adjacent unbounded quantifiers are of different type (existential never adjoins existential, nor universal universal).

Any Π_n formula is equivalent to one in which adjacent unbounded quantifiers are of different type.

Proof Fixing $n, k_1, k_2, \dots, k_n \in \mathbf{N}$ and $\phi \in \Sigma_0$, we must show that

$$\exists x_n^1 \exists x_n^2 \dots \exists x_n^{k_n} \forall x_{n-1}^1 \forall x_{n-1}^2 \dots \forall x_{n-1}^{k_{n-1}} \dots \mathcal{Q}x_1^1 \mathcal{Q}x_1^2 \dots \mathcal{Q}x_1^{k_1} \phi$$

(for the appropriate quantifier \mathcal{Q}) is equivalent to a formula of the form

$$\exists x_n \forall x_{n-1} \dots \mathcal{Q}x_1 \psi$$

with $\psi \in \Sigma_0$. Let us induce on n .

In the case $n = 1$, we must show

$$\exists x^1 \exists x^2 \dots \exists x^k \phi(x^1, x^2, \dots, x^k)$$

(with $\phi \in \Sigma_0$) is equivalent to a formula of the form $\exists x \psi(x)$ with $\psi \in \Sigma_0$. We can take $\psi(x) = \phi((x)_1, (x)_2, \dots, (x)_k)$.

So our inductive hypothesis is that the lemma holds for Σ_{n-1} . In the case n , by the inductive hypothesis, we may assume we have a formula of the form

$$\exists x^1 \exists x^2 \dots \exists x^k \forall x_2 \exists x_3 \dots \mathcal{Q}x_n \phi(x^1, x^2, \dots, x^k, x_2, x_3, \dots, x_n)$$

with $\phi \in \Sigma_0$. The argument follows that used in the case $n = 1$. Induction completes the proof for Σ_n formulas.

The argument for Π_n formulas is completely analogous. ||

Definition 5 For any set S of formulas, the set IS is the set of formulas of the form

$$\forall \vec{u} ([\theta(0, \vec{u}) \wedge \forall x (\theta(x, \vec{u}) \rightarrow \theta(x+1, \vec{u}))] \rightarrow \forall x \theta(x, \vec{u}))$$

for all $\theta \in S$.

The set LS is of formulas

$$\forall \vec{u} (\exists x \theta(\vec{u}, x) \rightarrow \exists x [\theta(\vec{u}, x) \wedge \forall z < x \neg \theta(\vec{u}, z)])$$

for $\theta \in S$.

The set BS is of formulas

$$\forall \vec{u} \forall y [\forall x < y \exists z \theta(\vec{u}, x, z) \rightarrow \exists t \forall x < y \exists z < t \theta(\vec{u}, x, z)]$$

for $\theta \in S$.

And the set $B'S$ is of formulas

$$\forall \vec{u} [\forall x \exists z \theta(\vec{u}, x, z) \rightarrow \forall y \exists t \forall x < y \exists z < t \theta(\vec{u}, x, z)]$$

for $\theta \in S$.

Definition 6 When a set B of sentences proves a set C of sentences, we will use the notation $B \rightarrow C$. $B \leftrightarrow C$ will be used to mean $B \rightarrow C \wedge C \rightarrow B$; $B \not\rightarrow C$ will be used to mean that B does not prove C . When $A \cup B \rightarrow C$, we will sometimes write in the presence of A , $B \rightarrow C$, or $A \vdash B \rightarrow C$.

Lemma 7 For each n , if $\phi, \theta \in \Sigma_n$ (Π_n), then $\phi \wedge \theta, \phi \vee \theta$ are logically equivalent, in the presence of P^- , to a sentence in Σ_n (Π_n).

Proof This is certainly true when $n = 0$.

Suppose our lemma is true for Σ_{n-1} and Π_{n-1} , and suppose $\phi, \theta \in \Sigma_n$ resp. Π_n). By Lemma 4, we may assume that $\phi = \exists x \psi$ and $\theta = \exists x \eta$ with $\psi, \eta \in \Pi_n$ (resp. $\phi = \forall x \psi$ and $\theta = \forall x \eta$ with $\psi, \eta \in \Sigma_n$).

Then $\phi \wedge \theta$ is $\exists x \psi \wedge \exists y \eta$, which is equivalent to $\exists s(\psi((s)_1) \wedge \eta((s)_2))$. (Respectively, $\phi \wedge \theta$ is $\forall x \psi \wedge \forall y \eta$, equivalent to $\forall s(\psi((s)_1) \wedge \eta((s)_2))$.) This is then in Σ_n (Π_n) by induction.

The argument for \vee is identical in every respect. ||

Lemma 8 If a set of sentences S is closed under binary \wedge and \vee , and contains all sentences of the form $\neg x < y$, then $BS \leftrightarrow B'S$.

Proof Suppose BS holds. Fix \vec{u} and suppose $\forall x \exists z \theta(\vec{u}, x, z)$ for some fixed $\theta \in S$. Fix y . Certainly $\forall x < y \exists z \theta(\vec{u}, x, z)$. By BS ,

$$\exists t \forall x < y \exists z < t \theta(\vec{u}, x, z)$$

which shows $B'S$.

Conversely, suppose $B'S$ and, fixing $\theta \in S, \vec{u}, y$, suppose $\forall x < y \exists z \theta(\vec{u}, x, z)$. Then $\forall x \exists z(x < y \rightarrow \theta(\vec{u}, x, z))$. $B'S$ applies to this, and implies

$$\forall y' \exists t \forall x < y' \exists z < t(x < y \rightarrow \theta(\vec{u}, x, z))$$

and, in particular,

$$\exists t \forall x < y \exists z < t(x < y \rightarrow \theta(\vec{u}, x, z));$$

i.e., $\exists t \forall x < y \exists z < t \theta(\vec{u}, x, z)$, which shows BS . ||

Corollary 9 For each n , $B\Sigma_n \leftrightarrow B'\Sigma_n$ and $B\Pi_n \leftrightarrow B'\Pi_n$.

Proof Immediate. ||

Theorem 10 In the presence of $P^- \cup I\Sigma_0$, we have, for all $n \geq 0$,

$$I\Sigma_{n+1} \rightarrow B\Sigma_{n+1} \leftrightarrow B\Pi_n \rightarrow I\Sigma_n \leftrightarrow I\Pi_n \leftrightarrow L\Sigma_n \leftrightarrow L\Pi_n.$$

We will prove this in several propositions.

Proposition 11 $P^- \vdash I\Sigma_n \leftrightarrow I\Pi_n$.

Proof Suppose $\theta(\vec{u}, x) \in \Sigma_n$ (resp. Π_n) and $I\Sigma_n$ (resp. $I\Pi_n$) fails for it; i.e., there is a model M of

$$\theta(\vec{u}, 0) \wedge \neg\theta(\vec{u}, a) \wedge \forall x(\theta(\vec{u}, x) \rightarrow \theta(\vec{u}, x+1)).$$

Let $\phi(\vec{u}, y) = \neg\theta(\vec{u}, a-y) \in \Pi_n$ (resp. Σ_n). Because $\forall x(\theta(\vec{u}, x) \rightarrow \theta(\vec{u}, x+1))$, we have, for each y ,

$$\neg\theta(\vec{u}, a-y) \rightarrow \neg\theta(\vec{u}, a-y-1)$$

and so $\phi(\vec{u}, y) \rightarrow \phi(\vec{u}, y+1)$. Thus,

$$M \models \phi(\vec{u}, 0) \wedge \neg\phi(\vec{u}, a) \wedge \forall y(\phi(\vec{u}, y) \rightarrow \phi(\vec{u}, y+1)).$$

$I\Pi_n$ (resp. $I\Sigma_n$) thus fails in M . ||

Proposition 12 $P^- \vdash I\Pi_n \leftrightarrow L\Sigma_n$.

Proof First, suppose $I\Pi_n$ holds, and, fixing \vec{u} , assume $\exists x\theta(\vec{u}, x)$ for some $\theta \in \Sigma_n$. Let $\phi(\vec{u}, x) = \forall y < x \neg\theta(\vec{u}, y) \in \Pi_n$.

We claim $\neg\forall x\phi(\vec{u}, x)$. Indeed, by $\exists x\theta(\vec{u}, x)$ we obtain

$$\exists x\exists y < x \ x = y+1 \wedge \theta(\vec{u}, y)$$

and so $\neg\forall x\forall y < x \neg\theta(\vec{u}, y)$, which is our claim.

Thus, we have $\phi(\vec{u}, 0) \wedge \neg\forall x\phi(\vec{u}, x)$, so, by $I\Pi_n$, we obtain

$$\neg\forall x(\phi(\vec{u}, x) \rightarrow \phi(\vec{u}, x+1)),$$

and thus $\exists x(\phi(\vec{u}, x) \wedge \neg\phi(\vec{u}, x+1))$, so that

$$\exists x(\forall y < x \neg\theta(\vec{u}, y)) \wedge \neg\forall y < x+1 \neg\theta(\vec{u}, y).$$

Thus, $\exists x(\theta(\vec{u}, x) \wedge \forall y < x \neg\theta(\vec{u}, y))$, and $L\Sigma_n$ holds.

Conversely, suppose $L\Sigma_n$ holds in some model M , and, fixing \vec{u} , assume $\theta(\vec{u}, 0) \wedge \exists x\neg\theta(\vec{u}, x)$ for some $\theta \in \Pi_n$. By $L\Sigma_n$ we have

$$\exists x(\neg\theta(\vec{u}, x) \wedge \forall z < x\theta(\vec{u}, z)).$$

Find such an x in M . We then have $\theta(\vec{u}, x-1) \wedge \neg\theta(\vec{u}, x)$, i.e., $\neg\forall x(\theta(\vec{u}, x) \rightarrow \theta(\vec{u}, x+1))$, which shows that $I\Pi_n$ holds in M . ||

Lemma 13 *Suppose $\phi(\vec{u}, y) \in \Sigma_n$. Then $\forall x < y\phi(\vec{u}, x)$ is logically equivalent, assuming $B\Sigma_{n-1}$ and $B\Pi_{n-1}$, to a formula in Σ_n .*

Proof This is trivially true for $n=0$. Suppose it is true for Σ_{n-1} , and suppose $\phi \in \Sigma_n$. We may, by Lemma 4, assume $\phi(\vec{u}, y) = \exists z\theta(\vec{u}, y, z)$ for some $\theta \in \Pi_{n-1}$. With $B\Pi_{n-1}$, $\forall x < y\exists z\theta(\vec{u}, x, z)$ is equivalent to

$$\exists t\forall x < y\exists z < t\theta(\vec{u}, x, z)$$

i.e.,

$$\neg\forall t\exists x < y\forall z < t\neg\theta(\vec{u}, x, z)$$

which, by our inductive hypothesis, is equivalent, given $B\Sigma_{n-1}$ and $B\Pi_{n-1}$, to a formula in Σ_n . ||

Remark 14 *Kirby [2] only states Lemma 13 assuming $B\Sigma_n$, which is unnecessary in its proof. Of course, Proposition 16 will show that our hypotheses are equivalent to his.*

Corollary 15 *Suppose $\phi(\vec{u}, y) \in \Pi_n$. Then $\exists x < y\phi(\vec{u}, x)$ is logically equivalent, assuming $B\Sigma_{n-1}$ and $B\Pi_{n-1}$, to a formula in Π_n .*

Proof Indeed, when $y \neq 0$, $\neg\exists x < y\phi(\vec{u}, x)$ is equivalent to $\forall x < y\neg\phi(\vec{u}, x)$, which is, by the lemma, in Σ_n . (If $y = 0$, $\exists x < y\phi(\vec{u}, x)$ is equivalent to $0 \neq 0 \in \Pi_n$.) ||

Proposition 16 $P^- \vdash B\Pi_n \leftrightarrow B\Sigma_{n+1}$.

Proof The “ \leftarrow ” direction is trivial. For the converse, assume $P^- \cup B\Pi_n$, and, for $\theta(\vec{u}, n, z, x, y) \in \Pi_n$ and fixed \vec{u}, y , suppose $\forall x < y\exists z\exists n\theta(\vec{u}, n, z, x, y)$. Then

$$\forall x < y\exists a\theta(\vec{u}, (a)_1, (a)_2, x, y).$$

By $B\Pi_n$,

$$\exists t\forall x < y\exists a < t\theta(\vec{u}, (a)_1, (a)_2, x, y).$$

Thus, $\exists t\forall x < y\exists z < t\exists n\theta(\vec{u}, n, z, x, y)$.² Since y, \vec{u} , and θ were arbitrary, we have, by Lemma 4, proven $B\Sigma_{n+1}$. ||

Remark 17 *Kirby [2] only states Proposition 16 assuming the presence of $P^- \cup I\Sigma_0$, which is unnecessary in its proof. Of course, Proposition 19 will show that our hypotheses are equivalent to his.*

Proposition 18 $P^- \vdash I\Sigma_n \rightarrow B\Sigma_n$ for each $n > 0$.

²Indeed, $\forall a (a)_1 + (a)_2 < a$, as $a + a = (((a)_1 + (a)_2) \cdot ((a)_1 + (a)_2 + 1)) + (a)_1 + (a)_1$.

Proof By Proposition 16, it is sufficient to prove, for each $n \geq 0$, $I\Sigma_{n+1} \rightarrow B\Pi_n$. Suppose, then, that $I\Sigma_{n+1}$ holds, and assume, for $\theta \in \Pi_n$ and fixed y, \vec{u} ,

$$\forall x < y \exists z \theta(\vec{u}, x, z). \quad (1)$$

Let $\psi(\vec{u}, u) = y < u \vee \exists t \forall x < u \exists z < t \theta(\vec{u}, x, z)$.

We claim that $\psi \in \Sigma_{n+1}$. Indeed, when $n = 0$ this is trivial. When $n > 0$, we are assuming $I\Sigma_{n+1}$ and are assuming $I\Sigma_n \rightarrow B\Pi_{n-1}$. Hence, we may (by Proposition 16) assume $B\Sigma_n$. Corollary 15 then applies, and shows us that $\psi \in \Sigma_{n+1}$.

We proceed to apply $I\Sigma_{n+1}$ to ψ . $\psi(\vec{u}, 0)$ is vacuously true. Suppose, for fixed u , $\psi(\vec{u}, 0)$ holds. Then $\psi(\vec{u}, u + 1)$ holds trivially if $u \geq y$; if $u < y$ then $\exists t \forall x < u \exists z < t \theta(\vec{u}, x, z)$. Find such t , and find (by (1)) z such that $\theta(\vec{u}, u, z)$. Let $\tau = t + z + 1$. Then $\forall x < u + 1 \exists z < \tau \theta(\vec{u}, x, z)$. Thus by induction, we have $\forall u \psi(\vec{u}, u)$. In particular, we have $\psi(\vec{u}, y)$. Hence, $\exists t \forall x < y \exists z < t \theta(\vec{u}, x, z)$. Since \vec{u} and θ were arbitrary, we have $B\Pi_n$. \parallel

Proposition 19 $P^- \cup I\Sigma_0 \vdash B\Sigma_{n+1} \rightarrow I\Sigma_n$.

Proof This is trivial when $n = 0$; suppose $B\Sigma_n \rightarrow I\Sigma_{n-1}$, and suppose $B\Sigma_{n+1}$ holds. Fixing $\theta \in \Pi_{n+1}, \vec{u}, a$, say

$$\exists w \theta(\vec{u}, w, 0) \wedge \neg \exists w \theta(\vec{u}, w, a).$$

By Lemma 4 it is sufficient to show

$$\exists x (\exists w \theta(\vec{u}, w, x) \wedge \neg \exists w \theta(\vec{u}, w, x + 1)).$$

Clearly we have

$$\forall x < a \exists w (\theta(\vec{u}, w, x) \vee (w = 0 \wedge \neg \exists v \theta(\vec{u}, v, x))).$$

$B\Sigma_{n+1} \rightarrow B\Pi_n$, which implies

$$\exists t \forall x < a \exists w < t (\theta(\vec{u}, w, x) \vee (w = 0 \wedge \neg \exists v \theta(\vec{u}, v, x)));$$

i.e., $\forall x < a (\exists w \theta(\vec{u}, w, x) \leftrightarrow \exists w < t \theta(\vec{u}, w, x))$. Hence, $\exists w < t \theta(\vec{u}, w, x) \wedge x < a$ which is, by Corollary 15, in Π_{n-1} . We are assuming $B\Sigma_{n+1}$; hence, $B\Sigma_n, I\Sigma_{n-1}$ by induction, and, by Proposition 11, we have $L\Pi_{n-1}$. Thus, we this is true when $x = 0$ but false for $x = a$, so that

$$\exists x ([\exists w < t \theta(\vec{u}, w, x) \wedge x < a] \wedge \neg [\exists w < t \theta(\vec{u}, w, x + 1) \wedge x + 1 < a]).$$

Thus, $\exists x (\exists w \theta(\vec{u}, w, x) \wedge \neg \exists w \theta(\vec{u}, w, x + 1))$, which is what was to have been shown. \parallel

Proposition 20 $P^- \vdash I\Sigma_n \leftrightarrow L\Pi_n$.

Proof First, suppose $L\Pi_n$ holds, and suppose, for $\theta \in \Sigma_n$ and fixed \vec{u} , we have $\theta(\vec{u}, 0) \wedge \exists x \neg \theta(\vec{u}, x)$. We must show $\neg \forall x (\theta(\vec{u}, x) \rightarrow \theta(\vec{u}, x + 1))$. By $L\Pi_n$ and $\exists x \neg \theta(\vec{u}, x)$, we have

$$\exists x (\neg \theta(\vec{u}, x) \wedge \forall y < x \theta(\vec{u}, y)),$$

which implies $\exists x (\neg \theta(\vec{u}, x) \wedge \theta(\vec{u}, x - 1))$, as was to have been shown.

Conversely, suppose $I\Sigma_n$ holds and suppose, for $\theta \in \Pi_n$ and fixed \vec{u} , $\exists y \theta(\vec{u}, y)$. Let $\phi(\vec{u}, x) = \forall y < x \neg \theta(\vec{u}, y)$. Note that $\phi \in \Sigma_n$ by Lemma 13, because we have $B\Sigma_n$ by Proposition 18. Thus, $I\Sigma_n$ applies to ϕ .

$\phi(\vec{u}, 0)$ is true vacuously. $\neg \forall x \phi(\vec{u}, x)$ is true.³ Thus,

$$\neg \forall x (\phi(\vec{u}, x) \rightarrow \phi(\vec{u}, x + 1));$$

i.e., $\exists x (\phi(\vec{u}, x) \wedge \neg \phi(\vec{u}, x + 1))$, which is the same as

$$\exists x ((\forall y < x \neg \theta(\vec{u}, y)) \wedge \exists y < x + 1 \theta(\vec{u}, y));$$

i.e., $\exists x (\theta(\vec{u}, x) \wedge \forall y < x \neg \theta(\vec{u}, y))$, which proves $L\Pi_n$. ||

Our proof of Theorem 10 is now complete.

Definition 21 Peano Arithmetic is $P^- \cup (\bigcup_n I\Sigma_n)$, and is denoted PA.

Theorem 22 The converses to the unidirectional arrows in the statement of Theorem 10 are false. That is, even in the presence of $P^- \cup I\Sigma_0$ we have (for each n) $B\Sigma_{n+1} \quad I\Sigma_{n+1}$ and $I\Sigma_n \quad B\Pi_{n+1}$.

The proof of this theorem is beyond the scope of this paper.

Corollary 23 There is no finite list of axioms S such that $P^- \vdash S \leftrightarrow \text{PA}$.

Proof Suppose $S \leftrightarrow \text{PA}$ in the presence of P^- , where S is some finite set of axioms. Assume the presence of P^- . $\bigcup_n I\Sigma_n \rightarrow S$, so that for large enough N , $\bigcup_{n=0}^N I\Sigma_n \rightarrow S$.⁴ But $I\Sigma_{N+1} \rightarrow \bigcup_{n=0}^N I\Sigma_n$, so $I\Sigma_{N+1} \rightarrow S \rightarrow \text{PA} \rightarrow \bigcup_n I\Sigma_n \rightarrow I\Sigma_{N+2}$. But this contradicts Theorem 22. ||

References

- [1] R. Kaye, J. Paris, C. Dimitracopoulos, “On Parameter-Free Induction Schemas”, *J. Symb. Logic* **53** (1988), pp. 1082–1097.
- [2] L. A. S. Kirby, *Initial Segments of Models of Arithmetic*. Ph.D. Thesis, Manchester, 1977.

³Indeed, we have $\exists y \theta(\vec{u}, y)$, so $\exists x \exists y < x \theta(\vec{u}, y)$ (just take, for example, $x = y + 1$). $\neg \forall x \phi(\vec{u}, x)$ follows immediately.

⁴Any axiomatization A of a finitely axiomatizable set T has a finite subset $F \subset A$ which axiomatizes T . This follows from a result called Compactness, which is beyond the scope of this paper

- [3] J. B. Paris, L. A. S. Kirby, “ Σ_n -Collection Schemas in Arithmetic”, *Logic Colloquium '77*, North-Holland 1978.
- [4] Ch. Parsons, “On a Number Theoretic Choice Schema and Its Relation to Induction”, *Intuitionism and Proof Theory*, North-Holland 1970.
- [5] A. J. Wilkie, “Some Results and Problems on Weak Systems of Arithmetic”, *Logic Colloquium '77*, North-Holland 1978.