UC Irvine
UC Irvine Previously Published Works

Title
Comparing Peano arithmetic, Basic Law V, and Hume's Principle

Permalink
https://escholarship.org/uc/item/0907s3cj

Journal
Annals of Pure and Applied Logic, 163(11)

ISSN
0168-0072

Author
Walsh, Sean

Publication Date
2012-11-01

DOI
10.1016/j.apal.2011.12.016

Peer reviewed

Comparing Peano Arithmetic, Basic Law V, and Hume's Principle

Sean Walsh

Department of Philosophy, Birkbeck, University of London
Mailing Address: Department of Philosophy, Birkbeck College, Malet Street, London WC1E 7HX, UK
Email: swalsh108@gmail.com or s.walsh@bbk.ac.uk
Website: http://www.swalsh108.org

Abstract

This paper presents new constructions of models of Hume's Principle and Basic Law V with restricted amounts of comprehension. The techniques used in these constructions are drawn from hyperarithmetic theory and the model theory of fields, and formalizing these techniques within various subsystems of second-order Peano arithmetic allows one to put upper and lower bounds on the interpretability strength of these theories and hence to compare these theories to the canonical subsystems of second-order arithmetic. The main results of this paper are: (i) there is a consistent extension of the hypearithmetic fragment of Basic Law V which interprets the hypearithmetic fragment of second-order Peano arithmetic (cf. Corollary 54 and Figure 2), and (ii) the hyperarithmetic fragment of Hume's Principle does not interpret the hyperarithmetic fragment of second-order Peano arithmetic (cf. Corollary 92 and Figure 2), so that in this specific sense there is no predicative version of Frege's Theorem.

Keywords: Second-order arithmetic, Basic Law V, Hume's Principle, hyperarithmetic, recursively saturated, interpretability
2000 MSC: 03F35, 03D65, 03C60, 03F25

Contents

1. Introduction, Definitions, and Overview of Main Results

1.1. Introduction

Second-order Peano arithmetic and its subsystems have been studied for many decades by mathematical logicians (cf. [35]), and the resulting theory continues to be the subject of current research and a source of open problems. More recently, philosophers of mathematics have begun to study systems closely related to second-order Peano arithmetic (cf. [8]). One of these systems, namely, Hume's Principle, constitutes an axiomatization of cardinality which is similar to the notion of cardinality defined in Zermelo-Frankel set theory. The contemporary philosophical interest in this principle stems from Crispin Wright's suggestion that it can serve as the centerpiece of a revitalized version of Frege's logicism (cf. [43], [44], [25]). Frege himself focused his logicism around a principle called Basic Law V, which in effect codified an alternative conception of set. While Russell's paradox shows that Basic Law V is inconsistent with the unrestricted comprehension schema (cf. Proposition 4), this principle has garnered renewed attention due to Ferreira and Wehmeier's recent proof that it is consistent with the hyperarithmetic comprehension schema ([13], cf. [41, 42] and Remark 52).

The goal of this paper is to apply methods from the subsystems of second-order Peano arithmetic to the subsystems of Basic Law V and Hume's Principle. In particular, we use methods from hyperarithmetic theory to build models of subsystems of Basic Law V (§ 3), and we use recursively saturated models and ideas from the model theory of fields to build models of subsystems of Hume's Principle and Basic Law V (§ 4). Our primary application of these new constructions is to compare the interpretability strength of the subsystems of second-order Peano arithmetic to the subsystems of Basic Law V and Hume's Principle. For, one of the few known ways to show that one theory is of *strictly* greater interpretability strength than another theory is to show that the first proves the consistency of the second (cf. Proposition 7). Hence, by formalizing our constructions, we can compare the interpretability strength of subsystems of Hume's Principle and Basic Law V to subsystems of Peano arithmetic. Our main results about interpretability are summarized in § 1.5 and on Figure 2. Prior to summarizing these results, we first present formal definitions of the theories and subsystems of Hume's Principle and Basic Law V (§§ 1.2-1.4) and then describe what is and is not known about the provability relation among these subsystems (§ 1.4 and Figure 1).

1.2. Definition of the Signatures and Theories of $\mathtt{PA}^2$, $\mathtt{BL}^2$ and $\mathtt{HP}^2$

The signature of $\mathtt{PA}^2$ is a many-sorted signature, with sorts for numbers as well as a sort for sets of numbers. The theory $\mathtt{PA}^2$ is a natural set of axioms for the following many-sorted structure in this signature:

$$(\omega, 0, s, +, \times, \leq, P(\omega)) \tag{1}$$

This structure satisfies the eight-axioms of Robinson's $\mathtt{Q}$

(Q1) $sx \neq 0$ (Q2) $sx = sy \to x = y$ (Q3) $x \neq 0 \to \exists\, w\; x = sw$
(Q4) $x + 0 = x$ (Q5) $x + sy = s(x + y)$ (Q6) $x \cdot 0 = 0$
(Q7) $x \cdot sy = x \cdot y + x$ (Q8) $x \leq y \leftrightarrow \exists\, z\; x + z = y$

and the mathematical induction axiom

$$\forall\, F\ [F(0)\ \&\ \forall\, n\ F(n) \to F(s(n))] \to [\forall\, n\ F(n)] \tag{2}$$

and each instance of the comprehension schema (where F does not occur free in φ)

$$\exists\, F\ \forall\, n\ [F(n) \leftrightarrow \varphi(n)] \tag{3}$$

Here, the formula φ is allowed to contain free object variables (in addition to n) and free set variables (with the exception of F). Hence, what an instance of this comprehension schema says is that if $\varphi(n)$ is a formula with parameters, then there is a set F corresponding to it. This all in place, we are now in a position to define:

Definition 1. The theory $\mathtt{PA}^2$ or $\mathtt{CA}^2$ or *second-order Peano arithmetic* consists of Q1-Q8, the mathematical induction axiom (2), and each instance of the comprehension schema (3) (cf. [35] p. 4).

The name $\mathtt{CA}^2$ is also given to $\mathtt{PA}^2$ because it reminds us of comprehension.

The signature of $\mathtt{HP}^2$ and $\mathtt{BL}^2$ is likewise a many-sorted signature, with sorts for objects as well as sorts for n-ary relations on objects, and with an additional function symbol from the unary relation sort to the object sort. The unary relations are written as $A, B, C, F, G, H, X, Y, Z$ and will be called *sets*, and the n-ary relation symbols for $n > 1$ are written as f, g, h, P, Q, R, S and will be called *relations*. Occasionally when we want to say something about both sets and relations, we will talk about all n-ary relations for $n \geq 1$. The additional function symbol is denoted by $\#$ in the case of $\mathtt{HP}^2$ and by ∂ in the case of $\mathtt{BL}^2$. So the signatures of $\mathtt{HP}^2$ and $\mathtt{BL}^2$ are exactly the same: it is merely for the sake of convenience and clarity that we use $\#$ in the context of $\mathtt{HP}^2$ and ∂ in the context of $\mathtt{BL}^2$. Hence, structures in this signature have the form

$$(M, S_1, S_2, \ldots, \#) \tag{4}$$

where M is a set, $S_n \subseteq P(M^n)$ and $\# : S_1 \to M$. Note that the function $\#$ only goes from S_1 to M, so that the relations from S_n for $n > 1$ are not in the domain of this function.

It is worth pausing for a moment to dwell on a technical point. Formally, the signature of $\mathtt{PA}^2$ also contains a binary relation symbol E which holds between an object and a set and which, in the standard model from (1), is interpreted by the $\in$ relation from the ambient set-theory. In structures where this holds, let us say that the symbol E is interpreted *absolutely*. It is easy to see that every structure in the signature of $\mathtt{PA}^2$ is isomorphic to a structure that interprets this symbol absolutely, and it is for this reason that this symbol is typically suppressed when describing structures. Likewise, formally the signature of $\mathtt{HP}^2$ and $\mathtt{BL}^2$ contains $(n+1)$-ary relation symbols E_n, which hold between n-tuples of objects and n-ary relations. Further, there is an obvious generalization of the notion of absoluteness for structures in this signature, such that the structure from (4) interprets E_n absolutely, and such that every structure in this signature is isomorphic to a structure which interprets E_n absolutely. Hence, as in the case of second-order Peano arithmetic, in what follows, these symbols will be suppressed when describing structures, and it will be assumed that every structure in this signature has the form of (4).

Hume's Principle and Basic Law V can now be defined. *Hume's Principle* is the following axiom in the signature of structure (4):

$$\#X = \#Y \iff \exists \text{ bijection } f : X \to Y \tag{5}$$

Here, the notion of bijectivity is defined in terms of functionality, injectivity, and surjectivity in the obvious manner. The axiom *Basic Law V* is the following sentence in this signature:

$$\partial X = \partial Y \iff X = Y \tag{6}$$

Here, two sets are said to be equal if they are coextensive; formally, the equality of coextensive sets can be taken to be an axiom of all the theories considered in this paper. The important thing to note here is that $(M, S_1, S_2, \ldots, \partial)$ is a model of Basic Law V if and only if the function $\partial : S_1 \to M$ is an injection. That is, Basic Law V mandates that a very simple relation holds between S_1 and M. There is no analogue of this in the case of Hume's Principle, since the right-hand side of (5) contains a higher-order quantifier.

Nevertheless, there are many natural models of Hume's Principle, and examining these models is the easiest way to define the theories $\mathtt{HP}^2$ and $\mathtt{BL}^2$. In particular, if α is an ordinal which is not a cardinal, and if $\#$ is interpreted as cardinality, then the following structure is a model of Hume's Principle:

$$(\alpha, P(\alpha), P(\alpha^2), \ldots, \#) \tag{7}$$

Restricting attention to ordinals α that are not cardinals serves the purpose of ensuring that $\#(\alpha) < \alpha$, so that $\mathrm{dom}(\#_\alpha)$ is $P(\alpha)$ and so that $\mathrm{rng}(\#_\alpha)$ is a subset of α. For all n-ary relation variables R and all $n \geq 1$, this structure also satisfies each instance of the following comprehension schema (where R does not occur free in $\varphi(\overline{z})$)

$$\exists R \, \forall \overline{n} \, [\overline{n} \in R \leftrightarrow \varphi(\overline{n})] \tag{8}$$

This comprehension schema is simply the generalization of the comprehension schema from $\mathtt{PA}^2$, namely (3), to the n-ary relations for all $n \geq 1$. Here, as with (3), the formula φ is allowed to include free object variables (in addition to $\overline{n}$) and free relation variables of any arity $m \geq 1$ (with the exception of R). Hence, we can now define the following theories:

Definition 2. The theory $\mathtt{HP}^2$ is the theory that is given by Hume's Principle (5) and the comprehension schema (8).

Definition 3. The theory $\mathtt{BL}^2$ is the theory which is given by Basic Law V (6) and the comprehension schema (8).

The primary focus of this paper is on subsystems of $\mathtt{HP}^2$ and $\mathtt{BL}^2$ that are generated by restrictions on the complexity of the formulas appearing in the comprehension schema (8). This is due to the fact that we seek to compare the interpretability strength of these subsystems to those of second-order Peano arithmetic. However, unlike in the case of $\mathtt{PA}^2$ and $\mathtt{HP}^2$, attention *must* be restricted to these subsystems in the case of $\mathtt{BL}^2$. For, it is not difficult to see that Russell's paradox shows that $\mathtt{BL}^2$ is inconsistent:

Proposition 4. $\mathtt{BL}^2$ is inconsistent.

Proof. By applying the comprehension schema (8) to the formula

$$\varphi(x) \equiv \exists\, Y\; \partial(Y) = x\; \&\; x \notin Y \tag{9}$$

it follows that $\mathtt{BL}^2$ proves that there is set X that satisfies

$$\forall\, x\; [x \in X \iff (\exists\, Y\; \partial(Y) = x\; \&\; x \notin Y)] \tag{10}$$

There are then two cases: either $\partial(X) \in X$ or $\partial(X) \notin X$. Case one: suppose that $\partial(X) \in X$. Then by the left-to-right direction of equation (10), it follows that there is Y such that $\partial(Y) = \partial(X)$ and $\partial(X) \notin Y$. But $\partial(Y) = \partial(X)$ and Basic Law V imply that $Y = X$, so that $\partial(X) \notin X$, which contradicts our case assumption. Case two: suppose that $\partial(X) \notin X$. Then by the right-to-left direction of equation (10), it follows that for any Y we have that $\partial(Y) = \partial(X)$ implies $\partial(X) \notin Y$. But then $\partial(X) = \partial(X)$ implies $\partial(X) \notin X$, which contradicts our case assumption. $\square$

Hence $\mathtt{BL}^2$ is inconsistent and does not have any models, unlike the theories $\mathtt{PA}^2$ and $\mathtt{HP}^2$, which respectively have the canonical models (1) and (7).

1.3. Definition of the Subsystems of $\mathtt{PA}^2$, $\mathtt{BL}^2$ and $\mathtt{HP}^2$

So if one wants to study Basic Law V, one needs to pass to subsystems of Basic Law V that do not allow instances of the comprehension schema (8) applied to formulas like the one in (9). To this end, let us introduce the following natural hierarchy of formulas in the signature of $\mathtt{BL}^2$ and $\mathtt{HP}^2$. A formula φ, perhaps with free object variables $\overline{z}$ and free relation variables $\overline{R}$ of different arities $m \geq 1$, is called *arithmetical or* Π_0^1 *or* Σ_0^1 if it does not contain any bound m-ary relation variables for any $m \geq 1$. Further, if $m \geq 1$ and R is an m-ary relation variable and $\varphi(R)$ is a Σ_n^1-formula, then $\exists\, R\; \varphi(R)$ is a Σ_n^1-formula and $\forall\, R\; \varphi(R)$ is a Π_{n+1}^1-formula. Likewise, if $m \geq 1$ and R is an m-ary relation variable and $\varphi(R)$ is Π_n^1-formula, then $\exists\, R\; \varphi(R)$ is a Σ_{n+1}^1-formula and $\forall\, R\; \varphi(R)$ is a Π_n^1-formula.

That is, in this hierarchy of formulas, one is allowed to accumulate arbitrarily many existential relation quantifiers of different arities $m \geq 1$ in front of a Σ_n^1-formula and still remain Σ_n^1, and likewise one is allowed to accumulate arbitrarily many universal relation quantifiers of different arities $m \geq 1$ in front of a Π_n^1-formula and still remain Π_n^1. It is only the change from a universal relation quantifier of some arity $m \geq 1$ to an existential relation quantifier of some arity $m \geq 1$ (or vice-versa) which increases the complexity of the sentence in this hierarchy. For instance, if X is set variable and R and S are binary relation variables, then the following formulas are respectively $\Sigma_1^1, \Pi_1^1, \Sigma_2^1, \Pi_2^1$:

$$\exists\, X\; \forall\, x\; R(x, \#X) \tag{11}$$
$$\forall\, R\; \forall\, X\; \exists\, y\; [\forall\, x\; R(x, y) \rightarrow y = \partial X] \tag{12}$$
$$\exists\, X\; \forall\, R\; [\exists\, x\; R(x, x) \rightarrow R(\#X, \#X)] \tag{13}$$
$$\forall\, R\; \exists\, X\; \exists\, S\; \forall\, y\; [(\forall\, x\; x \in X \leftrightarrow \neg Sxy) \rightarrow R(\partial X, y)] \tag{14}$$

Finally, it is worth explicitly noting that not all formulas are included in our hierarchy of formulas. For instance, we have said nothing about the complexity of formulas which include alternations of object quantifiers and set quantifiers, such as the following formula:

$$\forall\, X\; \exists\, y\; \forall\, Z\; [R(\#X, \#Z) \rightarrow R(y, \#Z)] \tag{15}$$

However, this is not a serious omission, since so long as one includes enough of the comprehension schema (8) to guarantee the existence of the singleton set $\{n\}$ for each element n, the above formula is equivalent to the following Π_3^1-formula

$$\forall\, X\, \exists\, Y\, \forall\, Z\, [\exists\, y \in Y\, \&\, \forall\, z \in Y\, z = y]\, \&\, [R(\#X, \#Z) \to R(y, \#Z)] \qquad (16)$$

That is, we can correct for this omission by treating object quantifiers as set quantifiers over singleton sets when they occur in alternation of object quantifiers and set quantifiers.

Using this hierarchy of formulas, one can define the subsystems of $\mathtt{BL}^2$ and $\mathtt{HP}^2$ by restricting the complexity of formulas which appear in the comprehension schema (8). For the following definition, let us recall that $\mathtt{CA}^2$ is another name for $\mathtt{PA}^2$ (cf. Definition 1). The idea behind the following definition is then that $\mathtt{AC}$ reminds us of the axiom of choice and is the result of inverting the letters in $\mathtt{CA}$, which reminds us of comprehension. So with the exception of the choice schema, each of the schemas which figure in the below definition asserts the existence of a certain class of definable sets and relations:

Definition 5. Suppose that $\mathtt{XY}^2$ is one of $\mathtt{CA}^2$, $\mathtt{BL}^2$, or $\mathtt{HP}^2$. Then we can define the following four subsystems of $\mathtt{XY}^2$:
(i) The subsystem $\mathtt{AXY}_0$ is $\mathtt{XY}^2$ but with the comprehension scheme (8) restricted to arithmetical formulas.
(ii) The subsystem $\Delta_1^1 - \mathtt{XY}_0$ is $\mathtt{XY}^2$ but with the comprehension scheme (8) replaced by the following schema, which is called the Δ_1^1-*comprehension schema* or *hyperarithmetic comprehension schema*, wherein φ is a Σ_1^1-formula and ψ is a Π_1^1-formula:

$$[\forall\, \overline{n}\, \varphi(\overline{n}) \leftrightarrow \psi(\overline{n})]\, \to\, [\exists\, R\, \forall\, \overline{n}\, \overline{n} \in R \leftrightarrow \varphi(\overline{n})] \qquad (17)$$

(iii) The subsystem $\Sigma_1^1 - \mathtt{YX}_0$ is $\mathtt{AXY}_0$ and the following schema, which is called the Σ_1^1-*choice schema*, wherein φ is a Σ_1^1-formula:

$$[\forall\, \overline{n}\, \exists\, P\, \varphi(\overline{n}, P)] \to [\exists\, R\, \forall\, \overline{n}\, \forall\, P\, (\forall\, \overline{m}\, (\overline{m} \in P \leftrightarrow \overline{nm} \in R)) \to \varphi(\overline{n}, P)] \qquad (18)$$

(iv) The subsystem $\Pi_n^1 - \mathtt{XY}_0$ is $\mathtt{XY}^2$ but with the comprehension schema (8) restricted to Π_n^1-formulas.

Further, in all these schemata, φ and ψ are allowed to contain free object variables (in addition to $\overline{n}$) and free relation variables of any arity $m \geq 1$ (with the exception of R).

The intuition behind the choice schema (18) can be made clearer as follows. Suppose that a structure $(M, S_1, S_2, \ldots, \#)$ is a model of $\Sigma_1^1 - \mathtt{PH}_0$ and that the antecedent of a given instance of the Σ_1^1-choice schema (18) holds. Then $\Sigma_1^1 - \mathtt{PH}_0$ asserts the existence of a relation R, which for the sake of simplicity we can assume to be a binary relation. For each object n in M, the following set is then guaranteed to exist in S_1 by the arithmetic comprehension schema (which is included in $\Sigma_1^1 - \mathtt{PH}_0$):

$$R_n = \{m : Rnm\} \qquad (19)$$

So it follows that $(M, S_1, S_2, \ldots, \#) \models \varphi(n, R_n)$ for every n in M. Hence, in the situation where for every n there is a choice of P such that $\varphi(n, P)$, the Σ_1^1-choice schema asserts
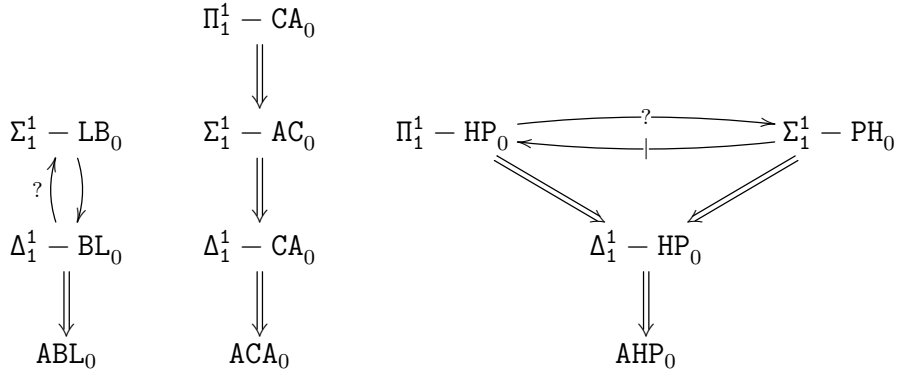
Figure 1: Provability Relation in Subsystems of $\mathtt{BL}^2$, $\mathtt{PA}^2$, and $\mathtt{HP}^2$

that there is a uniform way to make these choices, in that there is an R such that its columns R_n satisfy $\varphi(n, R_n)$ for each n.

Note, however, that the map $(R, n) \mapsto \#(R_n)$ is *not* a function symbol in the signature of $\mathtt{HP}^2$ or $\mathtt{BL}^2$. For instance, given a binary relation R, the comprehension schema (8) restricted to arithmetical formulas does *not* in general guarantee the existence of the binary relation

$$
\begin{aligned}
\{(n, m) : \#(R_n) = m\} &= \{(n, m) : \exists\, X\ (\forall\, x\ x \in X \leftrightarrow Rnx)\ \&\ \#X = m\} \\
&= \{(n, m) : \forall\, X\ (\forall\, x\ x \in X \leftrightarrow Rnx)\ \rightarrow\ \#X = m\} \qquad (20)
\end{aligned}
$$

For, as these definitions make evident, one will in general need the hyperarithmetic comprehension schema (17) in order to show that this relation exists (cf. Propositions 48-49). This example underscores an important fact: intuitively simple relations expressible via the maps $\#$ or ∂ may be quite complex when explicitly written out in terms of the primitives of the signature. Since our interest in this paper is on restrictions of the comprehension schema, this fact will be particularly important to keep in mind throughout this paper. (In § 5, we raise the question of what happens when one *does* include function symbols $(R, n) \mapsto \#(R_n)$ in the signature, so that relations like the one defined in equation (20) would count as arithmetical.)

1.4. Summary of Results about the Provability Relation

Our primary concern in this paper is with the interpretability relation between subsystems of $\mathtt{PA}^2$, $\mathtt{HP}^2$, and $\mathtt{BL}^2$, and we summarize our results in the next section (§ 1.5). However, since provability implies interpretability, and since the provability relation is intrinsically interesting, in this section we record what is known about this relation among the subsystems of $\mathtt{PA}^2$, $\mathtt{HP}^2$, and $\mathtt{BL}^2$. This is summarized in Figure 1, where the double arrows indicate that the provability implication is irreversible, and where the negated arrows indicate that the provability implication fails, and where the arrows with question marks beside them indicate that the provability implication is unknown.

Each of the positive provability relations in in Figure 1 follows immediately from the definitions, except for the fact that $\Pi^1_1 - \mathtt{CA}_0$ proves $\Sigma^1_1 - \mathtt{AC}_0$ and the fact that Σ^1_1-choice

implies Δ_1^1-comprehension. For the former, see Simpson [35] Theorem V.8.3 pp. 205-206. For the latter, the proof from Simpson [35] Theorem VII.6.6 (i) p. 295 carries over to the setting of $\mathtt{HP}^2$ and $\mathtt{BL}^2$, as we verify now:

Proposition 6. $\Sigma_1^1 - \mathtt{AC}_0 \to \Delta_1^1 - \mathtt{CA}_0$, and $\Sigma_1^1 - \mathtt{PH}_0 \to \Delta_1^1 - \mathtt{HP}_0$, and $\Sigma_1^1 - \mathtt{LB}_0 \to \Delta_1^1 - \mathtt{BL}_0$

Proof. Let $\mathcal{M} = (M, S, \dots)$ be a model of $\Sigma_1^1 - \mathtt{AC}_0$ (resp. $\Sigma_1^1 - \mathtt{PH}_0$, $\Sigma_1^1 - \mathtt{LB}_0$). By standard conventions, $\mathcal{M}$ is non-empty. However, nothing in these standard conventions requires that M be non-empty as opposed to say S. But, in the case of $\Sigma_1^1 - \mathtt{AC}_0$ we have that $0 \in M$, and in the case of $\Sigma_1^1 - \mathtt{PH}_0$ we have that $\#\emptyset \in M$, and likewise in the case of $\Sigma_1^1 - \mathtt{LB}_0$ we have that $\partial\emptyset \in M$. Hence, for the remainder of the proof, fix parameter $a \in M$. Suppose that $\mathcal{M} \models \forall \, \overline{z} \, \varphi(\overline{z}) \leftrightarrow \psi(\overline{z})$, where φ is Σ_1^1 and ψ is Π_1^1. Then $\mathcal{M} \models \forall \, \overline{z} \, \varphi(\overline{z}) \vee \neg\psi(\overline{z})$. Then by the arithmetical comprehension schema, $\mathcal{M} \models \forall \, \overline{z} \, \exists \, Z \, (\varphi(\overline{z}) \, \wedge \, a \in Z) \vee (\neg\psi(\overline{z}) \wedge a \notin Z)$. By the Σ_1^1-Choice Schema, there is R such that

$$\mathcal{M} \models \forall \, \overline{z} \, \forall \, Z \, (\forall x \, x \in Z \leftrightarrow R\overline{z}x) \to [(\varphi(\overline{z}) \, \wedge \, a \in Z) \vee (\neg\psi(\overline{z}) \wedge a \notin Z)] \qquad (21)$$

By the arithmetical comprehension schema, there is W such that $\overline{z} \in W$ if and only if $R\overline{z}a$. Then we claim that $\overline{z} \in W$ if and only if $\varphi(\overline{z})$. For, suppose that $\overline{z} \in W$, so that $R\overline{z}a$. Then $Z = \{x : R\overline{z}x\}$ exists by the arithmetical comprehension schema, and we have $a \in Z$. Then by (21), it follows that $\varphi(\overline{z})$. Conversely, suppose that $\overline{z} \notin W$, so that $\neg R\overline{z}a$. Then $Z = \{x : R\overline{z}x\}$ exists by the arithmetical comprehension schema, and we have $a \notin Z$. Then by (21), it follows that $\neg\psi(\overline{z})$ and hence $\neg\varphi(\overline{z})$. Hence, in fact we have established that $\overline{z} \in W$ if and only if $\varphi(\overline{z})$. So $\mathcal{M}$ models $\Delta_1^1 - \mathtt{CA}_0$ (resp. $\Delta_1^1 - \mathtt{HP}_0$, $\Delta_1^1 - \mathtt{BL}_0$). $\square$

The known non-provability relations in Figure 1 are not difficult to verify. In the case of the subsystems of $\mathtt{HP}^2$, we can read these results off of the results for the subsystems of $\mathtt{PA}^2$, as the proof of Proposition 46 indicates. In the case of the subsystems of $\mathtt{BL}^2$, the only known result we have is that $\mathtt{ABL}_0$ does not prove $\Delta_1^1 - \mathtt{BL}_0$, and this is shown in Proposition 44. In § 5, we list the remaining unknown questions about the provability relation, namely, the question of whether $\Delta_1^1 - \mathtt{BL}_0$ implies $\Sigma_1^1 - \mathtt{LB}_0$ and whether $\Pi_1^1 - \mathtt{HP}_0$ implies $\Sigma_1^1 - \mathtt{PH}_0$.

1.5. Summary of Results about the Interpretability Relation

Most of the formal work done on the the subsystems of $\mathtt{PA}^2$, $\mathtt{HP}^2$, $\mathtt{BL}^2$ has concerned the interpretability strength of these theories. A theory T_0 *is interpretable in* a theory T_1 ($T_0 \leq_\mathrm{I} T_1$) if every model M_1 of T_1 uniformly defines without parameters some model M_0 of T_0, where "uniform" has the sense that e.g. a binary relation symbol R in the signature of T_0 is defined by one and the same formula $\varphi(x, y)$ in each model M_1 of T_1. (For a more syntactic definition, see Lindström [23] p. 96 or Hájek and Pudlák [17] pp. 148-149). Since this relation is reflexive and transitive, one can define the associated notions

$$T_0 \equiv_\mathrm{I} T_1 \iff T_0 \leq_\mathrm{I} T_1 \, \& \, T_1 \leq_\mathrm{I} T_0 \qquad (22)$$

$$T_0 <_\mathrm{I} T_1 \iff T_0 \leq_\mathrm{I} T_1 \, \& \, T_1 \nleq_\mathrm{I} T_0 \qquad (23)$$

The relation $\leq_\mathrm{I}$ is then a partial order on the set of equivalence classes of theories under the equivalence relation $\equiv_\mathrm{I}$. Since this partial order is in fact a linear order in many

natural cases, it can be intuitively conceived as a measure of the strength of the theory. This order is also connected to the formal notion of consistency strength by the following proposition:

Proposition 7. Suppose T_1 is a finitely axiomatizable theory such that $\mathtt{ACA}_0 \subseteq T_1 \subseteq \mathtt{PA}^2$, and suppose that T_0 is a computable theory in a computable signature. Then

$$T_1 \vdash \mathrm{Con}(T_0) \implies T_1 \not\leq_{\mathrm{I}} T_0 \tag{24}$$

$$[T_0 \leq_{\mathrm{I}} T_1 \ \& \ T_1 \vdash \mathrm{Con}(T_0)] \implies T_0 <_{\mathrm{I}} T_1 \tag{25}$$

Proof. (Sketch) For (24), note that if $T_1 \vdash \mathrm{Con}(T_0)$, then T_1 proves that there is a model M_0 of T_0 (cf. Simpson [35] Theorem IV.3.3 p. 140). But if $T_1 \leq_{\mathrm{I}} T_0$ and T_1 is finitely axiomatizable, then this interpretation is due to a finite number of the axioms of T_0. Further, since T_0 is computable, this can be accurately represented in T_1, so that inside T_1 the model M_0 of T_0 defines a model M_1 of T_1, which likewise exists since the theory inside which we are working (namely T_1 itself) includes arithmetical comprehension. But then T_1 would prove $\mathrm{Con}(T_1)$, which contradicts Gödel's Second Incompleteness Theorem. (For a formal proof, see Lindström [23] Chapter 7 Corollary 1 p. 97). Note that (25) follows immediately from (24) and definition (23). $\square$

In what follows, we will apply this proposition to $T_1 = \mathtt{ACA}_0$ itself or $T_1 = \Pi^1_1 - \mathtt{CA}_0$, both of which are known to be finitely axiomatizable (cf. Simpson [35] Lemma VIII.1.5 pp. 311-312 and Lemma VI.1.1 pp. 217-218).

The major previous results on the interpretability strength of the subsystems of $\mathtt{PA}^2$, $\mathtt{HP}^2$, $\mathtt{BL}^2$ can be described as follows. In the 19th Century, Frege in essence showed that $\mathtt{PA}^2 \leq_{\mathrm{I}} \mathtt{HP}^2$ (cf. Frege [14], [4], Boolos and Heck [7]), and recently Heck ([21] p. 192) and Linnebo ([24] p. 161) noted that Frege's proofs in fact show that $\Pi^1_1 - \mathtt{CA}_0 \leq_{\mathrm{I}} \Pi^1_1 - \mathtt{HP}_0$ (cf. § 2.2, Corollary 21). Further, Boolos ([3]) showed that the converse holds (cf. Corollary 23), so that one has $\Pi^1_1 - \mathtt{CA}_0 \equiv_{\mathrm{I}} \Pi^1_1 - \mathtt{HP}_0$ (cf. Corollary 24). Heck ([20]) then showed that $\mathtt{ABL}_0$ interprets Robinson's $\mathtt{Q}$, and Ganea and Visser ([16], [40]) independently showed that the converse holds, so that $\mathtt{ABL}_0 \equiv_{\mathrm{I}} \mathtt{Q}$. Likewise, Burgess ([8]) showed that $\mathtt{AHP}_0$ interprets Robinson's Q. Finally, Ferreira and Wehmeier ([13]) showed that $\Delta^1_1 - \mathtt{BL}_0$ is consistent and a slight modification of their proof shows that $\Sigma^1_1 - \mathtt{LB}_0$ is consistent, and inspection of this proof shows that $\Sigma^1_1 - \mathtt{LB}_0 <_{\mathrm{I}} \Pi^1_1 - \mathtt{CA}_0$. These previous results and our new results are summarized in Figure 2, where the double arrows indicate that the provability relation is irreversible, and where the single arrows indicate that the provability relation may or may not be irreversible. That is, in the diagram $T_1 \Rightarrow T_0$ means $T_0 <_{\mathrm{I}} T_1$ and $T_1 \to T_0$ means $T_0 \leq_{\mathrm{I}} T_1$.

Our new results establish upper and lower bounds on consistent subsystems of $\mathtt{BL}^2$ and $\mathtt{HP}^2$ by (i) finding new constructions of models of these theories, (ii) noting that the constructions can be formalized in theories such as $\mathtt{ACA}_0$ and $\Pi^1_1 - \mathtt{CA}_0$, and (iii) applying Proposition 7. Our first main new result, Theorem 53, is a construction of a model $\mathcal{M}$ of $\Sigma^1_1 - \mathtt{LB}_0$ using ideas from higher recursion theory (cf. Sacks [33] Part A). This structure $\mathcal{M}$ models a finite extension of $\Sigma^1_1 - \mathtt{LB}_0$ called $\Sigma^1_1 - \mathtt{LB}_0 + \mathrm{Inf}$ which interprets $\Sigma^1_1 - \mathtt{AC}_0$. Moreover, since this construction is formalizable in $\Pi^1_1 - \mathtt{CA}_0$, we have that Proposition 7 implies that $\Sigma^1_1 - \mathtt{LB}_0 + \mathrm{Inf} <_{\mathrm{I}} \Pi^1_1 - \mathtt{CA}_0$.

Our second set of results concerns new constructions of models of $\Delta^1_1 - \mathtt{BL}_0$ and $\Sigma^1_1 - \mathtt{PH}_0$ and $\Delta^1_1 - \mathtt{HP}_0 + \neg\Sigma^1_1 - \mathtt{PH}_0$. These results are all based on a generalization of
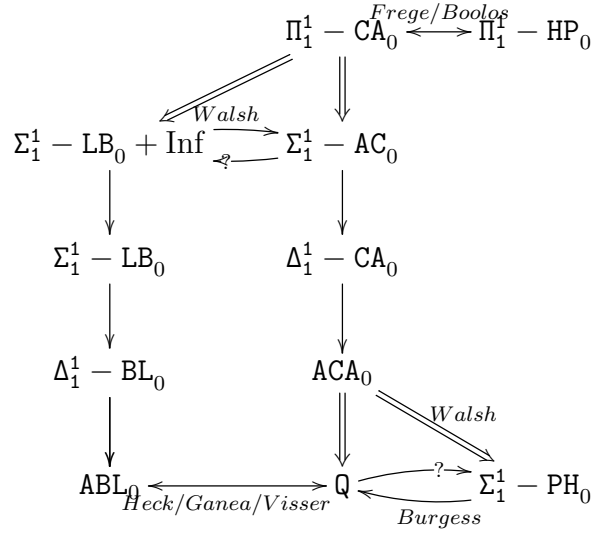
Figure 2: Interpretability Relation in Subsystems of $\mathtt{BL}^2$, $\mathtt{PA}^2$, and $\mathtt{HP}^2$

a theorem of Barwise-Schlipf and Ferreira-Wehmeier which allows us to built models of these theories on top of various recursively saturated structures (cf. Theorem 63). In particular, we show that if k is a countable recursively saturated o-minimal expansion of a real-closed field, then then there is a function $\# : D(k) \to k$, where $D(k^n)$ denotes the definable subsets of k^n, such that the structure

$$(k, D(k), D(k^2), \ldots, \#) \tag{26}$$

is a model of $\Sigma_1^1 - \mathtt{PH}_0$. Moreover, we note that this construction can be formalized in $\mathtt{ACA}_0$ for fields with $\mathtt{ACA}_0$-provable quantifier elimination, so that by Proposition 7, we have $\Sigma_1^1 - \mathtt{PH}_0 <_\mathrm{I} \mathtt{ACA}_0$ (cf. Corollary 92). Further, we show that if k is a saturated algebraically closed field, then there is a there is a function $\# : D(k) \to k$, where $D(k^n)$ denotes the definable subsets of k^n, such that the structure

$$(k, D(k), D(k^2), \ldots, \#) \tag{27}$$

is a model of $\Delta_1^1 - \mathtt{HP}_0 + \neg \Sigma_1^1 - \mathtt{PH}_0$. Further, we can use this construction to answer an open question of Linnebo (cf. Remark 74 and Proposition 76). However, we do not presently know whether this construction can be formalized in $\mathtt{ACA}_0$, although we have reduced it to the question of whether Ax's Theorem can be formalized in $\mathtt{ACA}_0$ (cf. Remark 71 and Question 104). Finally, we show that if k is a countable recursively saturated separably closed field of finite imperfection degree, then there is a function $\partial : D(k) \to k$, where $D(k^n)$ denotes the definable subsets of k^n, such that the structure

$$(k, D(k), D(k^2), \ldots, \partial) \tag{28}$$

is a model of $\Delta_1^1 - \mathtt{BL}_0$ (cf. Theorem 101). However, we do not presently know whether this construction can be formalized in $\mathtt{ACA}_0$, although we have reduced this question to the question of whether the proof of the elimination of imaginaries for separably closed fields can be formalized in $\mathtt{ACA}_0$ (cf. Remark 102 and Question 105).

11

2. Standard Models of $\mathtt{HP}^2$ and Associated Results

Prior to turning to the primary results of this paper in §§ 3-4, the relationship between $\mathtt{PA}^2$ and $\mathtt{HP}^2$ is briefly explored in this section. On the one hand, in § 2.2, a brief self-contained proof of Frege and Boolos's result that $\mathtt{PA}^2$ and $\mathtt{HP}^2$ are mutually interpretable is presented (cf. Corollary 24). On the other hand, in § 2.1, some of the ways in which the standard models of $\mathtt{HP}^2$ are similar to and different from the standard models of $\mathtt{PA}^2$ are examined. The standard model of $\mathtt{PA}^2$ is the structure from equation (1), namely, $(\omega, 0, s, +, \times, \leq, P(\omega))$, while the standard models of $\mathtt{HP}^2$ are the structures from equation (7), namely, structures of the form $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$, where α is an ordinal which is not a cardinal and where $\#_\alpha : P(\alpha) \to \alpha$ denotes cardinality. In § 2.1, it is shown that these standard models of $\mathtt{HP}^2$ depend only on the cardinality of α for $\alpha \geq \omega + \omega$ (Proposition 10 (i)), and further that they can have many automorphisms, unlike the standard model of $\mathtt{PA}^2$ (cf. Proposition 11 (iv)). Finally, it is shown that there is an analogue of the relative categoricity of $\mathtt{PA}^2$ in the setting of $\mathtt{HP}^2$ (cf. Proposition 14 and Remark 15).

2.1. Models of $\mathtt{HP}^2$ from Infinite Cardinals

Proposition 8. Suppose α, β are ordinals that are not cardinals, and consider the structures $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ and $(\beta, P(\beta), P(\beta^2), \ldots, \#_\beta)$, where $\#_\alpha : P(\alpha) \to \alpha$ and $\#_\beta : P(\beta) \to \beta$ denote cardinality.

   (i) The structures $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ and $(\beta, P(\beta), P(\beta^2), \ldots, \#_\beta)$ model $\mathtt{HP}^2$.
  (ii) If $\alpha = \omega + k + 1$ where $k \geq 0$, then $|\alpha - \mathrm{rng}(\#_\alpha)| = k$
 (iii) If $\alpha \geq \omega + \omega$, then $|\alpha - \mathrm{rng}(\#_\alpha)| = |\alpha|$.
  (iv) The structures $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ and $(\beta, P(\beta), P(\beta^2), \ldots, \#_\beta)$ are isomorphic if and only if $\alpha = \beta$ or $\alpha, \beta \geq \omega + \omega$ and $|\alpha| = |\beta|$.

Proof. For (i), note that restricting attention to ordinals α which are not cardinals serves the purpose of ensuring that $\#(\alpha) < \alpha$, so that $\mathrm{dom}(\#_\alpha)$ is $P(\alpha)$ and so that $\mathrm{rng}(\#_\alpha)$ is a subset of α. Further, note that $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ satisfies Hume's Principle by the definition of cardinality. Further, note that by the Power Set Axiom and the Separation Axiom, the structure $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ satisfies the full comprehension schema. Hence, in fact $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ is a model of $\mathtt{HP}^2$.

For (ii), note that $\alpha - \mathrm{rng}(\#_\alpha) = \{\omega + 1, \ldots, \omega + k\}$, which has cardinality k.

For (iii), note that since $\alpha \geq \omega + \omega$, we have that $\alpha - \omega$ is infinite, and hence $|\alpha| = |\alpha - \omega|$. Case One: α is a limit ordinal. Then the mapping from $\alpha - \omega$ to $\alpha - \mathrm{rng}(\#_\alpha)$ given by $\beta \mapsto \beta + 1$ is an injection. Case Two: α is a successor ordinal. Then $\alpha = \gamma + n$ where $n > 0$ and γ is a limit ordinal. Then $|\alpha| = |\alpha - \omega| = |\gamma - \omega|$. Then the mapping from $\gamma - \omega$ to $\alpha - \mathrm{rng}(\#_\alpha)$ given by $\beta \mapsto \beta + 1$ is an injection. Hence in both cases we have $|\alpha - \mathrm{rng}(\#_\alpha)| = |\alpha|$.

For (iv), suppose that the two structures are isomorphic. Then this isomorphism induces a bijection from α onto β, and hence α and β have the same cardinality. Further, suppose for the sake of contradiction that $\alpha \neq \beta$ and it is not the case that $\alpha, \beta \geq \omega + \omega$. If $\alpha < \beta < \omega + \omega$, then by part (ii) we have that $|\alpha - \mathrm{rng}(\#_\alpha)| < |\beta - \mathrm{rng}(\#_\beta)| < \omega$, and so the two structures are not elementarily equivalent and hence not isomorphic, which is a contradiction. If $\alpha < \omega + \omega \leq \beta$, then by parts (ii) and (iii) we have that

$|\alpha - \mathrm{rng}(\#_\alpha)| < \omega \leq |\beta - \mathrm{rng}(\#_\beta)|$, and so the two structures are not elementarily equivalent and hence not isomorphic, which is a contradiction. Hence, in fact, we must have that $\alpha = \beta$ or $\alpha, \beta \geq \omega + \omega$ and $|\alpha| = |\beta|$.

Conversely, suppose that $\alpha, \beta \geq \omega + \omega$ have the same cardinality, so that $\mathrm{rng}(\#_\alpha) = \mathrm{rng}(\#_\beta)$ by definition, and hence that $|\alpha - \mathrm{rng}(\#_\alpha)| = |\alpha| = |\beta| = |\beta - \mathrm{rng}(\#_\beta)|$ by part (iii). Hence choose a bijection $f : \alpha \to \beta$ such that $f(\gamma) = \gamma$ on $\mathrm{rng}(\#_\alpha)$. Extend f to a bijection $\overline{f} : P(\alpha) \to P(\beta)$ by setting $\overline{f}(X) = \{f(x) : x \in X\}$. Since $f(\gamma) = \gamma$ on $\mathrm{rng}(\#_\alpha)$ and since f is a bijection, we have that

$$\overline{f}(\#_\alpha(X)) = f(|X|) = |X| = |\{f(x) : x \in X\}| = \left|\overline{f}(X)\right| = \#_\beta(\overline{f}(X)) \qquad (29)$$

Hence, $\overline{f}$ is an isomorphism. $\qquad\square$

Definition 9. If κ is a cardinal, then define the ordinal

$$H_\kappa = \begin{cases} \omega + \kappa + 1 & \text{if } \kappa < \omega, \\ \omega + \omega & \text{if } \kappa = \omega \\ \kappa + 1 & \text{if } \kappa > \omega. \end{cases} \qquad (30)$$

and define the structure

$$\mathcal{H}_\kappa = (H_\kappa, P(H_\kappa), P(H_\kappa^2), \ldots, \#_\kappa) \qquad (31)$$

where $\#_\kappa : P(H_\kappa) \to H_\kappa$ denotes cardinality.

Proposition 10.
  (i) For every ordinal α that is not a cardinal, there is exactly one cardinal κ such that the structure $\mathcal{H}_\kappa$ is isomorphic to the structure $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$, where $\#_\alpha : P(\alpha) \to \alpha$ denotes cardinality.
  (ii) If κ is a cardinal then $|H_\kappa - \mathrm{rng}(\#_\kappa)| = \kappa$.
  (iii) If κ, λ are cardinals, then $\mathcal{H}_\kappa$ and $\mathcal{H}_\lambda$ are isomorphic if and only if $\kappa = \lambda$.

Proof. For (ii), there are three cases. First, suppose that $\kappa = k < \omega$. Then $H_\kappa - \mathrm{rng}(\#_\kappa) = \{\omega + 1, \ldots, \omega + k\}$. Second, suppose that $\kappa = \omega$. Then $H_\kappa - \mathrm{rng}(\#_\kappa) = \{\omega + n : 0 < n < \omega\}$. Third, suppose that $\kappa > \omega$. Then by Proposition 8 (iii), $|H_\kappa - \mathrm{rng}(\#_\kappa)| = |\kappa + 1 - \mathrm{rng}(\#)| = |\kappa + 1| = \kappa$.

For (iii), note that the right-to-left direction is trivial. For the left-to-right direction, suppose for the sake of contradiction that $\mathcal{H}_\kappa$ and $\mathcal{H}_\lambda$ are isomorphic and that $\kappa \neq \lambda$. Then without loss of generality, $\kappa < \lambda$. First suppose that $\kappa < \lambda < \omega$. Then part (ii) implies that $\mathcal{H}_\kappa$ and $\mathcal{H}_\lambda$ are not elementarily equivalent, since $\mathcal{H}_\kappa$ models that there are exactly κ elements not in the range of $\#$, whereas $\mathcal{H}_\kappa$ models that there are exactly λ elements not in the range of $\#$. Second suppose that $\kappa < \omega \leq \lambda$. Then likewise the structures $\mathcal{H}_\kappa$ and $\mathcal{H}_\lambda$ are not elementarily equivalent, since $\mathcal{H}_\kappa$ models that there are exactly κ many elements not in the range of $\#$, whereas $\mathcal{H}_\lambda$ models that there are at least $\kappa + 1$ many elements not in the range of $\#$. Third, suppose that $\kappa = \omega < \lambda$. But this cannot happen, since the isomorphism from $\mathcal{H}_\kappa$ and $\mathcal{H}_\lambda$ would induce a bijection between the first-order parts of these structures, which, respectively, have cardinality ω and $\lambda > \omega$. Fourth, suppose that $\omega < \kappa < \lambda$. Again this cannot happen, since the

isomorphism from $\mathcal{H}_\kappa$ and $\mathcal{H}_\lambda$ would induce a bijection between the first-order parts of these structures, which respectively, have cardinality κ and $\lambda > \kappa$.

For (i), note that uniqueness follows from part (iii). For existence, there are two cases. If $\alpha < \omega + \omega$, then $\alpha = \omega + k + 1$ where $k \geq 0$. Then of course the structure $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ is identical with the structure $\mathcal{H}_k$. If $\alpha \geq \omega + \omega$, then by Proposition 8 (iv), we have that $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#_\alpha)$ is isomorphic to $H_{|\alpha|}$. $\square$

Proposition 11. Suppose that κ is a cardinal.

(i) If $\beta, \gamma \in (H_\kappa - \mathrm{rng}(\#_\kappa))$ then there is $f \in \mathrm{Aut}(\mathcal{H}_\kappa)$ such that $f(\beta) = \gamma$.

(ii) If $X \subseteq H_\kappa$ is $\emptyset$-definable in $\mathcal{H}_\kappa$ then $X \subseteq \mathrm{rng}(\#_\kappa)$ or $(H_\kappa - \mathrm{rng}(\#_\kappa)) \subseteq X$.

(iii) If $\beta \in \mathrm{rng}(\#_\kappa)$ and $f \in \mathrm{Aut}(\mathcal{H}_\kappa)$ then $f(\beta) = \beta$.

(iv) $\mathrm{Aut}(\mathcal{H}_\kappa)$ and $\mathrm{Aut}(\kappa)$ are isomorphic, where we view κ as a structure in the empty signature.

Proof. (i) Let $f : H_\kappa \to H_\kappa$ by setting $f(\gamma) = \beta$, $f(\beta) = \gamma$, and let f be the identity otherwise, so that f is a bijection of H_κ. Extend f to a mapping $\overline{f} : \mathcal{H}_\kappa \to \mathcal{H}_\kappa$ by setting $\overline{f}(X) = \{f(x) : x \in X\}$. Then $\overline{f}$ is clearly a bijection since f is a bijection. To show that it is an automorphism of the structure $\mathcal{H}_\kappa$, it suffices to show that $\overline{f}(\#_\kappa X) = \#_\kappa \overline{f}(X)$. But, since f is the identity on $\mathrm{rng}(\#_\kappa)$, we have that $\overline{f}(\#_\kappa X) = f(\#_\kappa X) = \#_\kappa X$, and since f is a bijection, we have that $f \upharpoonright X : X \to \overline{f}(X)$ is a bijection, and so $\#_\kappa X = \#_\kappa \overline{f}(X)$. Hence, in fact $\overline{f}$ is an automorphism of $\mathcal{H}_\kappa$ which sends β to γ.

(ii) Suppose that $X \subseteq H_k$ is $\emptyset$-definable in $\mathcal{H}_\kappa$, but it is not the case that $X \subseteq \mathrm{rng}(\#_\kappa)$ or $(H_\kappa - \mathrm{rng}(\#_\kappa)) \subseteq X$. Then there is $\beta \in X \cap (H_\kappa - \mathrm{rng}(\#_\kappa))$ and $\gamma \in (H_\kappa - \mathrm{rng}(\#_\kappa)) \cap (H_\kappa - X)$. By part (i), there is $f \in \mathrm{Aut}(\mathcal{H}_\kappa)$ such that $f(\beta) = \gamma$. But since X is $\emptyset$-definable, we have that $\beta \in X$ if and only if $\gamma = f(\beta) \in X$, which is a contradiction.

(iii) Suppose that $\beta \in \mathrm{rng}(\#_\kappa)$ and $f \in \mathrm{Aut}(\mathcal{H}_\kappa)$ and $f(\beta) \neq \beta$. Since $\mathrm{rng}(\#_\kappa)$ is $\emptyset$-definable and $\beta \in \mathrm{rng}(\#_\kappa)$, we have that $f(\beta) \in \mathrm{rng}(\#_\kappa)$. Case One: $f(\beta) < \beta$. Note that the relation $<$ on $\mathrm{rng}(\#_\kappa)$ is $\emptyset$-definable, since on $\mathrm{rng}(\#_\kappa)$ we have

$$\lambda \leq \lambda' \Longleftrightarrow \mathcal{H}_\kappa \models \exists X \, \exists Y \, \#_\kappa(X) = \lambda \,\&\, \#_\kappa(Y) = \lambda' \,\&\, \exists \text{ injective } f : X \to Y \quad (32)$$

Then our case assumption $f(\beta) < \beta$ implies $f(f(\beta)) < f(\beta) < \beta$ and so we obtain an infinite decreasing sequence of ordinals, which is a contradiction. Case Two: $\beta < f(\beta)$. Since $f \in \mathrm{Aut}(\mathcal{H}_\kappa)$ we have that $f^{-1} \in \mathrm{Aut}(\mathcal{H}_\kappa)$, and since $\beta < f(\beta)$ we have $f^{-1}(\beta) < \beta$, since again the relation $<$ on $\mathrm{rng}(\#_\kappa)$ is $\emptyset$-definable. Hence, by iterating $f^{-1}(f^{-1}(\beta)) < f^{-1}(\beta) < \beta$ as before, we again obtain an infinite decreasing sequence of ordinals, which is a contradiction.

(iv) If X is a set viewed as a structure in the empty signature, then $\mathrm{Aut}(X)$ is just the set of permutations of X, and hence if X and Y have the same cardinality, then $\mathrm{Aut}(X)$ and $\mathrm{Aut}(Y)$ are isomorphic as groups. Hence by Proposition 10 (ii), we have that $\mathrm{Aut}(\kappa)$ and $\mathrm{Aut}(H_\kappa - \mathrm{rng}(\#))$ are isomorphic as groups. So it suffices to find a group isomorphism $F : \mathrm{Aut}(H_\kappa - \mathrm{rng}(\#)) \to \mathrm{Aut}(\mathcal{H}_\kappa)$.

To this end, given a bijection $f : H_\kappa \to H_\kappa$, extend f to a mapping $\overline{f} : \mathcal{H}_\kappa \to \mathcal{H}_\kappa$ by setting $\overline{f}(X) = \{f(x) : x \in X\}$, so that $\overline{f} : \mathcal{H}_\kappa \to \mathcal{H}_\kappa$ is a bijection. Then we claim that

$$\overline{f} \in \mathrm{Aut}(\mathcal{H}_\kappa) \Longleftrightarrow f \upharpoonright (\mathrm{rng}(\#_\kappa)) = \mathrm{id}_{\mathrm{rng}(\#_\kappa)} \quad (33)$$

The left-to-right direction follows directly from part (iii). For the right-to-left direction, it suffices to show that $\overline{f}(\#_\kappa X) = \#_\kappa \overline{f}(X)$. Since f is the identity on $\mathrm{rng}(\#_\kappa)$, we have that $\overline{f}(\#_\kappa X) = f(\#_\kappa X) = \#_\kappa X$, and since f is a bijection, we have that $f \restriction X : X \to \overline{f}(X)$ is a bijection, and so $\#_\kappa X = \#_\kappa \overline{f}(X)$. Hence, equation (33) does hold, and so we can define $F : \mathrm{Aut}(H_\kappa - \mathrm{rng}(\#_\kappa)) \to \mathrm{Aut}(\mathcal{H}_\kappa)$ by setting $F(g) = \overline{f}$, where f is g on $H_\kappa - \mathrm{rng}(\#_\kappa)$ and where f is the identity on $\mathrm{rng}(\#_\kappa)$. Since $F(g_1 \circ g_2) = F(g_1) \circ F(g_2)$, we have that F witnesses the group isomorphism between $\mathrm{Aut}(H_\kappa - \mathrm{rng}(\#_\kappa))$ and $\mathrm{Aut}(\mathcal{H}_\kappa)$. $\qquad\square$

Remark 12. The proof of the theorem above shows one how to construct many natural examples of sentences that are independent of HP^2. For instance, in equation (32), it was shown how to define the ordering in $\mathcal{H}_\kappa$. Using this, one can form a sentence φ such that $\mathcal{H}_\kappa \models \varphi$ if and only if κ is an infinite successor cardinal, so that $\mathcal{H}_{\omega_2} \models \mathrm{HP}^2 + \varphi$ and $\mathcal{H}_{\omega_\omega} \models \mathrm{HP}^2 + \neg\varphi$. This contrasts starkly with the case of PA^2, where there are comparatively few known examples of natural independent sentences.

Remark 13. The structures $\mathcal{H}_\kappa$ for $\kappa < \omega$ from Definition 9 are on one level very different: for, they are not elementarily equivalent since $\mathcal{H}_\kappa$ models that there are exactly κ-many elements that are not in the range of the #-function. However, on another level, these structures are very similar to each other: for, when $\kappa < \omega$, it is easy to see that $\mathcal{H}_\kappa$ is isomorphic to the structure $(\omega, P(\omega), P(\omega^2), \ldots, \#_\kappa^*)$, where $\#_\kappa^*(X) = 0$ if X is infinite and where $\#_\kappa^*(X) = \kappa + 1 + |X|$ if X is finite. Further, when one restricts to the ranges of the $\#_\kappa^*$-functions, the induced structures $(\mathrm{rng}(\#_\kappa^*), P(\omega) \cap P(\mathrm{rng}(\#_\kappa^*)), P(\omega) \cap P(\mathrm{rng}(\#_\kappa^*)^2), \ldots, \#_\kappa^*)$ are all isomorphic to the structure $(\omega, P(\omega), P(\omega^2), \ldots, \#^*)$ where $\#^*(X) = 0$ if X is infinite and where $\#^*(X) = 1 + |X|$ if X is finite. As the next theorem indicates, this is a very general phenomenon among models of HP^2: namely, so long as different #-functions on one and the same underlying set can in some sense see each other, they yield isomorphic structures when one restricts attention to their ranges.

Proposition 14. Suppose that $(M, S_1, S_2, \ldots, \#_1, \#_2)$ is a structure where $S_n \subseteq P(M^n)$ and where $\#_i : S_1 \to M$. Suppose further that the structures $(M, S_1, S_2, \ldots, \#_i)$ are models of HP^2 for $i \in \{1, 2\}$, and further that the structure $(M, S_1, S_2, \ldots, \#_1, \#_2)$ satisfies every instance of the comprehension schema (8), in the signature that includes both of the function symbols $\#_1, \#_2$. Finally, for $i \in \{1, 2\}$, define the following induced structure:

$$\mathcal{N}_i = (\mathrm{rng}(\#_i), S_1 \cap P(\mathrm{rng}(\#_i)), S_2 \cap P(\mathrm{rng}(\#_i)^2), \ldots, \#_i) \qquad (34)$$

Then $\mathcal{N}_1$ and $\mathcal{N}_2$ are isomorphic models of HP^2.

Proof. First we define a bijection $\Gamma : \mathrm{rng}\#_1 \to \mathrm{rng}\#_2$. If $\#_1 X \in \mathrm{rng}\#_1$ where $X \in S_1$, then we define $\Gamma(\#_1 X) = \#_2 X$. Note that $\Gamma : \mathrm{rng}\#_1 \to \mathrm{rng}\#_2$ is well-defined: if $\#_1 X = \#_1 Y$ then we need to show that $\#_2 X = \#_2 Y$. This follows, since

$$\#_1 X = \#_1 Y \implies [\exists \text{ bijection } f : X \to Y] \implies \#_2 X = \#_2 Y \qquad (35)$$

Next, note that $\Gamma : \mathrm{rng}\#_1 \to \mathrm{rng}\#_2$ is injective:

$$\Gamma(\#_1 X) = \Gamma(\#_1 Y) \implies \#_2 X = \#_2 Y \implies [\exists \text{ bijection } f : X \to Y] \implies \#_1 X = \#_1 Y \qquad (36)$$

Finally, note that $\Gamma : \mathrm{rng}\#_1 \to \mathrm{rng}\#_2$ is surjective: if $\#_2 X \in \mathrm{rng}\#_2$ then by definition $\Gamma(\#_1 X) = \#_2 X$. Hence, in fact $\Gamma : \mathrm{rng}\#_1 \to \mathrm{rng}\#_2$ is a bijection. Further, note that the graph of Γ is in S_2 since one has the equality

$$\mathrm{graph}(\Gamma) = \{(x, y) \in M^2 : \exists\, Z\ \#_1(Z) = x\ \&\ \#_2(Z) = y\} \tag{37}$$

and since it was assumed that the structure $(M, S_1, S_2, \ldots, \#_1, \#_2)$ satisfies every instance of the comprehension schema (8) in the signature that includes both of the function symbols $\#_1, \#_2$. Now, extend to $\overline{\Gamma} : \mathcal{N}_1 \to \mathcal{N}_2$ by setting $\overline{\Gamma}(X) = \{\Gamma(x) : x \in X\}$, which exists in S_1 since the graph of Γ is in S_2. Then $\overline{\Gamma} : \mathcal{N}_1 \to \mathcal{N}_2$ is an isomorphism, because

$$\overline{\Gamma}(\#_1 X) = \Gamma(\#_1 X) = \#_2 X = \#_2\{\Gamma(x) : x \in X\} = \#_2 \overline{\Gamma}(X), \tag{38}$$

where the first and second equalities follow respectively from the definitions of $\overline{\Gamma}$ and Γ, and where the third equality follows from the fact that $\Gamma : X \to \{\Gamma(x) : x \in X\}$ is a bijection whose graph is in S_2, and where the last equality follows from the definition of $\overline{\Gamma}$. $\qquad\square$

Remark 15. The previous proposition can be thought of as an analogue of the relative categoricity results for models of $\mathtt{PA}^2$. In the 19th Century, Dedekind showed that any two models $(M, +, \times, P(M), P(M^2), \ldots)$ and $(N, \oplus, \otimes, P(N), P(N^2), \ldots)$ of $\mathtt{PA}^2$ are isomorphic ([10] § 132, cf. Shapiro [34] Theorem 4.8 p. 82). However, it is not difficult to see that Dedekind's result can be relativized, in the following way: if $(M, +, \times, \oplus, \otimes, S_1, S_2, \ldots)$ is a structure where $S_n \subseteq P(M^n)$ such that $(M, +, \times, S_1, S_2, \ldots)$ and $(M, \oplus, \otimes, S_1, S_2, \ldots)$ are models of $\mathtt{PA}^2$ and such that $(M, +, \times, \oplus, \otimes, S_1, S_2, \ldots)$ satisfies every instance of the comprehension schema (8) in the signature of $+, \times, \oplus, \otimes$, then $(M, +, \times, S_1, S_2, \ldots)$ and $(M, \oplus, \otimes, S_1, S_2, \ldots)$ are isomorphic (cf. Parsons [31] § 49 pp. 279 ff). The previous proposition is simply the analogue of this phenomenon in the setting of $\mathtt{HP}^2$.

2.2. The Mutual Interpretability of $\mathtt{PA}^2$ and $\mathtt{HP}^2$

The goal of this section is to present a brief and self-contained proof of the result that $\mathtt{PA}^2$ is mutually interpretable with $\mathtt{HP}^2$ (Corollary 24). One half of this result, namely, the interpretability of $\mathtt{HP}^2$ in $\mathtt{PA}^2$ is due to Boolos (Corollary 23). The other half of the result, namely, the interpretability of $\mathtt{PA}^2$ in $\mathtt{HP}^2$ is now called Frege's Theorem, namely (Corollary 21). The proof of Frege's Theorem can be broken down into two steps: first, the proof that $\mathtt{PA}^2$ is interpretable in the theory consisting of (Q1)-(Q2) and the comprehension schema (3) (cf. Theorem 16), and second the argument that this latter theory is interpretable in $\mathtt{HP}^2$ (cf. Theorem 20). Elements of the first step can be found in Dedekind (cf. [10] § 72), and elements of this second step can be traced back to Frege (cf. Boolos and Heck [7]).

However, the modern presentation stems from Wright [43] pp. 154-169 (cf. also Boolos [5]). The warrant for including a proof of this result here is two-fold: (i) the proof presented here is slightly briefer than other published presentations, and (ii) the proof presented here is slightly different from other published presentations in that it is centered around the notion of Dedekind-finiteness, defined in terms of the lack of injective non-surjective functions, as opposed to Frege's ancestral notion (cf. the relation $X \not\prec X$ in Proposition 18 and Theorem 20).

The observations recorded in this section about the Π_n^1-comprehension schema are due to Heck ([21] p. 192) and Linnebo ([24] p. 161). The trick of defining the graph of addition and multiplication in terms of its initial segments in the proof of Theorem 16 is adapted from Burgess and Hazen [9] pp. 6-10, although their concern there was not with Frege's Theorem.

Theorem 16. $\mathtt{PA}^2$ is interpretable in the theory consisting of (Q1)-(Q2) and the comprehension schema (3). More generally, $\Pi_n^1 - \mathtt{CA}_0$ is interpretable in the theory consisting of (Q1)-(Q2) and the comprehension schema (3) restricted to Π_n^1-formulas for $n > 0$.

Proof. Suppose that we are working with structure $\mathcal{M} = (M, S_1, S_2, \ldots, 0, s)$ that satisfies (Q1)-(Q2) and the comprehension schema (3) restricted to Π_n^1-formulas for $n > 0$. In what follows, we will refer respectively to the element 0 and the function s as "zero" and "successor." It must be shown how to uniformly define a model of $\Pi_n^1 - \mathtt{CA}_0$ within this structure. We say that X in S_1 is *inductive* if it contains zero and is closed under successor. Let N be the intersection of all the inductive sets X in S_1, which exists in S_1 by Π_1^1-comprehension. Note that zero is in N by construction, and note that N is closed under successor: for, if a is in N then a is contained in every inductive set X, and by definition of inductive sets, it follows that the successor of a is contained in every inductive set X, which is to say that the successor of a is in N.

Hence, we can define the structure $\mathcal{N} = (N, S_1 \cap P(N), S_2 \cap P(N^2), \ldots, 0, s)$ uniformly within $\mathcal{M}$. This structure then satisfies (Q1)-(Q2) since $\mathcal{M}$ satisfies (Q1)-(Q2). Further, $\mathcal{N}$ satisfies the Mathematical Induction Axiom (2), since if $F \in S_1 \cap P(N)$ contains zero and is closed under successor, then $F \in S_1$ contains zero and is closed under successor, and so by definition of N, it follows that $N \subseteq F \subseteq N$. For (Q3), let X be the subset of N for which the conclusion holds, i.e., $X = \{a \in N : a \neq 0 \to \exists\, w \in N\ x = sw\}$. Clearly zero is in X, and suppose that $a \in X \subseteq N$: then of course $sa = sw$ for some $w \in N$, namely $w = a$, and hence $sa \in X$. Hence, by the Mathematical Induction Axiom (2), it follows that $X = N$. Finally, before turning to the remainder of the axioms of Robinson's Q, note that since $\mathcal{M}$ satisfies Π_n^1-comprehension, we have that $\mathcal{N}$ satisfies Π_n^1-comprehension as well, since the second-order parts of $\mathcal{N}$ are just the second-order parts of $\mathcal{M}$ restricted to subsets of N.

To verify axioms Q4-Q5 of Robinson's Q, we must first define addition. Let $x + y = z$ if and only if there is a graph of a partial function $G \subseteq N^3$ such that $(x, y, z) \in G \subseteq N^3$ and

$$(x, 0, x) \in G\ \&\ [(x, sy, z) \in G \to\ \exists\, w\ sw = z\ \&\ (x, y, w) \in G] \qquad (39)$$

That is, we define the graph of addition as the union of its initial segments. Note that this graph of addition exists by the Π_1^1-Comprehension Schema. Further, note that addition is well-defined on its domain. Suppose that G_0 and G_1 are partial functions which satisfy equation (39) and fix an arbitrary x and let $Y = \{y \in N : \forall\ z_0, z_1\ (x, y, z_0) \in G_0\ \&\ (x, y, z_1) \in G_1 \to z_0 = z_1\}$. Clearly, $0 \in Y$ and if $y \in Y$ and $(x, sy, z_0) \in G_0$ and $(x, sy, z_1) \in G_1$ then there is w_0, w_1 such that $sw_0 = z_0$ and $sw_1 = z_1$ and $(x, y, w_0) \in G_0$ and $(x, y, w_1) \in G_1$. Then since $y \in Y$ we have $w_0 = w_1$ and hence $z_0 = sw_0 = sw_1 = z_1$. Hence, in fact, addition is a well-defined function on its domain. To show that it is a total function, fix an arbitrary x and let $Y = \{y \in N : \exists\, z\ x + y = z\}$. Clearly, $0 \in Y$, since we can choose $G = \{(x, 0, x)\}$. Suppose that $y \in Y$, say, with $(x, y, z) \in G$. To see that $sy \in Y$, set $G' = G \cup \{(x, sy, sz)\}$. Then clearly G' also satisfies

17

equation (39). Hence, in fact, addition is a total function. Finally, the verification of Q4 and Q5 follows directly from our construction in equation (39). To verify Q6-Q7, just define multiplication analogously. □

Remark 17. Hence, it remains to show that the theory consisting of (Q1)-(Q2) and the comprehension schema (3) is interpretable in $\mathtt{HP}^2$. In preparation for this result (Theorem 20), we first record some elementary considerations in the following proposition.

Proposition 18. Suppose that $(M, S_1, S_2, \ldots, \#)$ models $\mathtt{AHP}_0$. For X, Y in S_1, define $X \prec Y$ if and only if there is injective non-surjective function $f : X \to Y$ such that $\mathrm{graph}(f)$ is in S_2. Then for $a, b \in M$ and X, U, A, B in S_1, it follows that

  (i) If $a \notin X$ and $X \cup \{a\} \prec X \cup \{a\}$ then $X \prec X$.
  (ii) If $a \notin X$ and $U \prec X \cup \{a\}$ then $U \prec X$ or $\#U = \#X$.
  (iii) If $a \in A, b \in B$, then $\#A = \#B$ if and only if $\#(A - \{a\}) = \#(B - \{b\})$
  (iv) If $X \neq \emptyset$ then $\emptyset \prec X$
  (v) $X \nprec \emptyset$

Proof. For (i), suppose that $f : X \cup \{a\} \to X \cup \{a\}$ is an injection that is not a surjection. If $f(X) \subseteq X$ and $f : X \to X$ is surjective, then $f(a) = a$ and hence $f : X \cup \{a\} \to X \cup \{a\}$ would be surjective, contrary to hypothesis; hence when $f(X) \subseteq X$, it must be the case that $f : X \to X$ is injective but not surjective. On the other hand, when $f(X) \nsubseteq X$ then say $f(y) = a$ where $y \in X$ and $f(a) = z \in X$, and hence define $g : X \to X$ by $g(y) = z$ and $g = f$ otherwise. Then g is injective and misses the same point that f does. Further, the graph of g exists by the arithmetical comprehension schema.

For (ii), suppose that $f : U \to X \cup \{a\}$ is an injection which is not a surjection. If $f(U) \subseteq X$ then $\#U = \#X$ when $f : U \to X$ is a bijection and $U \prec X$ otherwise. If $f(U) \nsubseteq X$ then say $f(y) = a$ and f misses $b \in X$, in which case we define an injective function $g : U \to X$ by $g(y) = b$ and $g = f$ otherwise. The graph of g exists by the arithmetical comprehension schema. If g is a bijection, then $\#U = \#X$ and $U \prec X$ otherwise.

For (iii), suppose that $a \in A$ and $b \in B$ and let us first establish the left-to-right direction. So suppose that $f : A \to B$ is a bijection. If $f(a) = b$ then $f \restriction (A - \{a\})$ is the desired bijection. If $f(a) = d$ for $d \neq b$ and $f(c) = b$ for $c \neq a$, then define a bijection $g : (A - \{a\}) \to (B - \{b\})$ by $g(c) = d$ and $g = f$ otherwise. The graph of this function g then exists by the arithmetical comprehension schema. Now let us establish the right-to-left direction. Suppose that $g : (A - \{a\}) \to (B - \{b\})$ is a bijection. Then define $f : A \to B$ by $f(a) = b$ and $f = g$ otherwise. Then the graph of f exists by the arithmetical comprehension schema and f is a bijection since g was a bijection.

For (iv), note that the "empty" binary relation witnesses that there is an injective non-surjective function from $\emptyset$ to X.

For (v), note that if $X \prec \emptyset$, then there would be an injective non-surjective function $f : X \to \emptyset$, which would imply that there was an element in $\emptyset \setminus \mathrm{rng}(f)$, which would imply that there was some element in $\emptyset$.

□

Remark 19. It is well-known that the chief difficulty in the proof of the following theorem is establishing the totality of the successor function (cf. remarks to this effect in

18

Wright [43] p. 161). Prior to looking at the proof, it is helpful to think about what happens on the standard models $(\alpha, P(\alpha), P(\alpha^2), \ldots, \#)$ from § 2.1, where α is an ordinal which is not a cardinal and where $\# : P(\alpha) \to \alpha$ is cardinality. It is easy to see that ω is uniformly definable in each of these structures. Further, it is easy to see that for each $n \in \omega$, it follows that

$$\{\#W : W \prec \{0, \ldots, n\}\} = \{0, \ldots, n\} \tag{40}$$

where as in the previous proposition, $X \prec Y$ if and only if there is injective non-surjective function $f : X \to Y$. From this we see that

$$\{0, \ldots, n\} \not\prec \{0, \ldots, n\} \ \& \ \#\{0, \ldots, n\} = \#\{\#W : W \prec \{0, \ldots, n\}\} \tag{41}$$

as well as

$$\begin{aligned} s(\#\{0, \ldots, n\}) &= s(n+1) = n + 2 = \#(\{0, \ldots, n\} \cup \{n+1\}) \\ &= \#(\{\#W : W \prec \{0, \ldots, n\}\} \cup \{\#(\{0, \ldots, n\})\}) \end{aligned} \tag{42}$$

The entire idea of the below proof is to show that we can replicate these considerations in arbitrary models of HP^2. So in such an arbitrary model, we will define an analogue N of ω, and for analogues X of $\{0, \ldots, n\}$, we will find that

$$s(\#X) = \#(\{\#W : W \prec X\} \cup \{\#X\}) \tag{43}$$

This, in any case, is the heuristic explanation of the proof of the totality of the successor function in the following theorem.

Theorem 20. The theory consisting of (Q1)-(Q2) and the comprehension schema (3) is interpretable in HP^2. More generally, the theory consisting of (Q1)-(Q2) and the comprehension schema (3) restricted to Π_n^1-formulas is interpretable in $\Pi_n^1 - \mathrm{HP}_0$ for $n > 0$.

Proof. Suppose that we are working with structure $\mathcal{M} = (M, S_1, S_2, \ldots, \#)$ that satisfies $\Pi_n^1 - \mathrm{HP}_0$. It must be shown how to uniformly define a model of (Q1)-(Q2) and the comprehension schema (3) restricted to Π_n^1-formulas. Define $0 = \#\emptyset$ and define $s(x, y)$ if and only if there is X, Y in S_1 such that $\#X = x, \#Y = y$, and there is $b \in Y$ such that $\#X = \#(Y - \{b\})$. That is, $s(x, y)$ says that x, y are respectively cardinalities of sets X, Y and the cardinality of X is equal to the cardinality of Y minus one point. Note that the relation s exists in S_2 by the Π_1^1-comprehension schema. In what follows, we will respectively refer to the element 0 and the relation s as "zero" and "successor," keeping in mind that formally s is a binary relation. Then say that X in S_1 is *inductive* if it contains zero and is closed under successors, that is, if $x \in X$ and $s(x, y)$ then $y \in X$. Then define N to be the intersection of all the inductive sets, so that N is in S_1 by the Π_1^1-comprehension schema. Now we show that (i) s is a well-defined function on its domain, and that (ii) s is a total function on N, and that (iii) s maps elements of N to elements of N, and that (iv) s satisfies axioms Q1-Q2 on N.

For (i), to see that s is well-defined, suppose that $s(x, y)$ and $s(x, z)$. Then $x = \#X$, $y = \#Y$, $z = \#Z$ and there exists $b \in Y, c \in Z$ such that $\#X = \#(Y - \{b\}) = \#(Z - \{c\})$. Then by the right-to-left direction of Proposition 18 (iii), it follows that $\#Y = \#Z$ and hence $y = \#Y = \#Z = z$. Hence, s is a well-defined function on its domain.

19

For (ii), recall from Proposition 18 that for X, Y in S_1, we say $X \prec Y$ if and only if there is an injective non-surjective function $f : X \to Y$ such that $\mathrm{graph}(f)$ is in S_2. Then by iterated applications of Π_1^1-comprehension, the following exist in S_2 and S_1 respectively

$$R = \{(\#W, \#X) : W \prec X\} \tag{44}$$
$$Z = \{\#X : X \not\prec X \ \& \ \exists Y \ (\forall w \ w \in Y \leftrightarrow (w, \#X) \in R) \ \& \ \#X = \#Y\} \tag{45}$$

Note that

$$Z = \{\#X : X \not\prec X \ \& \ \#X = \#(\{\#W : W \prec X\})\} \tag{46}$$

(It may be heuristically helpful to compare this with equation (41)). Suppose that $\#X$ is in Z. Then $X \not\prec X$ and $\#X = \#(\{\#W : W \prec X\})$. Then

$$s(\#X, \#(\{\#W : W \prec X\} \cup \{\#X\})) \tag{47}$$

(Likewise, it may be helpful to compare this with equation (43)). Hence, we have the inclusion $Z \subseteq \{x : \exists y \ s(x, y)\}$, and so it suffices to show that Z is inductive.

Clearly, $0 \in Z$. Suppose that $\#X$ is in Z, so that $X \not\prec X$ and $\#X = \#(\{\#W : W \prec X\})$. Then $s(\#X, \#(\{\#W : W \prec X\} \cup \{\#X\}))$. Since successor is well-defined on its domain by part (i), it suffices to show that $\#(\{\#W : W \prec X\} \cup \{\#X\})$ is in Z. We have $\{\#W : W \prec X\} \not\prec \{\#W : W \prec X\}$. Since $\#X \notin \{\#W : W \prec X\}$, it follows from Proposition 18 (i) that $\{\#W : W \prec X\} \cup \{\#X\} \not\prec \{\#W : W \prec X\} \cup \{\#X\}$. Hence, $\#(\{\#W : W \prec X\} \cup \{\#X\})$ satisfies the first conjunct of Z in equation (46). To see that $\#(\{\#W : W \prec X\} \cup \{\#X\})$ satisfies the second conjunct of Z in equation (46), it suffices to show that

$$\{\#W : W \prec X\} \cup \{\#X\} = \{\#U : U \prec \{\#W : W \prec X\} \cup \{\#X\}\} \tag{48}$$

For the left-to-right direction, suppose first that $W \prec X$. Since X is bijective with $\{\#W : W \prec X\}$, we have that $W \prec \{\#W : W \prec X\} \cup \{\#X\}$. Continuing with the left-to-right direction, suppose now that $\#U = \#X$. Since X is bijective with $\{\#W : W \prec X\}$, we have that $\#U = \#(\{\#W : W \prec X\})$ and hence $U \prec \{\#W : W \prec X\} \cup \{\#X\}$. For the right-to-left direction, suppose that $U \prec \{\#W : W \prec X\} \cup \{\#X\}$. Since $\#X \notin \{\#W : W \prec X\}$, we have by Proposition 18 (ii) that $\#U = \#(\{\#W : W \prec X\}) = \#X$ or $U \prec \{\#W : W \prec X\}$. Hence, in fact equation (48) holds. It follows that $\#(\{\#W : W \prec X\} \cup \{\#X\})$ is in Z. Hence, Z is an inductive set, and as mentioned at the close of the above paragraph, it thus follows that successor is a total function on N.

(iii) Now we show that successor maps elements of N to elements of N. Suppose that a is in N. Then by definition, a is contained in every inductive set, and by parts (i)-(ii), it follows that there is unique b such that $s(a, b)$, from which it follows that b is contained in every inductive set, so that b is contained in N as well. Hence, successor maps elements of N to elements of N.

(iv) Finally, we note that the successor function s satisfies axioms (Q1)-(Q2). To see that it satisfies (Q1), note that if $s\#X = 0 = \#\emptyset$, then $\emptyset$ would be bijective with a non-empty set, which is a contradiction. To see that it satisfies (Q2), suppose that $s\#X = s\#Y$. Then $s\#X = \#A$ where $\#X = \#(A - \{a\})$ for some $a \in A$ and $s\#Y = \#B$ where $\#Y = \#(B - \{b\})$ for some $b \in B$. Then the left-to-right direction of Proposition 18 (iii) implies that $\#X = \#(A - \{a\}) = \#(B - \{b\}) = \#Y$.

Putting this all together, we can uniformly define the structure $\mathcal{N} = (N, S_1 \cap P(N), S_2 \cap P(N^2), \dots, 0, s)$ which satisfies (Q1)-(Q2). Finally, note that since $\mathcal{M}$ satisfies Π_n^1-comprehension, we have that $\mathcal{N}$ satisfies Π_n^1-comprehension as well, since the second-order parts of $\mathcal{N}$ are just the second-order parts of $\mathcal{M}$ restricted to subsets of N. $\square$

Corollary 21. $\mathtt{PA}^2$ is interpretable in $\mathtt{HP}^2$. More generally, $\Pi_n^1 - \mathtt{CA}_0$ is interpretable in $\Pi_n^1 - \mathtt{HP}_0$ for $n > 0$.

Proof. This follows immediately from Theorem 20 and Theorem 16. $\square$

Remark 22. The following theorem was first noted by Boolos ([3]). We include here for the sake of having a relatively self-contained presentation of the main results in this area, and because we will use Boolos' construction to transfer facts about the provability relation from subsystems of $\mathtt{PA}^2$ to subsystems of $\mathtt{HP}^2$ (cf. the proofs of Proposition 46 and Proposition 48).

Theorem 23. $\mathtt{HP}^2$ is interpretable in $\mathtt{PA}^2$. More generally, $\Pi_n^1 - \mathtt{HP}_0$ is interpretable in $\Pi_n^1 - \mathtt{CA}_0$ for $n > 0$, and $\Sigma_1^1 - \mathtt{PH}_0$ is interpretable in $\Sigma_1^1 - \mathtt{AC}_0$ and $\mathtt{AHP}_0$ is interpretable in $\mathtt{ACA}_0$.

Proof. We begin with the proof of the interpretability of $\mathtt{AHP}_0$ in $\mathtt{ACA}_0$. We will note how this proof yields all the other results as well. Let us work in a model $\mathcal{M} = (M, S_1, S_2, \dots, \oplus, \otimes)$ of $\mathtt{ACA}_0$, where $S_n \subseteq P(M^n)$. We must show how to uniformly define a model of $\mathtt{AHP}_0$. Consider the model $\mathcal{N} = (M, S_1, S_2, \dots, \#)$ where $\#(X) = n + 1$ if $|X| = n$, and where $\#(X) = 0$ if X is infinite. Then $\mathcal{N}$ is clearly definable in $\mathcal{M}$ since the graph of X is arithmetically definable. Further, since this graph is arithmetically definable, it follows that $\mathcal{N}$ satisfies the arithmetical comprehension schema. Further, by Simpson [35] Lemma II.3.6 p. 70, $\mathtt{ACA}_0$ proves that any two infinite sets are bijective, so that $\mathcal{N}$ is a model of $\mathtt{AHP}_0$. Hence, in fact we have that $\mathtt{AHP}_0$ is interpretable in $\mathtt{ACA}_0$. Further, it is obvious from this construction that $\mathcal{N}$ will satisfy whatever comprehension schemas $\mathcal{M}$ satisfies. $\square$

Corollary 24. $\mathtt{PA}^2$ is mutually interpretable with $\mathtt{HP}^2$. More generally, $\Pi_n^1 - \mathtt{CA}_0$ is mutually interpretable with $\Pi_n^1 - \mathtt{HP}_0$ for $n > 0$.

Proof. This follows immediately from Corollary 21 and Theorem 23. $\square$

3. Standard Models of Subsystems of $\mathtt{BL}^2$ and Associated Results

The primary goal of this section is to study models of subsystems of $\mathtt{BL}^2$ that are standard in the sense that they have the form $(\omega, S_1, S_2, \dots, \partial)$, where the sets $S_n \subseteq P(\omega^n)$ all come from some antecedently fixed computational class (e.g. the recursive sets, the arithmetical sets, the hyperarithmetical sets, etc.). The main result of this section is Theorem 53 which gives a construction of a standard model of the hyperarithmetic subsystem of $\mathtt{BL}_0$ in terms of the hyperarithmetic subsets of natural numbers. Further, this construction isolates a certain sentence Inf (cf. Definition 51) such that $\Sigma_1^1 - \mathtt{AC}_0 \leq_I \Sigma_1^1 - \mathtt{LB}_0 + \mathrm{Inf} <_I \Pi_1^1 - \mathtt{CA}_0$ (cf. Corollary 54 and Figure 2).

In the preliminary section § 3.1, we record some elementary facts about arbitrary models of subsystems of $\mathtt{BL}^2$, focusing in particular on the fact that arbitrary models of

the hyperarithmetic subsystems of $\mathtt{BL}^2$ require the existence of injective non-surjective functions (cf. Proposition 31). Such functions are important both because they are used to define the sentence Inf (cf. Definition 51) and because such functions are not required to exist by the hyperarithmetic subsystems of $\mathtt{HP}^2$ (cf. Remark 30). Further, in the preliminary section § 3.2, we review some elementary facts about hyperarithmetic theory, which we will employ in § 3.3. We also use these facts to fill in some parts of the provability relation (cf. Propositions 40-46 and Figure 1). Finally, in § 3.3, we turn to the main results of this section, namely the aforementioned Theorem 53 and Corollary 54.

3.1. Generalities on Models of Subsystems of $\mathtt{BL}^2$

Proposition 25. Suppose that $Y \subseteq M$ is definable with parameters by an arithmetical formula in the structure $(M, S_1, S_2, \ldots, \partial)$ (resp. in the structure $(M, S_1, S_2, \ldots, \#)$). Then Y is definable with parameters by an arithmetical formula that does not contain any instances of ∂ (resp. does not contain any instances of $\#$).

Proof. If $Y \subseteq M$ is definable in $(M, S_1, S_2, \ldots, \partial)$ by an arithmetical formula φ, and if $\partial(P)$ appears in φ, then P is not free in φ but rather is a parameter from S_1 and hence $a = \partial(P)$ is a parameter from M. So, replacing parameters from S_1 with parameters from M, it follows that the set Y is also definable by an arithmetical formula that does *not* contain any instances of ∂. $\square$

Proposition 26. Suppose that M is a structure and $\partial : D(M) \to M$ is an injection, where $D(M^n)$ is the definable subsets of M^n. Then $(M, D(M), D(M^2), \ldots, \partial)$ is a model of $\mathtt{ABL}_0$.

Proof. It is a model of Basic Law V since ∂ is an injection (cf. discussion subsequent to (6)). Further, it satisfies the arithmetical comprehension schema, since if $X \subseteq M$ is defined by an arithmetical formula, then by Proposition 25 it is defined by an arithmetical formula which does not include any instances of ∂. Hence, since $D(M)$ is closed under arithmetical comprehension, it follows that X is in $D(M)$, so that the structure $(M, D(M), D(M^2), \ldots, \partial)$ satisfies the arithmetical comprehension schema. $\square$

Proposition 27. Suppose that $(M, S_1, S_2, \ldots, \partial)$ is a model of $\Delta^1_1 - \mathtt{BL}_0$. (a) Then there is a injective function $s : M \to M$ such that $s(x) = \partial(\{x\})$ and such that $\mathrm{graph}(s)$ is in S_2. (b) Further, there is a function $s : M^n \to M$ such that $s(x_1, \ldots, x_n) = \partial(\{x_1, \ldots, x_n\})$ and such that $\mathrm{graph}(s)$ is in S_{n+1}.

Proof. The proof of (b) is identical to the proof of (a), so we present only the proof of (a). It suffices to show three things: first, that the graph of this function is Δ^1_1, second that this function is well-defined and total, and third that the function is injective. Note that the following Σ^1_1 and Π^1_1-definitions of $s(x) = y$ agree:

$$[\exists\, X\ (\forall\, z\ z \in X \leftrightarrow z = x)\ \&\ \partial X = y] \Longleftrightarrow [\forall\, Y\ (\forall\, z\ z \in Y \leftrightarrow z = x) \to \partial Y = y] \quad (49)$$

Suppose that the left-hand-side of this equation holds and that $Y = \{x\}$. Then $Y = X$ and hence $\partial(Y) = \partial(X) = y$. Conversely, suppose that the right-hand-side of this equation holds. By arithmetical comprehension, form the set $X = \{x\}$. Then by the right-hand-side it is the case that $\partial(X) = y$. Hence, by Δ^1_1-comprehension, there is

an s such that $s(x,y)$ if and only if both the left-hand-side and the right-hand-side of the above equation holds with respect to x and y. To see that the function is well-defined, suppose that the left-hand-side holds both of x and y and of x and z. By arithmetical comprehension, form the set $Y = \{x\}$. Then the right-hand-side implies that $y = \partial(Y) = z$. Hence, the function is well-defined. Further, it is everywhere defined because given x one can use arithmetical comprehension to form $X = \{x\}$, and hence x and $y = \partial(X)$ will satisfy the right-hand-side. Finally, to see that the function X is injective, suppose that $s(x) = s(y)$. Then $\partial(\{x\}) = \partial(\{y\})$. By Basic Law V, it follows that $\{x\} = \{y\}$ and hence that $x = y$. $\qquad\square$

Remark 28. The following proposition generalizes the construction in the Russell Paradox (cf. Proposition (4)). Note that in the following proposition, the term rng∂ is employed to designate the range of the function ∂. However, this set need not exist in the second-order parts of any of the models under consideration, even though it is is defined by a Σ_1^1-formula in these models.

Proposition 29. Suppose that $(M, S_1, S_2, \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$. For every A in S_1 such that $A \subseteq \mathrm{rng}\partial$, there is B in S_1 such that $B \subseteq A$ and $\partial B \in \mathrm{rng}\partial - A$.

Proof. First we claim that for all x it is the case that

$$[\exists\, X\ x \in A\ \&\ \partial X = x\ \&\ x \notin X] \Longleftrightarrow [\forall\, Y\ x \in A\ \&\ (\partial Y = x \to x \notin Y)] \qquad (50)$$

Suppose that the left-hand-side holds, i.e., suppose that $x \in A\ \&\ \partial X = x\ \&\ x \notin X$, and further suppose that Y is such that $\partial Y = x$. Then $\partial X = x = \partial Y$ and Basic Law V implies that $X = Y$. Conversely, suppose that the right-hand-side holds, i.e., suppose it is the case that $\forall\, Y\ x \in A\ \&\ (\partial Y = x \to x \notin Y)$. Since $x \in A \subseteq \mathrm{rng}\partial$, there is X such that $\partial X = x$, and hence $x \notin X$. The claim is proved, and, hence, by the Δ_1^1-Comprehension Schema, there exists B such that $x \in B$ if and only if both the left-hand-side and right-hand-side of (50) hold with respect to x. Note that it follows automatically from the left-hand-side that $B \subseteq A$. So it remains to show that $\partial B \in \mathrm{rng}\partial - A$. Suppose not. Then $\partial B \in \mathrm{rng}\partial \cap A$. Then either $\partial B \in B$ or $\partial B \notin B$. If $\partial B \in B$ then by right-hand-side we have $\partial B \notin B$, which is a contradiction. If $\partial B \notin B$, then by the left-hand-side we have that $\forall\, X\ \partial B \notin A\ \vee\ \partial X \neq \partial B\ \vee\ \partial B \in X$. Applying this to $X = B$ we have that $\partial B \notin A\ \vee\ \partial B \neq \partial B\ \vee\ \partial B \in B$. Since by hypothesis we have that $\partial B \in \mathrm{rng}\partial \cap A$, we must conclude that $\partial B \in B$, which again contradicts our supposition. Hence, in fact, $\partial B \in \mathrm{rng}\partial - A$. $\qquad\square$

Remark 30. The following corollary is important because it shows that satisfying $\Delta_1^1 - \mathtt{BL}_0$ requires the existence of injective non-surjective functions. As we note in Proposition 32 and later in Corollary 73, this is not the case with $\mathtt{ABL}_0$ and $\Delta_1^1 - \mathtt{HP}_0$.

Corollary 31. Suppose that $(M, S_1, S_2, \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$. Then there is a injective non-surjective function $s : M \to M$ such that $\mathrm{graph}(s)$ is in S_2 and such that $s(x) = \partial(\{x\})$.

Proof. By Proposition 27 there is an injective function $s : M \to M$ such that $\mathrm{rng}(s) \subseteq \mathrm{rng}\partial$ and such that $\mathrm{graph}(s)$ is in S_2 and such that $s(x) = \partial(\{x\})$. By Proposition 29, there is B in S_1 such that $B \subseteq \mathrm{rng}(s)$ and $\partial B \in \mathrm{rng}\partial - \mathrm{rng}(s)$. Hence, $s : M \to M$ is not surjective. $\qquad\square$

Proposition 32. There is a structure $(M, S_1, S_2, \ldots)$ such that

(i) For any injection $\partial : S_1 \to M$ it is the case that $(M, S_1, S_2, \ldots, \partial)$ models both the theory $\mathtt{ABL}_0$ as well as the sentence that expresses that there are no injective non-surjective functions $f : M \to M$.

(ii) There is no injection $\partial : S_1 \to M$ such that $(M, S_1, S_2, \ldots, \partial)$ models $\Delta^1_1 - \mathtt{BL}_0$.

Proof. Let M be an algebraically closed field (cf. Marker [26] Example 4.3.10 p. 140) and let $S_n = D(M^n)$, i.e. the definable subsets of M^n. Suppose that $s : M \to M$ was an injective surjective function whose graph was in $S_2 = D(M^2)$. Then this implies that there is a definable injective non-surjective function $s : M \to M$, which contradicts Ax's Theorem (cf. Theorem 65). For (i), note that by Proposition 26, the structure $(M, S_1, S_2, \ldots, \partial)$ is a model of $\mathtt{ABL}_0$ for any injection $\partial : D(k) \to k$. For (ii), note that if there was such an injection $\partial : S_1 \to M$, then by Corollary 31, there would be an injective non-surjective $s : M \to M$ such that $\mathrm{graph}(s)$ is in S_2, which is a contradiction. $\qquad\square$

3.2. Hyperarithmetic Theory and Some Related Elementary Results

Definition 33. Suppose that $X, Y \in 2^\omega$. Then $X \leq_T Y$ if X is Turing computable from Y or if X is $\Delta^{0,Y}_1$. Further, $X \leq_a Y$ if X is arithmetical in Y or if there is $n > 0$ such that X is $\Delta^{0,Y}_n$. Finally, $X \leq_h Y$ if X is hyperarithmetic in Y or if X is $\Delta^{1,Y}_1$ (For computational definitions of these reducibilities and proofs that they correspond with the relevant definability notion, see respectively Soare [36] p. 64, Odifreddi [30] p. 375, Sacks [33] p. 44).

Definition 34. Suppose that $Y \in 2^\omega$. Then define

$$\mathrm{REC}(Y) = \{X \in 2^\omega : X \leq_T Y\} \tag{51}$$

$$\mathrm{ARITH}(Y) = \{X \in 2^\omega : X \leq_a Y\} \tag{52}$$

$$\mathrm{HYP}(Y) = \{X \in 2^\omega : X \leq_h Y\} \tag{53}$$

Further, let $\mathrm{REC} = \mathrm{REC}(\emptyset)$ and $\mathrm{ARITH} = \mathrm{ARITH}(\emptyset)$ and $\mathrm{HYP} = \mathrm{HYP}(\emptyset)$ (cf. Simpson [35] Remark I.7.5. p. 25, Example I.11.2 p. 39).

Remark 35. Recall that structures in the language of $\mathtt{HP}^2$ and $\mathtt{BL}^2$ have the form $(M, S_1, S_2, \ldots, \#)$, where $S_n \subseteq P(M^n)$ and $\# : S_1 \to M$ (cf. equation (4)). If $\# : \mathrm{HYP}(Y) \to \omega$, then $(\omega, \mathrm{HYP}(Y), \#)$ will be used as an abbreviation for the structure $(\omega, S_1, S_2, \ldots, \#)$, where $S_n \subseteq P(\omega^n)$ is the set of n-ary relations whose graph is in $\mathrm{HYP}(Y)$ under any standard computable pairing function. Similarly, in what follows, we will sometimes use the abbreviations $(\omega, \mathrm{REC}(Y), \#)$ and $(\omega, \mathrm{ARITH}(Y), \#)$.

Proposition 36. The relation $X \leq_h Y$ is Π^1_1.

Proof. See Sacks [33] p. 45. $\qquad\square$

Theorem 37. (Kleene's Theorem on Restricted Quantification) Suppose that $\varphi(X, Y)$ is a Π^1_1 predicate. Then $\exists\, X \leq_h Y \; \varphi(X, Y)$ is a Π^1_1-predicate. Moreover, this is provable in $\Pi^1_1 - \mathtt{CA}_0$.

Proof. See Kleene [22] and Moschovakis [29] Theorem 4D.3 p. 220. That this theorem is provable in $\Pi_1^1 - \mathtt{CA}_0$ was noted by Simpson [35] VIII.3.20 p. 330. $\square$

Theorem 38. (Spector-Gandy Theorem) Suppose that $\varphi(Y)$ is a Π_1^1-predicate. Then there is an arithmetic predicate $\psi(X, Y)$ such that $\varphi(Y) \leftrightarrow \exists\, X \leq_h Y\, \psi(X, Y)$.

Proof. See Spector and Gandy ([37], [15]), Sacks [33] Theorem III.3.5 p. 61 and Exercise III.3.13 p. 62. $\square$

Remark 39. The following proposition is non-trivial only because the second-order quantifiers must be evaluated with respect to the second-order part $S_1 \subseteq P(\omega)$ of the structure (ω, S_1) and not with respect to $P(\omega)$ itself. For instance, one cannot infer that $(\omega, \mathrm{HYP}(Y)) \models \neg\Pi_1^1 - \mathtt{CA}_0$ simply from the fact that $\mathcal{O}^Y$ is Π_1^1 but not Σ_1^1, since to say this is merely to say that $\mathcal{O}^Y$ is Π_1^1-definable but not Σ_1^1-definable on the structure $(\omega, P(\omega))$.

Proposition 40. Suppose that $Y \in 2^\omega$. Then $(\omega, \mathrm{ARITH}(Y)) \models \mathtt{ACA}_0 + \neg\Delta_1^1 - \mathtt{CA}_0$ and $(\omega, \mathrm{HYP}(Y)) \models \Sigma_1^1 - \mathtt{AC}_0 + \neg\Pi_1^1 - \mathtt{CA}_0$.

Proof. For the fact that $(\omega, \mathrm{ARITH}(Y)) \models \mathtt{ACA}_0$, see Simpson [35] Theorem VIII.1.13 p. 313. Suppose that $(\omega, \mathrm{ARITH}(Y)) \models \Delta_1^1 - \mathtt{CA}_0$. But note that

$$
\begin{aligned}
(n, m) \in Y^{(\omega)} &\Longleftrightarrow \exists\, X \in \mathrm{ARITH}(Y)\ X = \oplus_{i=1}^n Y^{(i)}\ \&\ m \in X \\
&\Longleftrightarrow \forall\, X \in \mathrm{ARITH}(Y)\ X = \oplus_{i=1}^n Y^{(i)} \to m \in X
\end{aligned}
\tag{54}
$$

and hence $Y^{(\omega)} \in \mathrm{ARITH}(Y)$, which would contradict Tarski's Theorem on Truth. Hence, in fact $(\omega, \mathrm{ARITH}(Y)) \models \neg\Sigma_1^1 - \mathtt{AC}_0$. For the fact that $(\omega, \mathrm{HYP}(Y)) \models \Sigma_1^1 - \mathtt{AC}_0$, see Simpson [35] Theorem VIII.4.5 p. 334 and Theorem VIII.4.8 p. 335. This proof uses Kleene's Theorem on Restricted Quantification 37, and below in Theorem 53 we will emulate this proof in the setting of $\mathtt{BL}^2$. Suppose for the sake of contradiction that $(\omega, \mathrm{HYP}(Y)) \models \Pi_1^1 - \mathtt{CA}_0$. Since $\mathcal{O}^Y$ is $\Pi_1^{1,Y}$, by the Spector-Gandy Theorem 38, there is an arithmetic predicate $\psi(n, X, Y)$ such that $n \in \mathcal{O}^Y \Longleftrightarrow \exists\, X \leq_h Y\, \psi(n, X, Y)$. Then $\mathcal{O}^Y$ is Σ_1^1-definable on $(\omega, \mathrm{HYP}(Y))$ and hence exists in $\mathrm{HYP}(Y)$ by $\Pi_1^1 - \mathtt{CA}_0$, which contradicts that $\mathcal{O}^Y$ is not in $\mathrm{HYP}(Y)$. $\square$

Corollary 41. Suppose that there is a Π_1^1-formula $\theta(X, Y, Z)$ such that for all $Z \in 2^\omega$ the set $G_Z = \{(X, Y) \in 2^\omega \times 2^\omega : \theta(X, Y, Z)\}$ is the graph of a function $g_Z : \mathrm{HYP}(Z) \to \mathrm{HYP}(Z)$. Then the graph G_Z of g_Z is Σ_1^1-definable in the structure $(\omega, \mathrm{HYP}(Z))$ uniformly in Z.

Proof. Note that since $g_Z : \mathrm{HYP}(Z) \to \mathrm{HYP}(Z)$, we have that for all $X, Y, Z \in 2^\omega$

$$
\theta(X, Y, Z) \Longrightarrow X \oplus Y \leq_h Z
\tag{55}
$$

By the Spector-Gandy Theorem 38, there is an arithmetical predicate $\psi(X, Y, Z, W)$ such that for all $X, Y, Z \in 2^\omega$

$$
\theta(X, Y, Z) \Longleftrightarrow \exists\, W \leq_h X \oplus Y \oplus Z\ \psi(X, Y, Z, W)
\tag{56}
$$

Putting the two previous equations together, we have that for all $X, Y, Z \in 2^\omega$

$$\theta(X, Y, Z) \Longleftrightarrow \exists\, W \leq_h Z\; \psi(X, Y, Z, W) \tag{57}$$

Then for all $X, Y, Z \in 2^\omega$

$$g_Z(X) = Y \Longleftrightarrow (\omega, \mathrm{HYP}(Z)) \models \exists\, W\; \psi(X, Y, Z, W) \tag{58}$$

Hence, in fact the graph G_Z of g_Z is Σ_1^1-definable in the structure $(\omega, \mathrm{HYP}(Z))$ uniformly in Z. $\qquad\square$

Theorem 42. (Kondo's Uniformization Theorem) Suppose that $\varphi(X, Y)$ is a Π_1^1 predicate. Then there is a Π_1^1-predicate $\varphi'(X, Y)$ such that

$$\forall\, X, Y\; [\varphi'(X, Y) \rightarrow \varphi(X, Y)] \tag{59}$$
$$\forall\, X\; [\exists\, Y\; \varphi(X, Y)] \rightarrow [\exists! Y\; \varphi'(X, Y)] \tag{60}$$

Moreover, this is provable in $\Pi_1^1 - \mathtt{CA}_0$.

Proof. See Moschovakis [29] pp. 235-236. Simpson notes that Kondo's theorem is provable in $\Pi_1^1 - \mathtt{CA}_0$ (cf. [35] Theorem VI.2.6 p. 225). $\qquad\square$

Remark 43. The following two propositions use some of the preceding material to fill in some information about the provability relation (cf. Figure 1).

Proposition 44. There are models of $\mathtt{ABL}_0 + \neg\Delta_1^1 - \mathtt{BL}_0$.

Proof. Choose any injection $\partial : \mathrm{ARITH} \rightarrow \omega$. Then by Proposition 26 the structure $(\omega, \mathrm{ARITH}, \partial)$ is a model of $\mathtt{ABL}_0$. Further, since the graphs of addition and multiplication are in ARITH, if $(\omega, \mathrm{ARITH}, \partial) \models \Delta_1^1 - \mathtt{BL}_0$, then one would have that $\emptyset^{(\omega)} \in \mathrm{ARITH}$ (cf. equation (54)), which would contradict Tarski's theorem on truth. $\qquad\square$

Remark 45. The construction in the following proposition is the same construction as Boolos used to prove the interpretability of $\mathtt{HP}^2$ in $\mathtt{PA}^2$ (cf. the proof of Theorem 23).

Proposition 46. There are models of $\mathtt{AHP}_0 + \neg\Delta_1^1 - \mathtt{HP}_0$ and $\Sigma_1^1 - \mathtt{PH}_0 + \neg\Pi_1^1 - \mathtt{HP}_0$ and $\Delta_1^1 - \mathtt{HP}_0 + \neg\Sigma_1^1 - \mathtt{PH}_0$

Proof. Define a function $\# : \mathrm{ARITH} \rightarrow \omega$ by $\#X = 0$ if X is infinite and $\#X = |X| + 1$ if X is finite. By Simpson [35] Lemma II.3.6 p. 70, $\mathtt{ACA}_0$ proves that any two infinite sets are bijective, and hence $(\omega, \mathrm{ARITH}, \#)$ is a model of Hume's Principle. Further, it satisfies the arithmetical comprehension schema, since if $X \subseteq \omega$ is defined by an arithmetical formula, then by Proposition 25 it is defined by an arithmetical formula that does not include any instances of $\#$. Hence, since ARITH is closed under arithmetical comprehension, it follows that X is in ARITH, so that the structure $(\omega, \mathrm{ARITH}, \#)$ satisfies the arithmetical comprehension schema. Since $\emptyset^{(\omega)} \notin \mathrm{ARITH}$ but $\emptyset^{(\omega)}$ is Δ_1^1-definable over ARITH using the graphs of addition and multiplication as parameters (cf. equation (54)), we have that $(\omega, \mathrm{ARITH}, \#)$ is a model of $\mathtt{AHP}_0 + \neg\Delta_1^1 - \mathtt{HP}_0$. Similarly, using the fact that the graph of $\#$ is arithmetical, we can argue that $(\omega, \mathrm{HYP}, \#)$ is a model of $\Sigma_1^1 - \mathtt{PH}_0 + \neg\Pi_1^1 - \mathtt{HP}_0$. Likewise, Steel constructs a sequence of reals G_n such that $(\omega, \bigcup_{n=1}^{\infty} \mathrm{HYP}^{G_1 \oplus \cdots \oplus G_n})$ is a model of $\Delta_1^1 - \mathtt{CA}_0 + \neg\Sigma_1^1 - \mathtt{AC}_0$ ([38] Theorem 4 pp. 68 ff), and we can argue as before that $(\omega, \bigcup_{n=1}^{\infty} \mathrm{HYP}^{G_1 \oplus \cdots \oplus G_n}, \#)$ is a model of $\Delta_1^1 - \mathtt{HP}_0 + \neg\Sigma_1^1 - \mathtt{PH}_0$. $\qquad\square$

Remark 47. The following two propositions use elementary considerations about arithmetical sets (cf. Definition 34) to record some observations about natural functions whose existence cannot be proven in $\mathtt{ABL_0}$ or $\mathtt{AHP_0}$. For the motivation for these propositions, see § 2.2, and in particular around equation (20). The only reason for including these propositions here (as opposed to earlier) is that it seemed prudent to delay their proof until the arithmetical sets had been introduced, which we did earlier in this section (cf. Definition 34). Note that the construction in the following proposition is analogous to the construction used by Boolos to prove the interpretability of $\mathtt{HP}^2$ in $\mathtt{PA}^2$ (cf. the proof of Theorem 23).

Proposition 48. There is a structure M and a function $\# : D(M) \to M$, where $D(M^n)$ is the definable subsets of M^n, such that $(M, D(M), D(M^2), \ldots, \#)$ is a model of $\mathtt{AHP_0}$, and further there is binary relation R in $D(M^2)$ such that the set $\{(n, m) : \#(R_n) = m\}$ does not exist in $D(M^2)$, where $R_n = \{x : Rnx\}$.

Proof. Let M be the standard model of first-order arithmetic $(\omega, +, \times)$ so that $D(M)$ are the arithmetical sets ARITH. Choose a real $Z \notin$ ARITH, such as $\emptyset^{(\omega)}$, and enumerate Z as $z_0, z_1, z_2, \ldots$. Define the function $\# : \text{ARITH} \to \omega$ by $\#(X) = z_n$ if X is finite and $|X| = n$ and define $\#(X) = z_\infty$ for some fixed $z_\infty \notin Z$ if X is infinite. This structure satisfies arithmetical comprehension, since if $X \subseteq M$ is defined by an arithmetical formula, then by Proposition 25 it is defined by an arithmetical formula which does not include any instances of $\#$. Hence, since $D(M)$ is closed under arithmetical comprehension, it follows that X is in $D(M)$, so that the structure $(M, D(M), D(M^2), \ldots, \#)$ satisfies the arithmetical comprehension schema. Further, by Simpson [35] Lemma II.3.6 p. 70, $\mathtt{ACA_0}$ proves that any two infinite sets are bijective, and hence the structure $(M, D(M), D(M^2), \ldots, \#)$ is a model of Hume's Principle. Hence, $(M, D(M), D(M^2), \ldots, \#)$ is a model of $\mathtt{AHP_0}$. Consider now the set $R = \{(n, m) : m < n\}$, which is clearly arithmetical and so exists in $D(M^2)$. Then $R_n = \{x : Rnx\} = \{0, \ldots, n-1\}$ and $\#(R_n) = z_n$. Then the set

$$\{(n, m) : \#(R_n) = m\} = \{(n, m) : z_n = m\} \tag{61}$$

is equal to the graph of $n \mapsto z_n$, which is not arithmetical: for, if it were arithmetical, then its range Z would be be arithmetical, which contradicts the hypothesis on Z. $\square$

Proposition 49. There is a structure M and an injection $\partial : D(M) \to M$, where $D(M^n)$ is the definable subsets of M^n, such that $(M, D(M), D(M^2), \ldots, \partial)$ is a model of $\mathtt{ABL_0}$, and further there is binary relation R in $D(M^2)$ such that the set $\{(n, m) : \partial(R_n) = m\}$ does not exist in $D(M^2)$, where $R_n = \{x : Rnx\}$.

Proof. Let M be the standard model of first-order arithmetic $(\omega, +, \times)$ so that $D(M)$ are the arithmetical sets ARITH. Choose a real $Z \notin$ ARITH, such as $\emptyset^{(\omega)}$, and enumerate Z as $z_0, z_1, z_2, \ldots$. Choose an injection $\partial : \text{ARITH} \to \omega$ such that $\partial(\{n\}) = z_n$, which we can do since Z is coinfinite (since it is not arithmetical). Then by Proposition 26, the structure $(M, D(M), D(M^2), \ldots, \partial)$ is a model of $\mathtt{ABL_0}$. Consider now the diagonal $R = \{(n, m) : n = m\}$ which is clearly arithmetical and so exists in $D(M^2)$. Then $R_n = \{x : Rnx\} = \{n\}$ and $\partial(R_n) = \partial(\{n\}) = z_n$. Then the set

$$\{(n, m) : \partial(R_n) = m\} = \{(n, m) : z_n = m\} \tag{62}$$

is equal to the graph of $n \mapsto z_n$, which is not arithmetical: for, if it were arithmetical, then its range Z would be be arithmetical, which contradicts the hypothesis on Z. $\square$

3.3. Standard Models of the Hyperarithmetic Subsystems of $\mathtt{BL}^2$

Remark 50. Recall from Proposition 27 that $\Delta_1^1 - \mathtt{BL}_0$ proves the existence of the graph of an injective function $s : M \to M$ such that $s(x) = \partial(\{x\})$. This function is is mentioned in the following axiom.

Definition 51. The following sentence Inf is a sentence in the signature of $\mathtt{BL}^2$:

$$\begin{aligned}
\text{Inf} \equiv{} & \exists\, s : M \to M \,[\forall\, x\ s(x) = \partial(\{x\})]\ \&\ \exists\, N\ [\partial(\emptyset) \in N\ \&\ \forall\, x\ x \in N \to sx \in N] \\
& \&\ \forall\, N'\ [\partial(\emptyset) \in N'\ \&\ \forall\, x\ x \in N' \to sx \in N'] \to N \subseteq N' \\
& \&\ \exists\, \oplus : N^2 \to N\ \exists\, \otimes : N^2 \to N\ \exists\, \preceq\, \subseteq\ N^2\ [(N, \partial(\emptyset), s, \oplus, \otimes, \preceq) \models (\mathrm{Q1}) - (\mathrm{Q8})]
\end{aligned}$$
(63)

Intuitively, Inf says that there is a smallest set N which contains the zero element $\partial(\emptyset)$ and which is closed under the successor function $s(x) = \partial(\{x\})$ and which has addition and multiplication functions $\oplus$ and $\otimes$ and an ordering relation $\preceq$ which satisfy the eight axioms of Robinson's Q.

Remark 52. The following theorem and its corollary is the main result of § 3. Recall that the Russell paradox showed that $\mathtt{BL}_0$ and $\Pi_1^1 - \mathtt{BL}_0$ is inconsistent (cf. Proposition (4)). Recently Ferreira and Wehmeier ([13]) showed that $\Delta_1^1 - \mathtt{BL}_0$ is consistent, using Barwise and Schlipf's recursively-saturated model construction. In § 4.1, we present a generalization of this construction (cf. Theorem 63), which we apply to $\Delta_1^1 - \mathtt{BL}_0$ and $\Delta_1^1 - \mathtt{HP}_0$ (cf. Proposition 76, Corollary 92, and Theorem 101). However, the recursively-saturated model construction does not provide one with natural models, simply because most natural structures are not recursively saturated (unless of course they are saturated *tout court*). Hence, this raises the question of whether there are natural models of $\Delta_1^1 - \mathtt{BL}_0$. The following theorem constructs a model of $\Delta_1^1 - \mathtt{BL}_0$ which is mutually interpretable with the minimal ω-model of $\Delta_1^1 - \mathtt{CA}_0$, namely, the model whose second-order part consists of the hyperarithmetic sets.

Theorem 53. For any real $Y \in 2^\omega$, there is a map $\partial_Y : \mathrm{HYP}(Y) \to \omega$ with $\Pi_1^{1,Y}$-graph such that (i) the structure $M_Y = (\omega, \mathrm{HYP}(Y), \partial_Y)$ is a model of (a) $\Sigma_1^1 - \mathtt{LB}_0$ and (b) the sentence Inf, and such that (ii) the two structures

$$M_Y = (\omega, \mathrm{HYP}(Y), \partial_Y), \quad (\omega, 0, s, +, \times, \leq, \mathrm{HYP}(Y))$$
(64)

are mutually interpretable uniformly in Y, in the following sense: (a) the map $\partial_Y : \mathrm{HYP}(Y) \to \omega$ is definable in $(\omega, \mathrm{HYP}(Y), 0, s, +, \times, \leq)$ uniformly in Y, and (b) an isomorphic copy H_Y of the structure $(\omega, \mathrm{HYP}(Y), 0, s, +, \times, \leq)$ is definable in the structure $M_Y = (\omega, \mathrm{HYP}(Y), \partial_Y)$ uniformly in Y. Moreover, all these facts are provable in $\Pi_1^1 - \mathtt{CA}_0$.

Proof. Define $P(Y \oplus X, n)$ iff $X \in \mathrm{HYP}(Y)$ and $n = \langle a, e \rangle$ is a hyperarithmetical-in-Y index of X:

$$P(Y \oplus X, \langle a, e \rangle) \equiv X \in \mathrm{HYP}(Y)\ \&\ a \in \mathcal{O}^Y\ \&\ X = \{e\}^{H_a^Y}$$
(65)

Since the relation $X \in \mathrm{HYP}(Y)$ is Π_1^1 and membership in H_a^Y is $\Delta_1^{1,Y}$ for $a \in \mathcal{O}^Y$, we have that $P(Y \oplus X, n)$ is a Π_1^1-predicate. By Kondo uniformization (Theorem 42), there

is a Π_1^1-uniformization P' of P. For $Y \in 2^\omega$, define $\partial_Y(X) = n$ if and only if $P'(Y \oplus X, n)$. Since $\partial_Y(X) = n$ implies that n is a hyperarithmetical-in-Y index of X, we have that $\partial_Y : \mathrm{HYP}(Y) \to \omega$ is an injection and hence $M_Y = (\omega, \mathrm{HYP}(Y), \partial_Y)$ is a model of Basic Law V. Note that since $\partial_Y : \mathrm{HYP}(Y) \to \omega$ has a $\Pi_1^{1,Y}$-graph, the Corollary to the Spector-Gandy Theorem (cf. Corollary 41) implies that $\partial_Y : \mathrm{HYP}(Y) \to \omega$ is definable in the structure $(\omega, \mathrm{HYP}(Y), 0, s, +, \times, \leq)$, and this establishes (ii)(a).

To establish (i)(a), note that since $\partial_Y : \mathrm{HYP}(Y) \to \omega$ is an injection, it follows that $M_Y = (\omega, \mathrm{HYP}(Y), \partial_Y)$ is a model $\mathtt{ABL}_0$ (as in the proof of Proposition 26). To see that it also models the Σ_1^1-choice schema (18), suppose that $M_Y \models \forall z \, \exists X \, \varphi(z, X, \partial_Y(X))$, where φ is an arithmetical formula. (The proof for the case where z is replaced by a tuple $\bar{z}$, or where there are multiple existential set quantifiers and multiple existential relation quantifiers, or where there are parameters from the model present in φ is exactly similar). Then $M_Y \models \forall z \, \exists X \, \exists e \, [\partial_Y(X) = e \, \wedge \, \varphi(z, X, e)]$. Define a relation $Q(Y \oplus \{z\}, X)$ as follows:

$$Q(Y \oplus \{z\}, X) \Longleftrightarrow X \in \mathrm{HYP}(Y) \,\&\, \exists e \, [\partial_Y(X) = e \, \wedge \, \varphi(z, X, e)] \qquad (66)$$

Then Q is a Π_1^1-predicate. By Kondo uniformization, there is a Π_1^1-uniformization Q' of Q. For $Y \in 2^\omega$, define $q_Y(z) = X$ if and only if $Q'(Y \oplus \{z\}, X)$ and let

$$R_Y = \{(z, x) : \exists X \in \mathrm{HYP}(Y) \, q_Y(z) = X \wedge x \in X\} \qquad (67)$$

Then by Kleene's Theorem on Restricted Quantification 37, R_Y is $\Pi_1^{1,Y}$-definable. Moreover, since Q' is a uniformization, we also have

$$R_Y = \{(z, x) : \forall X \in \mathrm{HYP}(Y) \, q_Y(z) = X \to x \in X\} \qquad (68)$$

Again, by Kleene's Theorem on Restricted Quantification (37), the set R_Y is $\Sigma_1^{1,Y}$-definable. Hence R_Y is $\Delta_1^{1,Y}$ and so $R_Y \in \mathrm{HYP}(Y)$. Finally, since Q' is a uniformization, we have that $M_Y \models \forall z \, \varphi(z, (R_Y)_z, \partial_Y((R_Y)_z))$, so in fact M_Y is a model of $\Sigma_1^1 - \mathtt{BL}_0$ and this establishes (i)(a).

To show (i)(b) and (ii)(b), we first prove (ii)(b) and then note how our proof of (ii)(b) in fact establishes (i)(b). Recall that by Proposition 27, there is an injective function $s_Y : \omega \to \omega$ whose graph is in $\mathrm{HYP}(Y)$ such that $s_Y(n) = \partial_Y(\{n\})$ for all $n \in \omega$. Define an s_Y-recursive function $f_Y : \omega \to \omega$:

$$f_Y(0) = \partial_Y(\emptyset) \quad \& \quad f_Y(n+1) = s_Y(f_Y(n)) \qquad (69)$$

Let N_Y be the range of f_Y, so that both the graph of f_Y and its range N_Y are in $\mathrm{HYP}(Y)$. Since $N_Y = \mathrm{rng}(f_Y)$ and $\mathrm{dom}(f_Y) = \omega$, the following induction principle holds:

$$\forall P \, [f_Y(0) \in P \,\&\, \forall n \in \omega \, f_Y(n) \in P \to f_Y(n+1) \in P] \to N_Y \subseteq P \qquad (70)$$

Using this form of induction, one can show that $f_Y : \omega \to N_Y$ is injective, so that its inverse $f_Y^{-1} : N_Y \to \omega$ exists and is likewise in $\mathrm{HYP}(Y)$. Further, one can arithmetically define from N_Y, f_Y and f_Y^{-1} the functions $\oplus_Y : N_Y^2 \to N_Y$ and $\otimes_Y : N_Y^2 \to N_Y$ as follows:

$$f_Y(x) \oplus_Y f_Y(y) = f_Y(f_Y^{-1}(x) + f_Y^{-1}(y)) \qquad f_Y(x) \otimes_Y f_Y(y) = f_Y(f_Y^{-1}(x) \cdot f_Y^{-1}(y)) \qquad (71)$$

and then arithmetically define a relation $\preceq$ on N_Y^2 by

$$x \preceq_Y y \iff \exists \, z \in N_Y \; x \oplus_Y z = y \tag{72}$$

Further one can extend the map to $\overline{f}_Y : \mathrm{HYP}(Y) \to (P(N_Y) \cap \mathrm{HYP}(Y))$ by setting

$$\overline{f}_Y(X) = \{f_Y(n) : n \in X\} \tag{73}$$

and define the following structure in the signature of $(\omega, 0, s, +, \times, \leq, \mathrm{HYP}(Y))$:

$$H_Y = (N_Y, \partial_Y(\emptyset), s_Y, \oplus_Y, \otimes_Y, \preceq_Y, \overline{f}_Y(\mathrm{HYP}(Y))) \tag{74}$$

Then the functions f_Y and $\overline{f}_Y$ witness that the two structures $(\omega, 0, s, +, \times, \leq, \mathrm{HYP}(Y))$ and H_Y are isomorphic.

Further, note that H_Y is definable within M_Y: for, by the induction principle (70) one can show that N_Y is the unique smallest set containing $\partial_Y(\emptyset)$ and closed under s_Y, and using equation (71) and the induction principle (70) one can show that $\oplus_Y$ and $\otimes_Y$ are the unique functions on N_Y satisfying the following recursion clauses

$$x \oplus_Y \partial_Y(\emptyset) = x \qquad\qquad x \oplus_Y (s_Y(z)) = s_Y(x \oplus_Y z) \tag{75}$$
$$x \otimes_Y \partial_Y(\emptyset) = \partial_Y(\emptyset) \qquad\qquad x \otimes_Y (s_Y(z)) = (x \otimes_Y z) \oplus_Y x \tag{76}$$

Hence, since H_Y and $(\omega, \mathrm{HYP}(Y), 0, s, +, \times, \leq)$ are isomorphic and since H_Y is definable in M_Y, we have established (ii)(b). Finally, note by construction that the structure H_Y witnesses that M_Y is a model of the axiom Inf, so that we have established (i)(b). $\qquad\square$

Corollary 54. $\Sigma_1^1 - \mathtt{AC}_0 \leq_\mathrm{I} \Sigma_1^1 - \mathtt{LB}_0 + \mathrm{Inf} <_\mathrm{I} \Pi_1^1 - \mathtt{CA}_0$.

Proof. Note that $\Sigma_1^1 - \mathtt{AC}_0 \leq_\mathrm{I} \Sigma_1^1 - \mathtt{LB}_0 + \mathrm{Inf}$ because the sentence Inf (cf. Definition 51) literally provides an interpretation. To see that $\Sigma_1^1 - \mathtt{LB}_0 + \mathrm{Inf} <_\mathrm{I} \Pi_1^1 - \mathtt{CA}_0$, note that since the previous theorem can be proven in $\Pi_1^1 - \mathtt{CA}_0$, it follows that $\Pi_1^1 - \mathtt{CA}_0$ proves the consistency of $\Sigma_1^1 - \mathtt{LB}_0 + \mathrm{Inf}$. Thus the result follows from Proposition 7. $\qquad\square$

4. Barwise-Schlipf Models of the Hyperarithmetic Subsystems of $\mathtt{BL}^2$ and $\mathtt{HP}^2$

In this section, we turn to building models of subsystems of $\mathtt{BL}^2$ and $\mathtt{HP}^2$ on top of various recursively saturated fields. In particular, § 4.1 is devoted to the statement and proof of a generalization of a theorem of Barwise-Schlipf and Ferreira-Wehmeier (Theorem 63). Then in §§ 4.2-4.4 three applications of this theorem are presented. The major result here is Corollary 92, which says that $\Sigma_1^1 - \mathtt{PH}_0 <_\mathrm{I} \mathtt{ACA}_0$, and this fills in a key piece of Figure 2 about the interpretability relation.

4.1. Generalization of the Barwise-Schlipf/Ferreira-Wehmeier Metatheorems

The main theorem of this section (Theorem 63) is a generalization of the way in which Barwise-Schlipf ([2]) built models of $\Delta_1^1 - \mathtt{CA}_0$ on top of recursively saturated models of Peano arithmetic, and the way in which Ferreira-Wehmeier ([13]) built models of $\Delta_1^1 - \mathtt{BL}_0$ on top of recursively saturated structures. The new addition is the concept of a uniformly definable function $\partial : D(M) \to M$ (Definition 55). Subsequent to defining this notion, the definitions of definable skolem functions and recursively saturated structures are recalled, and then Theorem 63 is stated and proven.

Definition 55. Suppose that M is an L-structure and let $D(M^n)$ be the definable subsets of M^n. Then $\partial : D(M) \to M$ is *uniformly definable* if for all L-formula $\theta(x, \overline{y})$ with all free variables displayed and with a non-empty set $\overline{y}$ of parameter variables, there is an L-formula $\theta'(x, \overline{y})$ with the same free variables, such that $\{\partial(\theta(\cdot, \overline{a}))\} = \{b : M \models \theta'(b, \overline{a})\}$ for all $\overline{a} \in M$, i.e.:

$$\overline{a}, b \in M \implies [\partial(\theta(\cdot, \overline{a})) = b \iff M \models \theta'(b, \overline{a})] \tag{77}$$

Definition 56. Suppose that L is countable and that M is an L-structure and that $B \in 2^\omega$. Then $\partial : D(M) \to M$ is *B-computably uniformly definable* if it is uniformly definable and the map $\theta \mapsto \theta'$ is B-computable.

Definition 57. Suppose that M is an L-structure. Then M has *definable skolem functions* if for every definable set $P \subseteq M^{m+n}$ there is a definable set $P' \subseteq M^{m+n}$ such that

$$M \models \forall \overline{x}, \overline{y} \, [P'\overline{xy} \to P\overline{xy}] \tag{78}$$
$$M \models \forall \, \overline{x} \, [\exists \, \overline{y} \, P\overline{xy}] \to [\exists! \, \overline{y} \, P'\overline{xy}] \tag{79}$$

Remark 58. Note that in this definition, the parameters used to define P' may exceed those used to define P. Note also the obvious similarity between definable skolem functions and the uniformization results, such as Kondo's Uniformization Theorem 42, which we employed in Theorem 53. In particular, equations (78)-(79) are nearly identical to equations (59)-(60).

Definition 59. Suppose that M is an L-structure and $A \subseteq M$. A set of A-formulas $p(\overline{v})$ in finitely many variables $\overline{v}$ is *realized* in M if there is an $\overline{b}$ in M such that $M \models \theta(\overline{b})$ for every A-formula $\theta(\overline{v})$ in $p(\overline{v})$. A set of A-formulas $p(\overline{v})$ is *finitely realized* in M if every finite subset $p_0(\overline{v})$ of $p(\overline{v})$ is realized in M. The structure M is *saturated* if for every $A \subseteq M$ with $|A| < |M|$ and every set of A-formulas $p(\overline{v})$, if $p(\overline{v})$ is finitely realized in M then $p(\overline{v})$ is realized in M.

Definition 60. Suppose that L and M are countable and $B \in 2^\omega$. Then M is *B-recursively saturated* if for every finite $A \subseteq M$ and every B-computable set of A-formulas $p(\overline{v})$, if $p(\overline{v})$ is finitely realized in M then $p(\overline{v})$ is realized in M.

Remark 61. The following proposition records the very elementary observation that saturated structures (resp. B-recursively saturated structures) have a kind of compactness property, in that each covering of M^n by definable sets has a finite sub-covering (resp. each B-recursive covering of M^n by definable sets has a finite sub-covering).

Proposition 62. Suppose that M is a saturated L-structure (resp. B-recursively saturated L-structure) and that $A \subseteq M$ with $|A| < |M|$. Further, suppose that $\{\theta_i(\overline{v})\}_{i \in I}$ is a set of A-formulas (resp. B-computable set of A-formulas). Then

$$[M \models \forall \, \overline{a} \bigvee_{i \in I} \theta_i(\overline{v})] \implies [\exists \text{ finite } I_0 \subseteq I \; M \models \forall \, \overline{a} \bigvee_{i \in I_0} \theta_i(\overline{v})] \tag{80}$$

Proof. The contrapositive of equation (80) says that if the set of A-formulas $p(\overline{v}) = \{\neg\theta_i(\overline{a}) : i \in I\}$ is finitely realized, then it is realized. $\square$

Theorem 63. Suppose that M is an L-structure and $\partial : D(M) \to M$ such that the structure $N = (M, D(M), D(M^2), \dots, \partial)$ models $\mathtt{ABL_0}$ (resp. $\mathtt{AHP_0}$). Suppose that $B \in 2^\omega$. Then

  (i) If $\partial : D(M) \to M$ is uniformly definable and M is saturated, then the structure N models $\Delta_1^1 - \mathtt{BL_0}$ (resp. $\Delta_1^1 - \mathtt{HP_0}$).

  (ii) If $\partial : D(M) \to M$ is uniformly definable and M is saturated, then the structure N models $\Sigma_1^1 - \mathtt{LB_0}$ (resp. $\Sigma_1^1 - \mathtt{PH_0}$) if and only if M has definable skolem functions.

  (iii) If $\partial : D(M) \to M$ is B-computably uniformly definable and M is B-recursively saturated, then the structure N models $\Delta_1^1 - \mathtt{BL_0}$ (resp. $\Delta_1^1 - \mathtt{HP_0}$).

  (iv) If $\partial : D(M) \to M$ is B-computably uniformly definable and M is B-recursively saturated, then the structure N models $\Sigma_1^1 - \mathtt{LB_0}$ (resp. $\Sigma_1^1 - \mathtt{PH_0}$) if and only if M has definable skolem functions.

Proof. In all four parts of this proof, the proof is identical between Basic Law V and Hume's Principle, and so we only include the proofs for the case of Basic Law V. Further, the proof of (i) and (iii) are parallel and the proof of (ii) and (iv) are parallel, and so we present the proofs of (i) and (iii) simultaneously and the proofs of (ii) and (iv) simultaneously. For (i) and (iii), suppose that $\partial : D(M) \to M$ is uniformly definable (resp. B-computably uniformly definable) and M is saturated (resp. B-recursively saturated). To see that N is a model of $\Delta_1^1 - \mathtt{BL_0}$, suppose that there is a subset Z of M^n which is defined on N by a Σ_1^1-formula $\varphi(\bar{z})$ and by a Π_1^1-formula $\psi(\bar{z})$. Let us suppose that $\varphi(\bar{z})$ and $\psi(\bar{z})$ use exactly one set parameter $A \in D(M)$ where

$$A = \{w \in M : M \models \rho(w, \bar{a})\} \tag{81}$$

and where $\rho(w, \bar{v})$ is an $\emptyset$-formula with $\bar{a} \in M$, since the proof in the case where there are multiple parameters, with some being objects, some sets, and some binary relations etc., is exactly identical. Further, let us suppose that $\varphi(\bar{z}) \equiv \exists\, X\, \varphi_0(\bar{z}, X, \partial(X), A)$ and that $\psi(\bar{z}) \equiv \forall\, X\, \psi_0(\bar{z}, X, \partial(X), A)$, since the proof in the case where there are multiple existential (resp. universal) set-quantifiers or relation-quantifiers in $\varphi(\bar{z})$ (resp. $\psi(\bar{z})$) is exactly identical. Then

$$\bar{z} \in Z \iff N \models \exists\, X\, \varphi_0(\bar{z}, X, \partial(X), A) \iff N \models \forall\, X\, \psi_0(\bar{z}, X, \partial(X), A) \tag{82}$$

Then

$$N \models \forall\, \bar{z}\, \exists\, X\, \varphi_0(\bar{z}, X, \partial(X), A) \vee \neg\psi_0(\bar{z}, X, \partial(X), A) \tag{83}$$

Let us abbreviate

$$\xi_0(\bar{z}, X, \partial(X), A) \equiv \varphi_0(\bar{z}, X, \partial(X), A) \vee \neg\psi_0(\bar{z}, X, \partial(X), A) \tag{84}$$

so that equation (83) becomes

$$N \models \forall\, \bar{z}\, \exists\, X\, \xi_0(\bar{z}, X, \partial(X), A) \tag{85}$$

Then this translates into M as

$$M \models \forall\, \bar{z} \bigvee_{\theta(x,\bar{y})} \exists\, \bar{b}\, \xi_0(\bar{z}, \theta(\cdot, \bar{b}), \partial(\theta(\cdot, \bar{b})), \rho(\cdot, \bar{a})) \tag{86}$$

where $\theta(x, \overline{y})$ ranges over $\emptyset$-formulas with non-empty set of parameter variables $\overline{y}$. Since the map $\partial : D(M) \to M$ is uniformly definable (resp. B-computably uniformly definable) via the map $\theta \mapsto \theta'$, we have

$$M \models \forall \, \overline{z} \bigvee_{\theta(x, \overline{y})} \exists \, \overline{b} \, \exists \, c \, (\theta'(c, \overline{b}) \; \& \; \xi_0(\overline{z}, \theta(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))) \tag{87}$$

Since M is saturated (resp. B-recursively saturated), an application of Proposition 62 implies that there is $K > 0$ and there are $\emptyset$-formulas $\theta_1(x, \overline{y}), \dots, \theta_K(x, \overline{y})$ such that

$$M \models \forall \, \overline{z} \bigvee_{i=1}^{K} \exists \, \overline{b} \, \exists \, c \, (\theta_i'(c, \overline{b}) \; \& \; \xi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))) \tag{88}$$

Then by definition of ξ_0 (cf. equation 84)), we have:

$$M \models \forall \, \overline{z} \bigvee_{i=1}^{K} \exists \, \overline{b} \, \exists \, c \, (\theta_i'(c, \overline{b}) \; \& \; (\varphi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a})) \vee \neg \psi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a})))) \tag{89}$$

It follows from equation (82) that

$$Z = \{\overline{z} \in M^n : M \models \bigvee_{i=1}^{K} \exists \, \overline{b} \, \exists \, c \, (\theta_i'(c, \overline{b}) \; \& \; (\varphi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))))\} \tag{90}$$

Hence $Z \in D(M^n)$ and so N satisfies $\Delta_1^1 - \mathtt{BL}_0$. Hence, this completes the proof of parts (i) and (iii).

We turn to the proofs of parts (ii) and (iv). First, we handle the proof of the right-to-left direction, which is quite similar to the proof from the above paragraph. Suppose that $\partial : D(M) \to M$ is uniformly definable (resp. B-computably uniformly definable) and M is saturated (resp. B-recursively saturated) and has definable skolem functions. To see that N is a model of $\Sigma_1^1 - \mathtt{LB}_0$, suppose that

$$N \models \forall \, \overline{z} \, \exists \, X \, \xi_0(\overline{z}, X, \partial(X), A) \tag{91}$$

where ξ_0 is arithmetical and where $A \in D(M)$ is a set parameter with

$$A = \{w \in M : M \models \rho(w, \overline{a})\} \tag{92}$$

and where $\rho(w, \overline{v})$ is an $\emptyset$-formula with $\overline{a} \in M$. (As in the proof in the previous paragraph, the case of multiple parameters or multiple set or relation quantifiers is exactly similar). Then equation (91) translates into M as

$$M \models \forall \, \overline{z} \bigvee_{\theta(x, \overline{y})} \exists \, \overline{b} \, \xi_0(\overline{z}, \theta(\cdot, \overline{b}), \partial(\theta(\cdot, \overline{b})), \rho(\cdot, \overline{a})) \tag{93}$$

where $\theta(x, \overline{y})$ ranges over $\emptyset$-formulas with non-empty set of parameter variables $\overline{y}$. Since $\partial : D(M) \to M$ is uniformly definable (resp. B-computably uniformly definable) via the map $\theta \mapsto \theta'$, we have

$$M \models \forall \, \overline{z} \bigvee_{\theta(x, \overline{y})} \exists \, \overline{b} \, \exists \, c \, (\theta'(c, \overline{b}) \; \& \; \xi_0(\overline{z}, \theta(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))) \tag{94}$$

Since M is saturated (resp. B-recursively saturated), an application of Proposition 62 implies that there is $K > 0$ and there are $\emptyset$-formulas $\theta_1(x, \overline{y}), \ldots, \theta_K(x, \overline{y})$ such that

$$M \models \forall \, \overline{z} \bigvee_{i=1}^{K} \exists \, \overline{b} \, \exists \, c \, (\theta_i'(c, \overline{b}) \,\&\, \xi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))) \tag{95}$$

Then by adding dummy variables if need be, we can move the disjunction to the right as follows:

$$M \models \forall \, \overline{z} \, \exists \, \overline{b} \, \exists \, c \bigvee_{i=1}^{K} (\theta_i'(c, \overline{b}) \,\&\, \xi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))) \tag{96}$$

and one can take the first such i as follows:

$$M \models \forall \, \overline{z} \, \exists \, \overline{b} \, \exists \, c \bigvee_{i=1}^{K} [(\theta_i'(c, \overline{b}) \,\&\, \xi_0(\overline{z}, \theta_i(\cdot, \overline{b}), c, \rho(\cdot, \overline{a}))) \,\&\, \bigwedge_{j < i} \neg(\theta_j'(c, \overline{b}) \,\&\, \xi_0(\overline{z}, \theta_j(\cdot, \overline{b}), c, \rho(\cdot, \overline{a})))] \tag{97}$$

Then since M has definable skolem functions, there is a possibly larger finite set of parameters $\overline{a}' \supseteq \overline{a}$ and $\overline{a}'$-definable functions f, g such that

$$M \models \forall \, \overline{z} \bigvee_{i=1}^{K} [(\theta_i'(g(\overline{z}), f(\overline{z})) \,\&\, \xi_0(\overline{z}, \theta_i(\cdot, f(\overline{z})), g(\overline{z}), \rho(\cdot, \overline{a})))$$
$$\&\, \bigwedge_{j < i} \neg(\theta_j'(g(\overline{z}), f(\overline{z})) \,\&\, \xi_0(\overline{z}, \theta_j(\cdot, f(\overline{z})), g(\overline{z}), \rho(\cdot, \overline{a})))] \tag{98}$$

Then there is a partition of M^n into the $\overline{a}'$-definable sets $P_1, \ldots, P_K$ which are defined as follows:

$$P_i = \{\overline{z} \in M^n : M \models [(\theta_i'(g(\overline{z}), f(\overline{z})) \,\&\, \xi_0(\overline{z}, \theta_i(\cdot, f(\overline{z})), g(\overline{z}), \rho(\cdot, \overline{a})))$$
$$\&\, \bigwedge_{j < i} \neg(\theta_j'(g(\overline{z}), f(\overline{z})) \,\&\, \xi_0(\overline{z}, \theta_j(\cdot, f(\overline{z})), g(\overline{z}), \rho(\cdot, \overline{a})))]\} \tag{99}$$

Then define the $\overline{a}'$-definable relation

$$R = \{(\overline{z}, w) : \bigvee_{i=1}^{K} [\overline{z} \in P_i \rightarrow \theta_i(w, f(\overline{z}))]\} \tag{100}$$

so that

$$\overline{z} \in P_i \implies R_{\overline{z}} = \{w \in M : (\overline{z}, w) \in R\} = \{w \in M : M \models \theta_i(w, f(\overline{z}))\} = \theta_i(\cdot, f(\overline{z})) \tag{101}$$

$$\overline{z} \in P_i \implies \{\partial(R_{\overline{z}})\} = \{\partial(\theta_i(\cdot, f(\overline{z})))\} = \{c \in M : M \models \theta_i'(c, f(\overline{z}))\} = \{g(\overline{z})\} \tag{102}$$

$$\overline{z} \in P_i \implies \partial(R_{\overline{z}}) = g(\overline{z}) \tag{103}$$

Putting these things together and glancing back at the definition of P_i in equation (99) we have,

$$\overline{z} \in P_i \implies M \models (\theta_i'(g(\overline{z}), f(\overline{z})) \,\&\, \xi_0(\overline{z}, \theta_i(\cdot, f(\overline{z})), g(\overline{z}), \rho(\cdot, \overline{a}))) \implies N \models \xi_0(\overline{z}, R_{\overline{z}}, \partial(R_{\overline{z}}), A) \tag{104}$$

Since the sets $P_1, \ldots, P_K$ partition M^n we have

$$N \models \forall \, \overline{z} \, \xi_0(\overline{z}, R_{\overline{z}}, \partial(R_{\overline{z}}), A) \tag{105}$$

and this implies that N models $\Sigma^1_1 - \mathtt{BL}_0$. Hence we have established the right-to-left direction of (ii) and (iv).

We want to establish the left-to-right direction of (ii) and (iv). Suppose that $\partial : D(M) \to M$ is uniformly definable (resp. B-computably uniformly definable) and M is saturated (resp. B-recursively saturated) and that N models $\Sigma^1_1 - \mathtt{BL}_0$. Suppose that $P \subseteq M^{m+n}$ is definable, perhaps with a finite set $\overline{a}$ of parameters from M. Note that for every $\overline{x} \in M^m$ with a tuple $\overline{y} \in M^n$ such that $P\overline{xy}$, we can arbitrarily choose one such $\overline{y} \in M^n$ and form the $\overline{y}$-definable singleton $\{\overline{y}\}$. This implies that

$$N \models \ \forall \, \overline{x} \, \exists \, R \, [(\exists! \, \overline{y} \, P\overline{xy}) \to ((\exists! \, \overline{y} \, R\overline{y}) \, \& \, (\forall \, \overline{y} \, R\overline{y} \to P\overline{xy}))] \tag{106}$$

Since $N \models \Sigma^1_1 - \mathtt{LB}_0$, one then has

$$N \models \exists \, P' \, \forall \, \overline{x} \, [(\exists \, \overline{y} \, P\overline{xy}) \to ((\exists! \, \overline{y} \, P'_{\overline{x}}\overline{y}) \, \& \, (\forall \, \overline{y} \, P'_{\overline{x}}\overline{y} \to P\overline{xy}))] \tag{107}$$

Since $P'_{\overline{x}}\overline{y}$ if and only if $P'\overline{xy}$, this implies that

$$N \models \exists \, P' \, \forall \, \overline{x} \, [(\exists \, \overline{y} \, P\overline{xy}) \to ((\exists! \, \overline{y} \, P'\overline{xy}) \, \& \, (\forall \, \overline{y} \, P'\overline{xy} \to P\overline{xy}))] \tag{108}$$

Finally, let $P'' = P' \cap P$. Then

$$M \models \forall \overline{x}, \overline{y} \, (P''\overline{xy} \to P\overline{xy}) \tag{109}$$
$$M \models \forall \, \overline{x} \, [(\exists \, \overline{y} \, P\overline{xy}) \to (\exists! \, \overline{y} \, P''\overline{xy})] \tag{110}$$

Hence, M has definable skolem functions. $\qquad\square$

4.2. Application to Algebraically Closed Fields

Remark 64. In this section, we apply Theorem 63 to construct models of $\Delta^1_1 - \mathtt{HP}_0$ on top of certain algebraically closed fields (cf. Theorem 70). The primary application of this construction is to answer a question posed by Linnebo (cf. Remark 74 and Theorem 76). Prior to doing this, we recall Ax's Theorem and note one elementary consequence of this theorem.

Theorem 65. (Ax's Theorem) Suppose that k is an algebraically closed field and $f : k \to k$ is a definable injective function. Then f is surjective.

Proof. See Ax [1] Theorem C pp. 241, 270 or Poizat [32] Lemma 4.3 pp. 70-71, in which is proved the stronger result wherein k is replaced by a definable subset of k^n. $\qquad\square$

Definition 66. A structure k is *strongly minimal* if every definable $X \subseteq k$ is finite or cofinite.

Proposition 67. Every algebraically closed field is strongly minimal.

Proof. See Marker [27] p. 5. $\qquad\square$

Proposition 68. Suppose that k is an algebraically closed field and that $X, Y \subseteq k$ are definable. Then the following are equivalent:

(i) There is definable bijection $f : X \to Y$

(ii) Either both X and Y are finite and of the same cardinality, or both X and Y are cofinite and $k \setminus X$ and $k \setminus Y$ are of the same cardinality.

Proof. Suppose that (i) holds. Then by strong minimality and the fact that an infinite set cannot be bijective with a finite set, either both X and Y are finite or both X and Y are cofinite. If X and Y are both finite then the fact that there is a definable bijection between them implies that X and Y have the same cardinality. If X and Y are both cofinite but $k \setminus X$ and $k \setminus Y$ are not of the same cardinality, then without loss of generality $k \setminus X = \{a_1, \dots, a_m\}$ and $k \setminus Y = \{b_1, \dots, b_n\}$ where $m < n$. Then define a function $\overline{f} : k \to k$ by $\overline{f} \restriction X = f$ and $f(a_i) = b_i$ for $i \leq m$. Then $f : k \to k$ is an injection that is not a surjection, since b_n is not the in the range of f. This contradicts Ax's Theorem 65. So, in fact, $k \setminus X$ and $k \setminus Y$ are of the same cardinality. Then (ii) holds.

Conversely, suppose that (ii) holds. If both X and Y are finite of the same cardinality, then simply enumerate the elements of X and Y and use these elements as parameters to define a bijection $f : X \to Y$. If X and Y are both cofinite and $k \setminus X$ and $k \setminus Y$ are of the same finite cardinality, then enumerate $k \setminus X = \{y_1, \dots, y_n\}$ and $k \setminus Y = \{x_1, \dots, x_n\}$. By renumbering, we can assume without loss of generality that $(k \setminus X) \cap (k \setminus Y) = \{x_1, \dots, x_m\} = \{y_1, \dots, y_m\}$ where $m \leq n$ and $x_1 = y_1, \dots, x_m = y_m$. If $m = n$ then this implies that $(k \setminus X) = (k \setminus Y)$ and $X = Y$, and we can choose the definable bijection $f : X \to Y$ to be the identity map. If $m < n$, then note that $\{x_{m+1}, \dots, x_n\} \subseteq X$ and $\{y_{m+1}, \dots, y_n\} \subseteq Y$ and $X \setminus \{x_{m+1}, \dots, x_n\} \subseteq Y$ and $Y \setminus \{y_{m+1}, \dots, y_n\} \subseteq X$. Then we can choose the definable bijection $f : X \to Y$ which is given by the identity on $X \setminus \{x_{m+1}, \dots, x_n\}$ and by $f(x_i) = y_i$ on $\{x_{m+1}, \dots, x_n\}$. $\square$

Proposition 69. Algebraically closed fields do not have definable skolem functions.

Proof. Let $\varphi(x, y) \equiv x = y^2$. Then $k \models \forall x \, \exists y \, x = y^2$. If k has definable skolem functions, then there is a definable function $f : k \to k$ such that $k \models \forall x \, x = (f(x))^2$. Then $\mathrm{rng}(f)$ is a definable set which includes exactly one square root for each $x \in k$. Then $\mathrm{rng}(f)$ is infinite and coinfinite, which contradicts strong minimality. $\square$

Theorem 70. Suppose that k is a saturated algebraically closed field of characteristic zero. Then there is a uniformly definable function $\# : D(k) \to k$ such that $(k, D(k), D(k^2), \dots, \#)$ is a model of $\Delta^1_1 - \mathrm{HP}_0 + \neg\Sigma^1_1 - \mathrm{PH}_0 + \neg\Pi^1_1 - \mathrm{HP}_0$. Further, there is *no* function $\partial : D(k) \to k$ such that $(k, D(k), D(k^2), \dots, \partial)$ is a model of $\Delta^1_1 - \mathrm{BL}_0$.

Proof. Since k is a field of characteristic zero, the prime field of k is $\mathbb{Q}$ and the integers $\mathbb{Z}$ are hence embedded into k via $\mathbb{Q}$. Using this embedding, define $\# : D(k) \to k$ by $\#X = |X|$ if X is finite and $\#X = -(|k \setminus X| + 1)$ if X is cofinite. Then by Proposition 68, the structure $(k, D(k), D(k^2), \dots, \#)$ is a model of Hume's Principle. To apply Theorem 63 (i)-(ii), we need to show that $\# : D(k) \to k$ is uniformly definable. Suppose that $\theta(x, \overline{y})$ is an $\emptyset$-formula with non-empty set $\overline{y}$ of parameter variables. Then by strong minimality, for any $\overline{a}$ we have that $\theta(\cdot, \overline{a})$ is finite or $\neg\theta(\cdot, \overline{a})$ is finite. Then

$$k \models \forall \overline{a} \bigvee_{N \geq 0} [|\theta(\cdot, \overline{a})| \leq N \vee |\neg\theta(\cdot, \overline{a})| \leq N] \tag{111}$$

36

Since k is saturated, by Proposition 62, there is an integer $N_\theta > 0$ such that

$$k \models \forall\, \overline{a}\, \bigvee_{i=0}^{N_\theta} [|\theta(\cdot, \overline{a})| \leq i \lor |\neg\theta(\cdot, \overline{a})| \leq i] \tag{112}$$

Then for each such formula $\theta(x, \overline{y})$ we define the following $\emptyset$-formula $\theta'(x, \overline{y})$ as follows:

$$\theta'(x, \overline{y}) \equiv \bigvee_{i=0}^{N_\theta} [|\theta(\cdot, \overline{y})| = i\ \&\ x = i] \lor [|\neg\theta(\cdot, \overline{y})| = i\ \&\ x = -(i+1)] \tag{113}$$

Hence, by definition, we have that for any $\overline{a}$

$$\{\#(\theta(\cdot, \overline{a}))\} = \{c : k \models \theta'(c, \overline{a})\} \tag{114}$$

The map $\# : D(k) \to k$ is uniformly definable. Hence, by Theorem 63 (i)-(ii) and Proposition 69, we have that $(k, D(k), D(k^2), \ldots, \#)$ is a model of $\Delta_1^1 - \mathtt{HP}_0 + \neg\Sigma_1^1 - \mathtt{PH}_0$. Further, since the set $\mathrm{rng}(\#) = \mathbb{Z}$ is definable by a Σ_1^1-formula in the structure $(k, D(k), D(k^2), \ldots, \#)$ but is *not* definable in k since k is strongly minimal, we have that $(k, D(k), D(k^2), \ldots, \#)$ is a model of $\neg\Pi_1^1 - \mathtt{HP}_0$.

Now let us note why there is *no* function $\partial : D(k) \to k$ such that $(k, D(k), D(k^2), \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$. If there was such a function, then by Corollary 31 it would follow that there was an injective non-surjective function $s : k \to k$ whose graph is in $D(k^2)$, which would contradict Ax's Theorem (65). $\qquad\square$

Remark 71. If we knew that all the parts of the proof of the above theorem were formalizable in $\mathtt{ACA}_0$, then we could infer from the proof of the above theorem and Proposition 7 that $\Delta_1^1 - \mathtt{HP}_0 <_\mathrm{I} \mathtt{ACA}_0$. It is clear from the proof that this comes down to determining whether or not Ax's Theorem 65 is provable in $\mathtt{ACA}_0$. However, note that in the next subsection, we will prove Corollary 92, which assures us that $\Delta_1^1 - \mathtt{HP}_0 <_\mathrm{I} \mathtt{ACA}_0$.

Remark 72. In conjunction with Corollary 31, the following corollary shows that there is a stark contrast between $\Delta_1^1 - \mathtt{HP}_0$ and $\Delta_1^1 - \mathtt{BL}_0$ on the score of whether they require the existence of injective non-surjective functions.

Corollary 73. There is a model $(M, S_1, S_2, \ldots, \#)$ of $\Delta_1^1 - \mathtt{HP}_0$ such that there is no injective non-surjective function $s : M \to M$ such that $\mathrm{graph}(s)$ is in S_2.

Proof. This follows immediately from the construction in Theorem 70 and Ax's Theorem 65. $\qquad\square$

Remark 74. Linnebo presented a description of properties that models of $\mathtt{AHP}_0$ and $\Delta_1^1 - \mathtt{HP}_0$ must have *if* they fail to model a certain sort of successor axiom ([24] pp. 164-165), and he additionally showed that there was a model of $\mathtt{AHP}_0$ which did not model this successor axiom ([24] Theorem 2 p. 164). Linnebo then remarked that it was unknown whether there was a model of $\Delta_1^1 - \mathtt{HP}_0$ that did not model the successor axiom (cf. [24] Remark 6 p. 168). Subsequent to defining this successor axiom, we now show that the model from the previous theorem does not model this axiom. We also explain why certain properties identified by Linnebo hold in this model.

Definition 75. The following are formulas in the language of HP^2 (cf. Linnebo [24] pp. 158-160):

(i) $P(n, m) \Longleftrightarrow \exists\, X, Y \; \#X = n \;\&\; \#Y = m \;\&\; \exists\, y \in Y \; X = Y \setminus \{y\}$

(ii) F is *hereditary* if Fn and $P(n, m)$ implies Fm

(iii) F is *closed* if $P(\#\emptyset, m)$ implies Fm

(iv) n is a *pseudo-number* if $n = \#\emptyset$ or n is contained in all hereditary, closed F.

(v) The successor axiom (SA) says that for any pseudo-number n, there is m such that $P(n, m)$.

Proposition 76. Suppose that k is a saturated algebraically closed field of characteristic zero. Suppose that $\# : D(k) \to k$ by $\#X = |X|$ if X is finite and $\#X = -(|k \setminus X| + 1)$ if X is cofinite. Then $(k, D(k), D(k^2), \ldots, \#) \models \Delta_1^1 - \text{HP}_0 + \neg\text{SA}$.

Proof. Before we begin, it is perhaps helpful to informally state the definition of $\#$ given above and describe how it interacts with the predicate $P(n, m)$. If X is a finite set with n elements, then $\#X = n$, and if X is a cofinite set with n elements in its complement, then $\#X = -(n + 1)$. So, for example, the set $X = \{\sqrt{2}, -1\}$ has $\#X = 2$, and the set $X = \{a \in k : k \models a^2 + 1 \neq 0\}$ has $\#X = -(2 + 1) = -3$, and the set $X = k$ has $\#X = -1$, and the set $X = \emptyset$ has $\#X = 0$. Further, if X is finite, then by choosing an element $y \notin X$, we have $P(\#X, \#(X \cup \{y\}))$. For example, if X is finite and has n elements and $y \notin X$, we have that $\#X = n$ and $\#(X \cup \{y\}) = n + 1$, so that $P(n, n+1)$. Conversely, if X is cofinite and has $n > 0$ elements in its complement and $y \notin X$, then we have that $X \cup \{y\}$ has $n - 1$ elements in its complement, so that $\#X = -(n+1) = -n-1$ and $\#(X \cup \{y\}) = -((n-1)+1) = -n$ and hence $P(-n-1, -n)$. For example, we have $P(0, 1), P(1, 2), P(2, 3), \ldots$ and $\ldots, P(-4, -3), P(-3, -2), P(-2, -1)$.

Now we begin the proof. In particular, we want to begin by describing what the hereditary, closed sets $F \in D(k)$ look like. So suppose that $F \in D(k)$ is hereditary and closed. First we claim that $\mathbb{N} \setminus \{0\} \subseteq F$. For, by the definition of $P(n, m)$ and $\#$, we have that F's being closed implies that $P(0, 1)$ and hence $1 \in F$. So suppose that $n \in (\mathbb{N} \setminus \{0\}) \cap F$. Then by the definition of $P(n, m)$ and $\#$, we have that F's being hereditary implies that $P(n, n+1)$ and hence $n + 1 \in F$. By induction, we have that if $F \in D(k)$ is hereditary and closed then $\mathbb{N} \setminus \{0\} \subseteq F$.

We want to claim that $\{n \in \mathbb{Z} : n \neq 0\} \subseteq F$. Suppose not. That is, suppose that there are some negative integers that are not in F. Then, since $F \in D(k)$ is infinite, strong minimality implies that F is co-finite. So there are at most finitely many negative integers that are not in F. Suppose that we write these negative integers in increasing order as $a_1 < a_2 < \cdots < a_n$. (E.g. if $\mathbb{Z} \setminus F = \{-5, -10, -12\}$ then $a_1 = -12, a_2 = -10$ and $a_3 = -5$). This implies that $a_1 - 1 \in F$. But then by the definition of $P(n, m)$ and $\#$, we have that F's being hereditary implies that $P(a_1 - 1, a_1)$ and hence Fa_1, which is a contradiction. Hence, in fact we have that $\{n \in \mathbb{Z} : n \neq 0\} \subseteq F$. So, what we have shown in this paragraph is that if $F \in D(k)$ is hereditary and closed, then $\{n \in \mathbb{Z} : n \neq 0\} \subseteq F$.

This, of course, implies that every element of $\mathbb{Z}$ is a pseudo-number. Conversely, it is not difficult to see that all the pseudo-numbers are elements of $\mathbb{Z}$. Suppose that $a \in k$ is not an integer. Then the set $F = k \setminus \{a\}$ is a hereditary closed set that does not contain a. Hence, what we have shown in this paragraph is that the pseudo-numbers in the structure $(k, D(k), D(k^2), \ldots, \#)$ are precisely the integers.

Now we are in a position to show that $(k, D(k), D(k^2), \ldots, \#) \models \neg\mathtt{SA}$. For, consider the set $k \in D(k)$. By definition $\#k = -(|k \setminus k| + 1) = -1$. Hence, by the results of the previous paragraph, we have that $\#k$ is a pseudo-number. So suppose that $\mathtt{SA}$ held on the structure $(k, D(k), D(k^2), \ldots, \#)$. Then there would be m such that $P(\#k, m)$. Then by definition, there would be sets $X, Y \in D(k)$ such that $\#k = \#X$ and $m = \#Y$ and $\exists\, y \in Y \; X = Y \setminus \{y\}$. Since Hume's Principle holds on the structure $(k, D(k), D(k^2), \ldots, \#)$, we have that $\#k = \#X$ implies that there is a bijection $f : X \to Y$ that is definable in the structure k. By Proposition 68, we have that $k \setminus k$ and $k \setminus X$ are of the same cardinality, so that $X = k$. But then the condition that $y \in Y \setminus X$ implies that $y \in k \setminus k$, which is a contradiction. So, in fact, $\mathtt{SA}$ does not hold on the structure $(k, D(k), D(k^2), \ldots, \#)$. $\quad\square$

Remark 77. In the course of his proof of the existence of a model of $\mathtt{AHP}_0 + \neg\mathtt{SA}$, Linnebo noted several properties which must be had by such models (cf. [24] pp. 164-165). Since models of $\Delta_1^1 - \mathtt{HP}_0 + \neg\mathtt{SA}$ are automatically models of $\mathtt{AHP}_0 + \neg\mathtt{SA}$, Linnebo's results predict several properties of the model from the previous proposition. In this remark, we briefly explain why the properties identified by Linnebo hold on this structure. First, Linnebo notes that the example of a pseduo-number n witnessing that $\mathtt{SA}$ fails on the structure $(k, D(k), D(k^2), \ldots, \#)$ must be such that $n = \#k$. In the last paragraph of the previous proposition, we showed that $n = \#k$ was such a counterexample. Second, Linnebo notes that the example of a structure $(k, D(k), D(k^2), \ldots, \#) \models \neg\mathtt{SA}$ must be such that $k \setminus X \neq \emptyset$ implies $\#k \neq \#X$. In the context of the model constructed in the previous proposition, this is a consequence of Ax's Theorem (or Proposition 68). Finally, Linnebo notes that the example of a structure $(k, D(k), D(k^2), \ldots, \#) \models \neg\mathtt{SA}$ must contain a copy of both ω and ω^* ordered by the P-relation, that is, this structure must contain a copy of the positive integers and the negative integers ordered by the P-relation. In the model constructed in the previous theorem, this is reflected in the fact that the pseduo-numbers are precisely the integers.

4.3. Application to O-Minimal Expansions of Real-Closed Fields

Remark 78. In this section, we apply Theorem 63 to construct models of $\Sigma_1^1 - \mathtt{PH}_0$ on top of certain o-minimal expansions of real-closed fields (cf. Theorem 90), and an effectivization of this construction allows us to conclude that $\Sigma_1^1 - \mathtt{PH}_0 <_I \mathtt{ACA}_0$ (cf. Corollary 92), thus filling in a key piece of the interpretability relation (cf. Figure 2). Prior to doing this, we recall some basic notions pertaining to the model theory of o-minimal expansions of real-closed fields, such as dimension and Euler characteristic; the reader who is already familiar with these notions may wish to proceed directly to Theorem 90.

Definition 79. Suppose that L is a signature extending the signature of linear orders, and suppose that M is an L-structure such that $(M, \leq)$ is a dense linear order. Then M is *o-minimal* if every definable set is a finite union of points and intervals.

Proposition 80. Every real-closed ordered field is o-minimal.

Proof. See Marker [27] Corollary 2.5 p. 11. $\quad\square$

Definition 81. Suppose that M is an o-minimal structure. If X is a definable subset of M^n, then let $C(X)$ be the set of definable continuous functions $f : X \to M$, and let

$C_\infty(X)$ be $C(X)$ plus the two constant functions $-\infty, \infty$. Further, if $f, g \in C_\infty(X)$ and $f < g$ on X, then let

$$(f, g)_X = \{(x, r) \in X \times R : f(x) < r < g(x)\} \tag{115}$$

Then inductively define the notion of a σ-cell, where $\sigma \in 2^{<\omega}$ is a finite sequence of zeros and ones. First, 0-cells are points and 1-cells are open intervals, including $(-\infty, a)$, (a, ∞). Second, given a σ-cell X, the $\sigma 0$-cells are graphs of functions $f \in C(X)$, and the $\sigma 1$-cells are sets $(f, g)_X$ where $f, g \in C_\infty(X)$.

Definition 82. Suppose that M is an o-minimal structure. A *decomposition* of M^n is defined inductively as follows. A decomposition of M^1 is a finite partition of M with the following form:

$$\{(-\infty, a_1), (a_1, a_2), \ldots, (a_k, +\infty), \{a_1\}, \ldots, \{a_k\}\} \tag{116}$$

where $a_1 < a_2 < \cdots < a_k$. A decomposition of $M^{m+1} = M^m \times M$ is a finite partition of M^{m+1} into cells $\{A_1, \ldots, A_n\}$ such that the set of projections $\{\pi(A_1), \ldots, \pi(A_n)\}$ is a decomposition of M^m, where $\pi : M^{m+1} \to M^m$ by $\pi(x_1, \ldots, x_{m+1}) = (x_1, \ldots, x_m)$. A decomposition of M^m is said to *partition* a definable set $X \subseteq M^m$ if X can be written as a finite union of cells in the decomposition.

Theorem 83. (Cell Decomposition Theorem) Suppose that M is an o-minimal structure. For any finite sequence of B-definable sets $A_1, \ldots, A_k \subseteq M^m$, there is a decomposition of M^m partitioning each of the A_i. Moreover, the cells in the decomposition are B-definable.

Proof. See van den Dries [39] Theorem 2.11 p. 52. □

Definition 84. Suppose that M is an o-minimal structure and that $X \subseteq M^n$. Then define

$$\dim(X) = \max\{i_1 + \cdots + i_n : X \text{ contains a } (i_1, \ldots, i_n)\text{-cell}\} \tag{117}$$

$$E(X) = k_0 - k_1 + k_2 - \cdots = \sum_{d=0}^{n} k_d(-1)^d \tag{118}$$

where k_d is the number of d-dimensional cells contained in some cell decomposition of X.

Remark 85. Note that if $X \subseteq M$, then $\dim(X) > 0$ if and only if X contains an open interval. Note that the above definition of Euler dimension can be shown to be independent of the choice of the cell decomposition (cf. [39] Proposition 2.2 p. 70).

Proposition 86. Suppose that M is an o-minimal structure and that $\theta(\overline{x}, \overline{y})$ is a $\emptyset$-formula. Then there is a positive integer $N_\theta > 0$ such that for all $\overline{b} \in M$, it is the case that

$$\left|\dim(\theta(\cdot, \overline{b})\right|, \left|E(\theta(\cdot, \overline{b}))\right| < N_\theta \tag{119}$$

Further, for each integer k, it is the case that the sets

$$\{\overline{b} \in M : \dim(\theta(\cdot, \overline{b})) = k\} \quad \& \quad \{\overline{b} \in M : E(\theta(\cdot, \overline{b})) = k\} \tag{120}$$

are $\emptyset$-definable. Moreover, the formulas that define these sets and the positive integer N_θ can be uniformly computed from θ.

Proof. See van den Dries [39] Proposition 1.5 p. 65 and Proposition 2.10 p. 72. □

Proposition 87. Suppose that M is an o-minimal expansion of a real-closed field, and suppose that $X \subseteq M^n$ and $Y \subseteq M^m$ are definable sets. Then there is a definable bijection $f : X \to Y$ if and only if $\dim(X) = \dim(Y)$ and $E(X) = E(Y)$.

Proof. See van den Dries [39] p. 132. □

Remark 88. As a simple illustration of this fact, consider the example of the two sets

$$X = (-2, -1) \sqcup \{0\} \sqcup (1, 2) \qquad Y = (-1, 1) \tag{121}$$

Both have dimension 1, since they both contain intervals, and their Euler characteristics are the same, namely, $E(X) = 1 - 2 = -1$ and $E(Y) = 0 - 1 = -1$. Hence, the above proposition predicts that there is a definable bijection $f : X \to Y$, and in fact this is the case: one simply sends $(-2, -1)$ to $(-1, 0)$ and one sends 0 to 0 and one sends $(1, 2)$ to $(0, 1)$.

Proposition 89. O-minimal expansions of real closed fields have definable skolem functions.

Proof. See van den Dries [39] p. 94. □

Theorem 90. Suppose that k is a recursively-saturated o-minimal expansion of a real-closed field. Then there is a computably uniformly definable function $\# : D(k) \to k$ such that $(k, D(k), D(k^2), \ldots, \#)$ is a model of $\Sigma_1^1 - \mathtt{PH}_0 + \neg \Pi_1^1 - \mathtt{HP}_0$. Further, there is no function $\partial : D(k) \to k$ such that $(k, D(k), D(k^2), \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$.

Proof. Since k is a field of characteristic zero, the prime field of k is $\mathbb{Q}$ and the integers $\mathbb{Z}$ are hence embedded into k via $\mathbb{Q}$. Choose a recursive bijection $\langle \cdot, \cdot \rangle : \mathbb{Z}^2 \to \mathbb{Z}$. Using this embedding and this bijection, define $\# : D(k) \to k$ by $\#X = \langle \dim(X), E(X) \rangle$. Then by Proposition 87, the structure $(k, D(k), D(k^2), \ldots, \#)$ is a model of Hume's Principle. To apply Theorem 63 (iii)-(iv), we need to show that $\# : D(k) \to k$ is computably uniformly definable. So suppose that $\theta(x, \overline{y})$ is an $\emptyset$-formula with non-empty set $\overline{y}$ of parameter variables. Then by Proposition 86, from the formula $\theta(x, \overline{y})$ we can uniformly compute a positive integer $N_\theta > 0$ such that

$$k \models \forall \overline{b} \left[\left| \dim(\theta(\cdot, \overline{b})) \right|, \left| E(\theta(\cdot, \overline{b})) \right| < N_\theta \right] \tag{122}$$

as well as $\emptyset$-formulas defining the sets $\{\overline{b} : \dim(\theta(\cdot, \overline{b}) = n\}$ and $\{\overline{b} : E(\theta(\cdot, \overline{b})) = n\}$. Then for each such formula $\theta(x, \overline{y})$ we define the following $\emptyset$-formula $\theta'(x, \overline{y})$ as follows:

$$\theta'(x, \overline{y}) \equiv \bigvee_{i=0}^{N_\theta} \bigvee_{j=0}^{N_\theta} [\dim(\theta(\cdot, \overline{y})) = i \ \& \ E(\theta(\cdot, \overline{y})) = j] \to x = \langle i, j \rangle \tag{123}$$

Hence, by definition, we have that for any $\overline{a}$

$$\{\#(\theta(\cdot, \overline{a}))\} = \{c : k \models \theta'(c, \overline{a})\} \tag{124}$$

Hence, by Theorem 63 (iii)-(iv) and Proposition 89, we have that $(k, D(k), D(k^2), \ldots, \#)$ is a model of $\Sigma_1^1 - \mathtt{PH}_0$. Further, since the set $\mathrm{rng}(\#) = \mathbb{Z}$ is definable by a Σ_1^1-formula

41

in the structure $(k, D(k), D(k^2), \ldots, \#)$ but is *not* definable in k since k is o-minimal, we have that $(k, D(k), D(k^2), \ldots, \#)$ is a model of $\neg \Pi_1^1 - \mathtt{HP}_0$.

Now let us note why there is *no* function $\partial : D(k) \to k$ such that $(k, D(k), D(k^2), \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$. If there was such a function, then by Proposition 27, there would be a function $s : k^2 \to k$ whose graph was in $D(k^2)$ and which satisfied $s(x, y) = \partial(\{x, y\})$. Consider the definable set $X = \{(x, y) \in k^2 : x < y\}$, and note that $\dim(X) = 2$. Then $s \restriction X : X \to k$ is an injection. For, suppose $s(x, y) = s(x', y')$ for $(x, y), (x', y') \in X$. Then $\partial(\{x, y\}) = \partial(\{x', y'\})$ and $x < y$ and $x' < y'$. Then by Basic Law V, $\{x, y\} = \{x', y'\}$ and $x < y$ and $x' < y'$. Then $x = x'$ and $y = y'$. Hence, in fact, $s \restriction X : X \to k$ is an injection. Then trivially $s \restriction X : X \to \mathrm{rng}(s \restriction X)$ is a bijection whose graph is in $D(k^2)$. Then by the left-to-right direction of Proposition 87, it would follow that

$$2 = \dim(X) = \dim(\mathrm{rng}(s \restriction X)) \leq \dim(k) = 1 \tag{125}$$

which is a contradiction. $\qquad\square$

Remark 91. It is our claim that all of the results quoted and proved in this subsection can be proven in $\mathtt{ACA}_0$ for o-minimal structures M with $\mathtt{ACA}_0$-provable quantifier-elimination, such as real-closed fields (cf. Marker [27] Theorem 2.3 p. 10, Simpson [35] Lemma II.9.6 p. 98). The reason for this is that (i) the proofs from van den Dries [39] all concern properties of definable sets, as opposed to properties of the defining formula, and (ii) the proofs from van den Dries [39] all proceed by induction on the cartesian power of the definable set. It is worthwhile to say a little bit more about each of these points.

In regard to (i), the proofs in this section from van den Dries [39] are all concerned with properties of a definable set X, so that the definable set X has the property regardless of which particular formula is used to define X. For instance, the property of X's being a cell has this feature, since a definable set $X \subseteq M$ is e.g. an interval or a point regardless of whether the formula φ or the formula ψ is being used to define it (where φ and ψ are two formulas that do in fact define X). By the same token, the proofs in this section from van den Dries [39] are *not* concerned with the syntactic complexity of given formulas, for instance, whether or not they are Π_3^0-formulas or Π_4^0-formulas. Hence, if M has quantifier-elimination, then for the purposes of the proofs in this section from van den Dries [39], we can take the quantifier-free formulas as representatives for the definable sets. For instance, in proving the Cell Decomposition Theorem in this manner, we would in fact prove that e.g. for every finite sequence of quantifier-free formulas $\varphi_1(\overline{x}), \ldots, \varphi_k(\overline{x})$ in m-free variables, there is a quantifier-free decomposition of M^m partitioning each of the $\varphi_i(\overline{x})$.

In regard to (ii), the proofs in this section from van den Dries [39] all proceed by induction, where it is first shown that the definable subsets of M have a given property, and then it is shown that if the definable subsets of M^n have a given property, then the definable subsets of M^{n+1} have this given property. Given our discussion in the previous paragraph, when proving these theorems in $\mathtt{ACA}_0$, we would in fact prove that the quantifier-free formulas $\varphi(x)$ have a given property, and that if the quantifier-free formulas $\varphi(x_1, \ldots, x_n)$ have a given property, then the quantifier-free formulas $\varphi(x_1, \ldots, x_{n+1})$ have a given property. Since $\mathtt{ACA}_0$ has the mathematical induction axiom for all sets X, it suffices to note that $\mathtt{ACA}_0$ has enough comprehension to show that the sets X on which it is doing induction exist. Here it suffices to note that the proofs in this section from van den

Dries [39] all concern properties of the definable sets that can be expressed by (iii) finitely many quantifiers over quantifier-free definable sets and by (iv) finitely many quantifiers over the structure M. For instance, to reiterate the point made in the last paragraph, in proving the Cell Decomposition Theorem in this fashion, we must show that for every m and every finite sequence of quantifier-free formulas $\varphi_1(\overline{x}), \ldots, \varphi_k(\overline{x})$ in m-free variables, there is a quantifier-free decomposition of M^m partitioning each of the $\varphi_i(\overline{x})$. In terms of (iii), this involves a universal quantifier over quantifier-free definable sets followed by an existential quantifier over quantifier-free definable sets. In terms of (iv), this involves a universal quantifier to say that e.g. the cells in the decomposition are disjoint and another universal quantifier to say that e.g. $\varphi_i(\overline{x})$ can be written as a finite union of pairwise disjoint cells in the decomposition. Since the number of quantifiers in (iii) and (iv) is fixed in advance, $\mathtt{ACA}_0$ can prove that the set on which one is doing induction exists. In this way, the proofs from van den Dries [39] can be translated word-for-word into proofs in $\mathtt{ACA}_0$ for o-minimal structures M which have $\mathtt{ACA}_0$-provable quantifier-elimination, such as real-closed fields.

Corollary 92. $\Sigma_1^1 - \mathtt{PH}_0 <_{\mathrm{I}} \mathtt{ACA}_0$.

Proof. This follows from Proposition 7, the fact that $\mathtt{ACA}_0$ proves the existence of recursively saturated elementary extensions (cf. Simpson [35] Lemma IX.4.2 pp. 379), and the fact that the proof of the previous theorem can be formalized in $\mathtt{ACA}_0$ for o-minimal expansions of real-closed fields with $\mathtt{ACA}_0$-provable quantifier-elimination, such as real-closed fields. $\square$

4.4. Application to Separably Closed Fields of Finite Imperfection Degree

Remark 93. In the two previous subsections, we applied Theorem 63 to construct models of $\Delta_1^1 - \mathtt{HP}_0$ on top of various fields, such as certain algebraically closed fields and o-minimal expansions of real-closed fields. We noted in both Theorem 70 and Theorem 90 that this construction *cannot* result in models of $\Delta_1^1 - \mathtt{BL}_0$. Hence, this raises the question of whether there is some natural field such that one can apply Theorem 63 to it to obtain models of $\Delta_1^1 - \mathtt{BL}_0$. In this section, we isolate certain model-theoretic conditions on a field (such a uniform elimination of imaginaries) which suffice to ensure that such a construction can succeed (cf. Theorem 99). Then we note that separably closed fields of finite imperfection degree satisfy these model-theoretic conditions (cf. Theorem 101).

Definition 94. Suppose that M is an L-structure. Then M has *uniform elimination of imaginaries* if for every $\emptyset$-definable equivalence relation E on M^n there is an $\emptyset$-definable function $f : M^n \to M^m$ for some $m > 0$ such that

$$\overline{z}E\overline{y} \iff f(\overline{z}) = f(\overline{y}) \tag{126}$$

Definition 95. Suppose that M is an L-structure. Then M has a $\emptyset$-*definable pairing function* if there is an $\emptyset$-definable injection $\iota : M^2 \to M$.

Proposition 96. Suppose that M has uniform elimination of imaginaries and an $\emptyset$-definable pairing function. Then for every $\emptyset$-definable equivalence relation E on M^n there is an $\emptyset$-definable function $f : M^n \to M$ such that

$$\overline{z}E\overline{y} \iff f(\overline{z}) = f(\overline{y}) \tag{127}$$

43

Proof. By hypothesis, M has an $\emptyset$-definable pairing function $\iota : M^2 \to M$. Then define injections $j_n : M^n \to M$ recursively as follows:

$$j_1(x_1) = x_1 \tag{128}$$
$$j_2(x_1, x_2) = \iota(x_1, x_2) \tag{129}$$
$$j_{n+1}(x_1, \ldots x_n, x_{n+1}) = \iota(j_n(x_1, \ldots, x_n), x_{n+1}) \tag{130}$$

Finally, given a function $f : M^n \to M^m$ for some $m > 0$ which witnesses the uniform elimination of imaginaries, simply define $f^* = j_m \circ f$. $\qquad\square$

Proposition 97. Suppose that M has an $\emptyset$-definable pairing function and that $\mathrm{dcl}(\emptyset)$ has at least two elements. Then there is a uniformly computable sequence of injections $\iota_n : M \to M$ such that $n \neq m$ implies $\mathrm{rng}(\iota_n) \cap \mathrm{rng}(\iota_m) = \emptyset$.

Proof. Suppose that $\iota : M^2 \to M$ is the $\emptyset$-definable injection and that $b, c \in \mathrm{dcl}(\emptyset)$ are distinct. Then define injections $\iota_n : M \to M$ recursively as follows:

$$\iota_0(x) = \iota(c, \iota(c, x)) \tag{131}$$
$$\iota_{2s+1}(x) = \iota(b, \iota_{2s}(x)) \tag{132}$$
$$\iota_{2s+2}(x) = \iota(c, \iota_{2s+1}(x)) \tag{133}$$

By construction, all the functions $\iota_n : M \to M$ are injections. So it remains to show by induction on $m \leq n$ that $\mathrm{rng}(\iota_n) \cap \mathrm{rng}(\iota_m) = \emptyset$ when $m \neq n$. Clearly this holds for $n = 0$. So suppose it holds for n. If n is even then $n = 2s$ and $n + 1 = 2s + 1$. Suppose that $m < n + 1$ is such that $\mathrm{rng}(\iota_{n+1}) \cap \mathrm{rng}(\iota_m) \neq \emptyset$. Then there are x, y such that $\iota_{n+1}(x) = \iota_m(y)$. Expanding this equation on the left, we have $\iota(b, \iota_{2s}(x)) = \iota_{2s+1}(x) = \iota_{n+1}(x) = \iota_m(y)$. Then by construction, $\iota_m(y) = \iota(b, \iota_{2t}(y))$ for some $2t + 1 = m$. Then $\iota_{m-1}(y) = \iota_{2t}(y) = \iota_{2s}(x) = \iota_n(x)$, which contradicts our induction hypothesis on n. On the other hand, if n is odd then $n = 2s + 1$ and $n + 1 = 2s + 2$. Suppose that $m < n + 1$ is such that $\mathrm{rng}(\iota_{n+1}) \cap \mathrm{rng}(\iota_m) \neq \emptyset$. Then there are x, y such that $\iota_{n+1}(x) = \iota_m(y)$. Then expanding this equation on the left we have $\iota(c, \iota_{2s+1}(x)) = \iota_{2s+2}(x) = \iota_{n+1}(x) = \iota_m(y)$. There are then two cases. First suppose that $m = 0$. Then by construction $\iota_m(y) = \iota(c, \iota(c, y))$. Then $\iota(b, \iota_{2s}(x)) = \iota_{2s+1}(x) = \iota(c, y)$, and so $b = c$, which is a contradiction. Second, suppose that $m > 0$. Then by construction, $\iota_m(y) = \iota(c, \iota_{2t+1}(y))$ for some $2t + 2 = m$. Then $\iota_{m-1}(y) = \iota_{2t+1}(y) = \iota_{2s+1}(x) = \iota_n(x)$, contradicting our induction hypothesis on n. $\qquad\square$

Remark 98. Prior to proving the following theorem, let us here attempt to sketch the intuitive proof idea. Suppose that M has uniform elimination of imaginaries and a $\emptyset$-definable pairing function. Then given a formula $\theta(x, \overline{y})$ with a set of parameter variables $\overline{y}$ of length $\ell > 0$, these assumptions yield an $\emptyset$-definable function $\partial_\theta : M^\ell \to M$ such that

$$M \models [\forall\, x\, \theta(x, \overline{a}) \leftrightarrow \theta(x, \overline{b})] \iff \partial_\theta(\overline{a}) = \partial_\theta(\overline{b}) \tag{134}$$

Intuitively, the idea is to build a model $(M, D(M), D(M^2), \ldots, \partial)$ of Basic Law V by setting

$$\partial(\theta(\cdot, \overline{a})) = \partial_\theta(\overline{a}) \tag{135}$$

However, there are two potential problems. First, such a function will not be well-defined, since a given set $X \in D(M)$ will be defined by many formulas $\theta_1(\cdot, \overline{a}), \theta_2(\cdot, \overline{b}), \ldots$. Second, it is not obvious that such a function will be injective, which is required by Basic Law V. Overcoming these problems is the only thing that makes the below proof non-trivial. In particular, the first problem is overcome simply by fixing beforehand an enumeration of the all potential defining formulas $\theta_1(x, \overline{y}), \ldots, \theta_n(x, \overline{y}), \ldots$, and then defining $\partial(X)$ to be $\partial_{\theta_n}(\overline{a})$ for the first $\theta_n(x, \overline{a})$ in the enumeration that defines X for some tuple $\overline{a}$. The second problem is overcome by including additional hypotheses on M which ensure that we can partition $M = \bigsqcup_n M_n$ and likewise ensure that $\partial_{\theta_n}(\overline{a})$ always takes values in M_n. The previous proposition was in effect devoted to explaining why the hypothesis of a $\emptyset$-definable pairing function with $|\mathrm{dcl}(\emptyset)| > 1$ ensure that we can construct such a partition.

Theorem 99. Suppose that M is a $Th(M)$-computably saturated structure such that (i) M has uniform elimination of imaginaries, (ii) M has an $\emptyset$-definable pairing function, and (iii) $\mathrm{dcl}(\emptyset)$ has at least two elements. Then there is a $Th(M)$-computably uniformly definable function $\partial : D(M) \to M$ such that $(M, D(M), D(M^2), \ldots, \partial)$ is a model of $\Delta_1^1 - \mathsf{BL}_0$.

Proof. To apply Theorem 63 (iii)-(iv), we need to define an injection $\partial : D(M) \to M$ that is $Th(M)$-computably uniformly definable. Choose a fixed computable enumeration of the $\emptyset$-formulas $\theta(x, \overline{y})$ with non-empty set $\overline{y}$ of parameter variables of length ℓ_n as $\theta_1(x, \overline{y}), \ldots, \theta_n(x, \overline{y}), \ldots$. For each $n > 0$ and $0 < m \leq n$, consider the following $\emptyset$-definable sets $U_{n,m} \subseteq M^{\ell_n}$, where again ℓ_n is the length of the tuple $\overline{y}$ in $\theta_n(x, \overline{y})$:

$$U_{1,1} = M^{\ell_1} \tag{136}$$

$$U_{2,1} = \{\overline{a} \in M^{\ell_2} : \exists \overline{b} \in M^{\ell_1} \, [\forall x \, \theta_2(x, \overline{a}) \leftrightarrow \theta_1(x, \overline{b})]\} \tag{137}$$

$$U_{2,2} = M^{\ell_2} \setminus U_{2,1} \tag{138}$$

$$U_{3,1} = \{\overline{a} \in M^{\ell_3} : \exists \overline{b} \in M^{\ell_1} \, [\forall x \, \theta_3(x, \overline{a}) \leftrightarrow \theta_1(x, \overline{b})]\} \tag{139}$$

$$U_{3,2} = \{\overline{a} \in M^{\ell_3} : \exists \overline{b} \in M^{\ell_2} \, [\forall x \, \theta_3(x, \overline{a}) \leftrightarrow \theta_2(x, \overline{b})]\} \setminus U_{3,1} \tag{140}$$

$$U_{3,3} = M^{\ell_3} \setminus (U_{3,1} \cup U_{3,2}) \tag{141}$$

Note that for a fixed $n > 0$ that the sets $U_{n,1}, \ldots, U_{n,n}$ partition M^{ℓ_n} and that the formulas defining these sets are uniformly computable from n. Then define $\emptyset$-definable equivalence relations on M^{ℓ_n} as follows:

$$\overline{y} E_n \overline{z} \iff M \models [\forall x \, \theta_n(x, \overline{y}) \leftrightarrow \theta_n(x, \overline{z})] \tag{142}$$

Note by definition that any two elements $\overline{y}$ and $\overline{z}$ which are E_n-equivalent are in the same member of the partition $U_{n,1}, \ldots, U_{n,n}$ of M^{ℓ_n}. By Proposition 96 from $\theta_n(x, \overline{y})$ we can uniformly $Th(M)$-compute a $\emptyset$-definable function $f_n : M^{\ell_n} \to M$ such that

$$M \models [\forall x \, \theta_n(x, \overline{y}) \leftrightarrow \theta_n(x, \overline{z})] \iff \overline{y} E_n \overline{z} \iff f_n(\overline{y}) = f_n(\overline{z}) \tag{143}$$

By Proposition 97, we can uniformly compute a sequence of injections $\iota_n : M \to M$ with disjoint ranges, and we can define $g_n = \iota_n \circ f_n$. Finally, define $\partial : D(M) \to M$ by

$$\partial(\theta_n(\cdot, \overline{a})) = c \iff \bigwedge_{m=1}^{n} [\overline{a} \in U_{n,m} \to (\exists \overline{b} \in M^{\ell_m} \, \& \, \forall x \, \theta_n(x, \overline{a}) \leftrightarrow \theta_m(x, \overline{b}) \, \& \, c = g_m(\overline{b}))] \tag{144}$$

First let us show that $\partial : D(M) \to M$ is a well-defined function. So suppose that $\theta_n(\cdot, \overline{a})$ and c satisfy the right-hand side of equation (144) and that $\theta_{n'}(\cdot, \overline{a}')$ and c' also satisfy the right-hand side of equation (144), and suppose that $\theta_n(\cdot, \overline{a})$ and $\theta_{n'}(\cdot, \overline{a}')$ define the same set. Then we must show that $c = c'$. Without loss of generality, $n' \leq n$. If $n' = n$, then since $\theta_n(\cdot, \overline{a})$ and $\theta_{n'}(\cdot, \overline{a}')$ define the same set, we have that $\overline{a}$ and $\overline{a}'$ are E_n-equivalent and hence are in the same set $U_{n,m}$. Then by the right-hand side of equation (144), we have that there are $\overline{b}, \overline{b}' \in M^{\ell_m}$ such that

$$M \models \forall x \; \theta_m(x, \overline{b}) \leftrightarrow \theta_n(x, \overline{a}) \leftrightarrow \theta_n(x, \overline{a}') \leftrightarrow \theta_m(x, \overline{b}') \tag{145}$$

$$c = g_m(\overline{b}) \tag{146}$$

$$c' = g_m(\overline{b}') \tag{147}$$

But by equation (145), we have that $\overline{b}$ and $\overline{b}'$ are E_m-equivalent, and hence by equation (143), we have that $f_m(\overline{b}) = f_m(\overline{b}')$ and so by equations (146)-(147) we have that

$$c = g_m(\overline{b}) = \iota_m \circ f_m(\overline{b}) = \iota_m \circ f_m(\overline{b}') = g_m(\overline{b}') = c' \tag{148}$$

In the case where $n' < n$, we have that $\overline{a} \in U_{n,m}$ and $\overline{a}' \in U_{n',m'}$ and so by the right-hand side of equation (144), we have that there is $\overline{b} \in M^{\ell_m}, \overline{b}' \in M^{\ell_{m'}}$ such that

$$M \models \forall x \; \theta_m(x, \overline{b}) \leftrightarrow \theta_n(x, \overline{a}) \leftrightarrow \theta_n(x, \overline{a}') \leftrightarrow \theta_{m'}(x, \overline{b}') \tag{149}$$

$$c = g_m(\overline{b}) \tag{150}$$

$$c' = g_{m'}(\overline{b}') \tag{151}$$

Then by equation (149) and the definition of the sets $U_{n,m}$, we must have that $m = m'$. Then by equation (149) again, we have that $\overline{b}$ and $\overline{b}'$ are E_m-equivalent, and, hence, by equation (143), we have that $f_m(\overline{b}) = f_m(\overline{b}')$, and so by equations (150)-(151) we have that

$$c = g_m(\overline{b}) = \iota_m \circ f_m(\overline{b}) = \iota_m \circ f_m(\overline{b}') = g_m(\overline{b}') = c' \tag{152}$$

Therefore, $\partial : D(M) \to M$ is a well-defined function.

Now let us show that $\partial : D(M) \to M$ is an injection. Suppose that $\theta_n(\cdot, \overline{a})$ and c satisfy the right-hand side of equation (144) and that $\theta_{n'}(\cdot, \overline{a}')$ and c' satisfy the right-hand side of equation (144) and suppose that $c = c'$. Then we must show that $\theta_n(\cdot, \overline{a})$ and $\theta_{n'}(\cdot, \overline{a}')$ define the same set. We have that $\overline{a} \in U_{n,m}$ and $\overline{a}' \in U_{n',m'}$, and by the right-hand side of equation (144), we have that there is $\overline{b} \in M^{\ell_m}, \overline{b}' \in M^{\ell_{m'}}$ such that

$$M \models \forall x \; \theta_n(x, \overline{a}) \leftrightarrow \theta_m(x, \overline{b}) \tag{153}$$

$$M \models \forall x \; \theta_{n'}(x, \overline{a}') \leftrightarrow \theta_{m'}(x, \overline{b}') \tag{154}$$

$$g_m(\overline{b}) = c = c' = g_{m'}(\overline{b}') \tag{155}$$

Since $g_m = \iota_m \circ f_m$ and since the functions ι_m have distinct ranges, equation (155) implies that $m = m'$ and since $g_m = \iota_m \circ f_m$ and ι_m is an injection, we have that equation (155) implies that $f_m(\overline{b}) = f_{m'}(\overline{b}')$, which by equation (143) implies that $\theta_m(\cdot, \overline{b})$ and $\theta_{m'}(\cdot, \overline{b}')$ define the same set. This in turn implies with equations (153)-(154) that

$\theta_n(\cdot, \overline{a})$ and $\theta_{n'}(\cdot, \overline{a}')$ define the same set, which is what we wanted to show. Hence, in fact $\partial : D(M) \to M$ is an injection.

So, $\partial : D(M) \to M$ is well-defined and indeed an injection. Note that by its very definition in equation (144), we have that $\partial : D(M) \to M$ is $Th(M)$-computably uniformly definable. Hence, by Theorem 63 (iii)-(iv), we have that $(M, D(M), D(M^2), \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$. $\qquad\square$

Definition 100. Suppose that k is field of characteristic $p > 0$. Then k is a *separably closed field of finite imperfection degree* if (i) there is a finite set $B \subseteq k$ such that the set of monomials $\{b_1^{m_1} \cdots b_e^{m_e} : 0 \leq m_i < p \,\&\, b_1, \ldots, b_e \in B\}$ is a basis for k over k^p, and if (ii) every $f \in k[x]$ such that $f' \neq 0$ has a root in k.

Theorem 101. Suppose that k is a recursively saturated separably closed field of finite imperfection degree. Then there is a computably uniformly definable function $\partial : D(k) \to k$ such that $(k, D(k), D(k^2), \ldots, \partial)$ is a model of $\Delta_1^1 - \mathtt{BL}_0$.

Proof. This follows immediately from the fact that such fields satisfy the antecedents of the previous theorem and have a computable theory when names are added for the finite set B from the previous definition (cf. Messmer [28] Proposition 4.2 p. 140, p. 143, Remark 4.4 p. 141). $\qquad\square$

Remark 102. If we knew that all the elements of the proof of the previous theorem were formalizable in $\mathtt{ACA}_0$, then we could infer from the proof of the above theorem and Proposition 7 that we have $\Delta_1^1 - \mathtt{BL}_0 <_I \mathtt{ACA}_0$. It is clear from the proof that this comes down to determining whether or not the uniform elimination of imaginaries for separably closed fields of finite imperfection degree is provable in $\mathtt{ACA}_0$.

5. Further Questions

Question 103. In Figure 1, we summarized what is known about the provability relation. Two questions which remain open are the following: does $\Delta_1^1 - \mathtt{BL}_0$ imply $\Sigma_1^1 - \mathtt{LB}_0$ and does $\Pi_1^1 - \mathtt{HP}_0$ imply $\Sigma_1^1 - \mathtt{PH}_0$?

Question 104. In Remark 71, we noted that if Ax's Theorem 65 is provable in $\mathtt{ACA}_0$, then we would have another proof of $\Delta_1^1 - \mathtt{HP}_0 <_I \mathtt{ACA}_0$ besides the proof from Corollary 92. Hence, is Ax's Theorem 65 provable in $\mathtt{ACA}_0$?

Question 105. In Remark 102, we noted that if the uniform elimination of imaginaries for separably closed fields is provable in $\mathtt{ACA}_0$, then we would have $\Delta_1^1 - \mathtt{BL}_0 <_I \mathtt{ACA}_0$. Hence, is the uniform elimination of imaginaries for separably closed fields provable in $\mathtt{ACA}_0$?

Question 106. The results in Heck [19], Ganea [16], and Visser [40] imply that $\mathtt{ABL}_0$ is mutually interpretable with Robinson's $\mathtt{Q}$. Is $\Delta_1^1 - \mathtt{BL}_0$ mutually interpretable with Robinson's Q?

Question 107. What is the exact interpretability strength of $\mathtt{AHP}_0$ and $\Delta_1^1 - \mathtt{HP}_0$? Are these theories interpretable in Robinson's Q?

Question 108. In § 2.2, and in particular around equation (20), we pointed out that there is no function symbol in our language for the mapping $(R, n) \mapsto \#(R_n)$, where R is a binary relation and $R_n = \{m : Rnm\}$. The inclusion of such a function symbol will not affect systems which contain the Δ_1^1-comprehension schema, since the graph of this function is Δ_1^1-definable (cf. equation (20)). However, in Propositions 48-49, we pointed that $\mathtt{AHP}_0$ and $\mathtt{ABL}_0$ do not prove the existence of the graph of this function $(R, n) \mapsto \#(R_n)$, in the sense that $\mathtt{AHP}_0$ and $\mathtt{ABL}_0$ do not prove that the binary relation $\{(n, m) : \#(R_n) = m\}$ exists for every binary relation R. Does the addition of this function symbol affect the interpretability strength of $\mathtt{AHP}_0$ and $\mathtt{ABL}_0$? In particular, do the Heck-Visser-Ganea results about the mutual interpretability of $\mathtt{ABL}_0$ and Robinson's Q mentioned in § 1.5 still hold if we add a function symbol for $(R, n) \mapsto \#(R_n)$?

6. Acknowledgements

[1] James Ax. The Elementary Theory of Finite Fields. *Annals of Mathematics*, 88:239–271, 1968.

[2] Jon Barwise and John Schlipf. On Recursively Saturated Models of Arithmetic. In A. Dold and B. Eckmann, editors, *Model Theory and Algebra*, volume 498 of *Lecture Notes in Mathematics*, pages 42–55. Springer, Berlin, 1975.

[3] George Boolos. The Consistency of Frege's Foundations of Arithmetic. In Judith Jarvis Thomson, editor, *On Being and Saying: Essays in Honor of Richard Cartwright*, pages 3–20. MIT Press, Cambridge, 1987. Reprinted in [6], [12].

[4] George Boolos. Frege's Theorem and the Peano Postulates. *The Bulletin of Symbolic Logic*, 1(3):317–326, 1995. Reprinted in [6].

[5] George Boolos. On the Proof of Frege's Theorem. In Adam Morton and Stephen P. Stich, editors, *Benacerraf and His Critics*, pages 143–159. Blackwell, 1996. Reprinted in [6].

[6] George Boolos. *Logic, Logic, and Logic.* Harvard University Press, Cambridge, MA, 1998. Edited by Richard Jeffrey.

[7] George Boolos and Richard G. Heck Jr. Die Grundlagen der Arithmetik, Sections 82-83. In Matthias Schirn, editor, *Philosophy of Mathematics Today*, pages 407–428. Clarendon Press, 1998. Reprinted in [6].

[8] John P. Burgess. *Fixing Frege.* Princeton Monographs in Philosophy. Princeton University Press, Princeton, 2005.

[9] John P. Burgess and A. P. Hazen. Predicative Logic and Formal Arithmetic. *Notre Dame Journal of Formal Logic*, 39(1):1–17, 1998.

[10] Richard Dedekind. *Was sind und was sollen die Zahlen?* Vieweg, Braunschweig, 1888. Second edition 1893. Reprinted in [11] vol. 3 pp. 335-391.

[11] Richard Dedekind. *Gesammelte mathematische Werke.* Vieweg, Braunschweig, 1930-1932. Three volumes. Edited by Robert Fricke, Emmy Noether, and Øystein Ore.

[12] William Demopoulos, editor. *Frege's Philosophy of Mathematics.* Harvard University Press, Cambridge, 1995.

[13] Fernando Ferreira and Kai F. Wehmeier. On the Consistency of the Δ_1^1-CA fragment of Frege's Grundgesetze. *Journal of Philosophical Logic*, 31(4):301–311, 2002.

[14] Gottlob Frege. *Die Grundlagen der Arithmetik.* Koebner, Breslau, 1884.

[15] R. O. Gandy. Proof of Mostowski's Conjecture. *Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques, Astronomiques et Physiques*, 8:571–575, 1960.

[16] Mihai Ganea. Burgess' *PV* is Robinson's *Q*. *Journal of Symbolic Logic*, 72(2):618–624, 2007.

[17] Petr Hájek and Pavel Pudlák. *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer, Berlin, 1998.

[18] Bob Hale and Crispin Wright. *The Reason's Proper Study*. Oxford University Press, Oxford, 2001.

[19] Richard G. Heck, Jr. The Development of Arithmetic in Frege's Grundgesetze der Arithmetik. *The Journal of Symbolic Logic*, 58(2):579–601, 1993.

[20] Richard G. Heck, Jr. The Consistency of Predicative Fragments of Frege's Grundgesetze der Arithmetik. *History and Philosophy of Logic*, 17(4):209–220, 1996.

[21] Richard G. Heck, Jr. Cardinality, Counting, and Equinumerosity. *Notre Dame Journal of Formal Logic*, 41(3):187–209, 2000.

[22] Stephen Cole Kleene. Quantification of Number-Theoretic Functions. *Compositio Mathematica*, 14:23–40, 1959.

[23] Per Lindström. *Aspects of Incompleteness*, volume 10 of *Lecture Notes in Logic*. Association for Symbolic Logic, Urbana, IL, second edition, 2003.

[24] Øystein Linnebo. Predicative Fragments of Frege Arithmetic. *The Bulletin of Symbolic Logic*, 10(2):153–174, 2004.

[25] Fraser MacBride. Speaking with the Shadows: A Study of Neo-Logicism. *The British Journal for the Philosophy of Science*, 54(1):103–163, 2003.

[26] David Marker. *Model Theory*, volume 217 of *Graduate Texts in Mathematics*. Springer, New York, 2002.

[27] David Marker. Introduction to the Model Theory of Fields. In *Model Theory of Fields*, volume 5 of *Lecture Notes in Logic*, pages 1–37. Association for Symbolic Logic, La Jolla, CA, second edition, 2006.

[28] Margit Messmer. Some Model Theory of Separably Closed Fields. In *Model Theory of Fields*, volume 5 of *Lecture Notes in Logic*, pages 135–152. Association for Symbolic Logic, La Jolla, CA, second edition, 2006.

[29] Yiannis N. Moschovakis. *Descriptive Set Theory*, volume 100 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1980.

[30] Piergiorgio Odifreddi. *Classical Recursion Theory. Vol. I*, volume 125 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1989.

[31] Charles Parsons. *Mathematical Thought and Its Objects*. Harvard University Press, Cambridge, 2008.

[32] Bruno Poizat. *Stable Groups*, volume 87 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[33] Gerald E. Sacks. *Higher Recursion Theory*. Perspectives in Mathematical Logic. Springer, Berlin, 1990.

[34] Stewart Shapiro. *Foundations without Foundationalism: A Case for Second-Order Logic*, volume 17 of *Oxford Logic Guides*. The Clarendon Press, New York, 1991.

[35] Stephen G. Simpson. *Subsystems of Second Order Arithmetic*. Cambridge University Press, Cambridge, second edition, 2009.

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

[37] Clifford Spector. Hyperarithmetical Quantifiers. *Fundamenta Mathematicae*, 48:313–320, 1959/1960.

[38] John R. Steel. Forcing with Tagged Trees. *Annals of Mathematical Logic*, 15(1):55–74, 1978.

[39] Lou van den Dries. *Tame Topology and O-Minimal Structures*, volume 248 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1998.

[40] Albert Visser. The Predicative Frege Hierarchy. *Annals of Pure and Applied Logic*, 160(2):129–153, 2009.

[41] Kai F. Wehmeier. Consistent Fragments of Grundgesetze and the Existence of Non-Logical Objects. *Synthese*, 121(3):309–328, 1999.

[42] Kai F. Wehmeier. Russell's Paradox in Consistent Fragments of Frege's Grundgesetze der Arithmetik. In *One Hundred Years of Russell's Paradox*, volume 6 of *de Gruyter Series in Logic and its Applications*, pages 247–257. de Gruyter, Berlin, 2004.

[43] Crispin Wright. *Frege's Conception of Numbers as Objects*, volume 2 of *Scots Philosophical Monographs*. Aberdeen University Press, Aberdeen, 1983.

[44] Crispin Wright. Is Hume's Principle Analytic? *Notre Dame Journal of Formal Logic*, 40(1):6–30, 1999. Reprinted in [18].