FROM BOUNDED ARITHMETIC TO SECOND ORDER ARITHMETIC VIA AUTOMORPHISMS

ALI ENAYAT

ABSTRACT. In this paper we examine the relationship between automorphisms of models of $I\Delta_0$ (bounded arithmetic) and strong systems of arithmetic, such as PA, ACA_0 (arithmetical comprehension schema with restricted induction), and Z_2 (second order arithmetic). For example, we establish the following characterization of PA by proving a "reversal" of a theorem of Gaifman:

Theorem. *The following are equivalent for completions T of $I\Delta_0$:*
(a) $T \vdash PA$;
(b) *Some model $\mathfrak{M} = (M, \cdots)$ of T has a proper end extension $\mathfrak{N}$ which satisfies $I\Delta_0$ and for some automorphism j of $\mathfrak{N}$, M is precisely the fixed point set of j.*
Our results also shed light on the metamathematics of the Quine-Jensen system NFU of set theory with a universal set.

1. INTRODUCTION

The classical work of Ehrenfeucht and Mostowski introduced the powerful method of indiscernibles to show that any first order theory with an infinite model has a proper class of models with rich automorphism groups [CK, Section 3.3]. In the context of models of arithmetic, the first substantial results concerning automorphisms that extend the work of Ehrenfeucht and Mostowski are to be found in Gaifman's seminal work [G] on the model theory of *Peano arithmetic* PA. Gaifman refined the MacDowell-Specker method [MS] of building elementary end extensions by introducing the machinery of *minimal types*, which can be used to produce a variety of models of PA with special properties. For example, they can be used to establish the striking result below. Here $Aut(\mathfrak{N})$ is the group of automorphisms of $\mathfrak{N}$, and $Aut(\mathfrak{N}, M)$ is the pointwise stabilizer of M (i.e., the subgroup of $Aut(\mathfrak{N})$ consisting of automorphisms of $\mathfrak{N}$ that fix every element of M).

THEOREM 1.1. (Gaifman) *Suppose $\mathfrak{M} = (M, \cdots)$ is a model of PA, and $\mathbb{L}$ is a linear order.*

(a) *There is an elementary end extension $\mathfrak{N}$ of $\mathfrak{M}$ such that $Aut(\mathfrak{N}, M) \cong Aut(\mathbb{L})$* [G, Theorem 4.11].
(b) *There is an elementary end extension $\mathfrak{N}$ of $\mathfrak{M}$ such that for some $j \in Aut(\mathfrak{N})$, M is the fixed point set of j* [G, Theorems 4.9 - 4.11].

Schmerl [Sc] has recently established a strong generalization of part (a) of Theorem 1.1 by showing that $Aut(\mathbb{L})$ can be replaced by any *closed subgroup* of $Aut(\mathbb{L})$. This shows that the class of left-orderable groups coincides with the class of groups that can occur as $Aut(\mathfrak{M})$ for models $\mathfrak{M}$ of PA. A major trend in the study of

Date: August 24, 2005.

automorphism groups of models of PA was initiated in the early 1980's with the work of Smorynski and Kotlarski (independently) on automorphisms of *countable recursively saturated models.* This has proved to be a fertile area of research, and has resulted in a number of striking results by Kaye, Kossak, Kotlarski, Lascar, and Schmerl, to name a few. The reader interested in becoming familiar with the rudiments of the subject is referred to the volume [KM].

This paper provides model theoretic characterizations of the strong systems of arithmetic PA, ACA_0, and Z_2 in terms of automorphisms of models of the weak system of arithmetic $I\Delta_0$ (commonly known as *bounded arithmetic*). Previously, Ressayre [Re] provided elegant characterizations of PA and the fragment $I\Sigma_1$ of PA in terms of *endomorphisms,* but there is no overlap between Ressayre's results and ours. For other model theoretic characterizations of PA, see [Kay].

The plan of the paper is as follows. After dealing with preliminaries in Section 2, we concentrate on the relationship between automorphisms of models of bounded arithmetic and the axiomatic systems PA and ACA_0 in Section 3. The principal results of Section 3 are Theorems A and B. Theorem A (Section 3.1) establishes a strong reversal of Theorem 1.1(b), while Theorem B (Section 3.2) is a refined form of Theorem 1.1(b) for models of ACA_0 (Theorem B is implicit in Gaifman [G], but the proof here is new). Theorems A and B together yield a model theoretic characterization of ACA_0 in terms of automorphisms. Section 4 focuses on the relationship between automorphisms of models of bounded arithmetic and models of second order arithmetic. The key notion in Section 4 is that of an "M-amenable automorphism", shown in Theorems C and D to be closely tied to models of full second order arithmetic. Section 5 includes a brief discussion of the consequences of the results in Sections 3 and 4 for the metamathematics of NFU set theory, and a discussion of further work and open questions.

The results of this paper were discovered in the context of the study of Jensen's modification NFU [Jen] of Quine's *New Foundations* system NF of set theory [Q] with a universal set. They have been used by Robert Solovay and the author to pinpoint the "arithmetical content" of certain natural extensions of NFU, such as the theory $NFUA^{-\infty}$ obtained by strengthening NFU with the axioms "every set is finite" and "every Cantorian set is strongly Cantorian". This topic will be fully treated in a forthcoming paper and we have therefore provided only a brief summary of our results for the metamathematics of NFU in Section 5.1. We should mention that there is also a *set theoretical* counterpart to the theme of this paper. This is partly explained in [E-1], in which automorphisms of models of weak systems of set theory are shown to be intimately connected to ZF-set theory with Mahlo cardinals. Roughly speaking, the results in [E-1] are the set theoretical analogues of Theorems A and B of this paper. The set theoretical analogues of Theorems D and E will appear in [E-3].

Brief history: In the early 1990's Holmes [Ho-1] made a breakthrough by using a large cardinal hypothesis (measurability) to establish the consistency of certain natural extensions ($NFUA$ and $NFUB$) of the Quine-Jensen system NFU. Holmes' work prompted Solovay[1] to work out the precise consistency strengths of $NFUA$

[1]Solovay's work on $NFUB$ appears in [Sol], but his work on $NFUA$ is unpublished. Holmes [Ho-1] contains an extension of one direction of Solovay's equiconsistency result on $NFUA$, and [E-1] contains a generalization of both directions of Solovay's equiconsistency result on $NFUA$.

and $NFUB$, by showing that (a) $NFUA$ is equiconsistent with

$$ZFC + \{\text{“there is an } n\text{-Mahlo cardinal”} : n \in \omega\},$$

and that (b) $NFUB$ is equiconsistent with

$$ZFC\backslash\{\text{Power Set}\} + \text{“there is a weakly compact cardinal”}.$$

The work of Holmes and Solovay unearthed a deep, unexpected relationship between strong set theoretical hypotheses and models of $NFUA/B$ in which the axiom of infinity holds. This inspired the author to seek a parallel relationship between strong *arithmetical* hypotheses and models of $NFUA/B$ in which the axiom of infinity *fails*. My initial result in this direction (a slightly weaker form of Theorem C) was an arithmetical analogue of a key result in [Sol]. The communication of this result to Solovay in January 2002 led to an extensive (e-mail) correspondence during the following year. It was during the course of this intense and inspiring period that I managed to obtain the results of this paper in their current form. Solovay has also established a number of results concerning the metamathematics of NFU that remain unpublished, which will hopefully appear in the near future.

Acknowledgments: I am indebted to Robert Solovay for his patience and insights offered through meticulously crafted e-mail communiqués. I also wish to thank Randall Holmes for helpful discussions about NFU; Roman Kossak and Joel Hamkins for inviting me to present my results at the CUNY Logic Workshop; Albert Visser for formulating probing questions which led me to the results in Section 3.3; Steve Simpson for alerting me to the crucial role of the dependent choice scheme in Mostowski's forcing construction [Mo-1, 2]; and Iraj Kalantari and Mojtaba Moniri for unfailing camaraderie. I am also grateful to Andreas Blass and the anonymous referees for detailed constructive comments on earlier drafts of this paper.

2. PRELIMINARIES

2.1. Bounded Arithmetic

- The language of first order arithmetic, $\mathcal{L}_A$, is $\{+, \cdot, Succ(x), <, 0\}$.
- Models of $\mathcal{L}_A$ are of the form $\mathfrak{M} = (M, +^{\mathfrak{M}}, \cdots), \mathfrak{N} = (N, +^{\mathfrak{N}}, \cdots)$, etc. For models $\mathfrak{M}$ and $\mathfrak{N}$ of $\mathcal{L}_A$, we say that $\mathfrak{N}$ *end extends* $\mathfrak{M}$ (equivalently: $\mathfrak{M}$ is an *initial* submodel of $\mathfrak{N}$), written $\mathfrak{M} \subseteq_e \mathfrak{N}$, if $\mathfrak{M}$ is a submodel of $\mathfrak{N}$ and $a < b$ for every $a \in M$, and $b \in N\backslash M$. We abbreviate the phrase “elementary end extension” by “e.e.e.”.
- I is a *cut* of $\mathfrak{M}$, where $\mathfrak{M}$ is a model of Robinson's Q, if I is a *proper* initial segment of $\mathfrak{M}$ with no last element.
- A first order $\mathcal{L}_A$-formula φ is said to be a Δ_0-*formula* if all the quantifiers of φ are bounded, i.e., they are of the form $\exists x \leq y$, or of the form $\forall x \leq y$, where x and y are (meta)variables. Δ_0-formulae are also known as *bounded* formulae.
- *Bounded arithmetic*, or $I\Delta_0$, is the fragment of Peano arithmetic with the induction scheme limited to Δ_0-formulae. More specifically, it is a theory formulated in the language $\mathcal{L}_A$, and is obtained by adding the scheme of induction for Δ_0-formulae to Robinson's arithmetic Q. The metamathematical study of bounded arithmetic has close ties with the subject of computational complexity. See [HP] or [Kr] for thorough introductions.

- Bennett [Be] showed that the graph of the exponential function $y = 2^x$ can be defined by a Δ_0-predicate in the standard model of arithmetic. Later, Paris found another Δ_0-predicate $\varphi(x, y)$ which does the job, and $I\Delta_0$ can prove the familiar algebraic laws about exponentiation for $\varphi(x, y)$ [DG, Appendix][2]. By a classical theorem of Parikh [Pa] however, $I\Delta_0$ can only prove the totality of functions with a polynomial growth rate, hence

$$I\Delta_0 \nvdash \forall x \exists y \varphi(x, y).$$

  It is now known that the graphs of many other fast growing recursive functions, such as the superexponential function $Superexp$[3], the Ackerman function, and indeed all functions $\{F_\alpha : \alpha < \varepsilon_0\}$ in the (fast growing) Wainer hierarchy, can be defined by Δ_0-predicates for which $I\Delta_0$ can prove appropriate recursion schemes. This remarkable discovery is due to Sommer [Som-1], [Som-2], but the reader is also referred to D'Aquino's paper [D] for a perspicuous Δ_0-treatment of the superexponential function and the Ackerman function.

The following result is well known: a routine proof by contradiction proves (a), with Δ_0 induction applied to $\varphi^*(v) := \forall x \leq v \, \neg\varphi(v)$, (b) follows from (a) since the maximum of S_φ is the least upper bound of S_φ, and (c) follows from (b).

LEMMA 2.1. *Suppose $\mathfrak{M}$ is a model of $I\Delta_0$, and let S_φ be the solution set of some Δ_0-formula $\varphi(x, \overrightarrow{a})$ in $\mathfrak{M}$, where $\overrightarrow{a}$ is a sequence of parameters from M. If $S_\varphi \neq \emptyset$, then:*

(a) [Δ_0-MIN] *S_φ has a minimum element;*
(b) [Δ_0-MAX] *If S_φ is bounded in $\mathfrak{M}$, then S_φ has a maximum element;*
(c) [Δ_0-OVERSPILL] *If S_φ includes a cut I of $\mathfrak{M}$, then for some $b \in M \backslash I$, $[0, b]^{\mathfrak{M}} \subseteq S_\varphi$.*

2.2. The Strength of $I\Delta_0$+Exp

Let $\varphi(x, y)$ be a reasonable Δ_0-formula expressing "$2^x = y$". $I\Delta_0 + Exp$ is the extension of $I\Delta_0$ obtained by adding the axiom

$$Exp := \forall x \exists y \varphi(x, y).$$

At first sight $I\Delta_0 + Exp$ is a rather weak theory since it cannot even prove the totality of the superexponential function or any faster growing function. But, experience has shown that it is a remarkably robust theory that is able to prove a large variety of theorems of number theory and finite combinatorics[4]. One explanation for this phenomenon is offered by the fact that one can use *Ackermann coding* to simulate a workable set theory within $I\Delta_0 + Exp$. Let $E(x, y)$ be a Δ_0-predicate that expresses

[2]Independently, Pudlák [Pu-1] also provided an $I\Delta_0$-treatment of the exponential function. A detailed exposition is provided in [Bu] and [HP, Ch. V, Sec.3(c)]

[3]The superexponential function, $Superexp(n, x)$, is defined by the recursion scheme: $Superexp(0, x) = x$, $Superexp(n + 1, x) = 2^{Superexp(n,x)}$. Thus for $n > 0$, $Superexp(n, x)$ is an exponential stack of length $n+1$, where the top element is x, and the remaining n entries form a tower of 2's.

[4]Indeed, Harvey Friedman has conjectured that all "arithmetical theorems" proved in the journal *Annals of Mathematics* (such as Wiles' proof of Fermat's Last Theorem), can be implemented within $I\Delta_0 + Exp$. We refer the reader to Avigad's paper [Av] for an excellent discussion of the foundational role of $I\Delta_0 + Exp$.

"the x-th digit in the binary expansion of y is a 1". We shall henceforth refer to E as "*Ackermann's* $\in$". It is well known that $\mathfrak{M}$ is a model of PA iff (M, E) is a model of $ZF\backslash\{\text{Infinity}\} \cup \{\neg\text{Infinity}\}$, but if $\mathfrak{M}$ is a model of $I\Delta_0 + Exp$, then (M, E) is still a model of a decent fragment of $ZF\backslash\{\text{Infinity}\} \cup \{\neg\text{Infinity}\}$. More specifically:

THEOREM 2.2. (Dimitracopoulos-Gaifman ([DG], [HP, Ch.I., Sec.1(b)]). *If* $\mathfrak{M} \vDash I\Delta_0 + Exp$, *and* E *is Ackermann's* $\in$, *then* (M, E) *satisfies the following axioms:*

(1) *Extensionality;*
(2) *Pairs;*
(3) *Union;*
(4) *Powerset;*
(5) Δ_0*-Comprehension Scheme; and*
(6) *the negation of Infinity.*

- Suppose $\mathfrak{M} \vDash I\Delta_0$ and E is Ackermann's $\in$ in the sense of $\mathfrak{M}$.
  (a) For $c \in M$, $c_E := \{m \in M : mEc\}$.
  (b) $X \subseteq M$ is *coded* in $\mathfrak{M}$ if there is some $c \in M$ such that $X = c_E$.
  (c) Suppose I is a cut of $\mathfrak{M}$,

$$SSy_I(\mathfrak{M}) := \{c_E \cap I : c \in N\}.$$

  In particular, if I is the standard cut of $\mathfrak{M}$, then $SSy_I(\mathfrak{M})$ is what is known in the literature as the *standard system* of $\mathfrak{M}$.

2.3. Second Order Arithmetic

- The systems Z_2 and ACA_0 are fully discussed in Simpson's encyclopedic reference [Si-2]. Z_2 is often referred to as *second order arithmetic*[5], or as *analysis*. ACA_0 is the subsystem of Z_2 with the comprehension scheme limited to formulas with no second order quantifiers.
- Models of second order arithmetic (and its subsystems) are of the *two-sorted* form $(\mathfrak{M}, \mathcal{A})$, where $\mathfrak{M}$ is a model in the language $\mathcal{L}_A$, and $\mathcal{A}$ is a family of subsets of M. Since coding apparatus is available in the models of arithmetic $\mathfrak{M}$ considered here, we shall use expressions such as "$f \in \mathcal{A}$", where f is a function, as a substitute for the more precise but lengthier expression "the canonical code of f is in $\mathcal{A}$".
- For $\mathcal{L} \supseteq \mathcal{L}_A$, $PA(\mathcal{L})$ is PA augmented by the induction scheme for all $\mathcal{L}$-formulas. Note that if $(\mathfrak{M}, \mathcal{A}) \vDash ACA_0$, then $(\mathfrak{M}, S)_{S\in\mathcal{A}} \vDash PA(\mathcal{L})$, where $\mathcal{L}$ is the extension of $\mathcal{L}_A$ obtained by adding a unary predicate for each $S \in \mathcal{A}$.

[5]Some authors, especially those belonging to the Polish school of logic (e.g., [Mo-1]), use A_2^- for the system Z_2 (and A_2 for Z_2 plus the choice scheme).

3. AUTOMORPHISMS AND $\mathbf{ACA_0}$

The main results of this section are Theorem A and Theorem B. Theorem A establishes a strong "reversal" of Theorem 1.1(b), and Theorem B is the analogue of Theorem 1.1(b) for models of ACA_0.

3.1. $\mathbf{ACA_0}$ from Automorphisms

THEOREM A. *If $\mathfrak{N} \vDash I\Delta_0$ and j is an automorphism of $\mathfrak{N}$ such that the fixed point set M of j is a proper initial segment of $\mathfrak{N}$, then $(\mathfrak{M}, SSy_M(\mathfrak{N})) \vDash ACA_0$.*

Before presenting the proof of Theorem A, let us point out an important corollary obtained by coupling Theorem A with Theorem 1.1(b):

COROLLARY 3.1 *The following are equivalent for completions T of $I\Delta_0$:*

(a) $T \vdash PA$;
(b) *Every model $\mathfrak{M}$ of T has a proper e.e.e. $\mathfrak{N}$ such that for some automorphism j of $\mathfrak{N}$, M is the fixed point set of j;*
(c) *Some model $\mathfrak{M}$ of T has a proper end extension $\mathfrak{N}$ such that $\mathfrak{N} \vDash I\Delta_0$ and for some automorphism j of $\mathfrak{N}$, M is the fixed point set of j.*

The proof of Theorem A relies on Lemmas A.0 through A.4 below. Lemmas A.0 and A.1 show the preliminary result that $\mathfrak{M}$ satisfies $I\Delta_0 + Exp + Superexp$ (where $Superexp$ is the axiom stating that the function $Superexp(x, x)$ is total). Indeed, the strategy of the proof of Lemma A.1 can be used to establish that M is closed under all primitive recursive functions, thus showing that $\mathfrak{M}$ is a model of PRA (primitive recursive arithmetic). However, the totality of the Ackermann function does not seem to be obtainable via this strategy. These first two Lemmas are used in Lemma A.2 to show that we can replace the end extension $\mathfrak{N}$ of $\mathfrak{M}$ in Theorem A, if necessary, by a model of $I\Delta_0 + Exp$. Lemma A.2 and Theorem 2.2 together allow us the luxury of accessing a decent amount of set theory within an initial segment of $\mathfrak{N}$ containing M via Ackermann coding, thereby providing streamlined proofs of the central Lemmas A.3 and A.4 without having to go through laborious calculations dealing with Ackerman coding.

- For the rest of this section we make the blanket assumption that $\mathfrak{M}, \mathfrak{N}$, and j are as in the statement of Theorem A. In particular, M is the fixed point set of j, and $\mathfrak{N}$ is a proper end extension of $\mathfrak{M}$.

LEMMA A.0. $\mathfrak{M} \vDash I\Delta_0$.

PROOF: Clearly M is closed under the operations of $\mathfrak{N}$. Since Δ_0-predicates are absolute for end extensions, this shows that $\mathfrak{M}$ inherits $I\Delta_0$ from $\mathfrak{N}$.
□

LEMMA A.1. *Exp and Superexp both hold in $\mathfrak{M}$.*

PROOF: We only verify Exp in $\mathfrak{M}$ since the verification of the totality of the superexponential function uses an identical strategy and is left to the reader. Recall that there is a Δ_0-predicate that reasonably expresses "$2^x = y$". Let

$$I := \{x \in N : \mathfrak{N} \vDash \exists y(2^x = y)\}.$$

Note that I is closed downward in $\mathfrak{M}$ and $I \cap M$ has no last element since $\mathfrak{M} \vDash I\Delta_0$ and $I\Delta_0$ is able to prove that the set of numbers x on which 2^x is defined is closed under both predecessors and immediate successors. To show that Exp holds in $\mathfrak{M}$, it suffices to show that $M \subseteq I$ since if x is fixed by j, and 2^x exists in $\mathfrak{N}$, then 2^x is definable from x within $\mathfrak{N}$ and must therefore also be fixed by j. Next, let

$$J := \{y \in N : \mathfrak{N} \vDash \exists x \ (2^x = y)\}.$$

It is easy to see that if J is unbounded in N then $M \subseteq I$, so our proof would be complete once we establish that J is unbounded in $\mathfrak{N}$. Suppose, on the contrary, that some $a \in N$ is an upper bound of J. Then the set

$$\{y < a : \mathfrak{N} \vDash \exists x < a \ (2^x = y)\}$$

has a maximum element by Lemma A.0 and Δ_0-MAX (Lemma 2.1(b)) since it is the solution set of a Δ_0-predicate, thus leading to the absurd conclusion that I has a maximum element.
□

LEMMA A.2. *There is an initial segment N^* of $\mathfrak{N}$ that properly contains M such that $\mathfrak{N}^* := (N^*, \cdots)$ is a model of $I\Delta_0 + Exp$, and $j \restriction \mathfrak{N}^*$ is an automorphism.*

PROOF: Let $\psi(x, y)$ be a Δ_0 predicate for $y = SuperExp(x, x)$ and let $b \in N \backslash M$. By Lemma A.1 for every $m \in M$,

$$\mathfrak{N} \vDash \exists y < b \ \psi(m, y).$$

Therefore, by Δ_0-OVERSPILL (Lemma 2.1(c)) there is an element $a \in N \backslash M$ for which $SuperExp(a, a)$ is well-defined in $\mathfrak{N}$. This implies that the elements

$$2^a, 2^{2^a}, \cdots, SuperExp(n, a), \cdots (n \in \omega)$$

are all well-defined within $\mathfrak{N}$. To define N^*, assume without loss of generality that $a < j(a)$ (else replace j by j^{-1}), and let

$$N^* := \bigcup_{k \in \omega} \bigcup_{n \in \omega} [0, SuperExp(n, j^k(a))]^{\mathfrak{N}}.$$

It is easy to verify that $\mathfrak{N}^* \vDash I\Delta_0 + Exp$, *and* $j \restriction \mathfrak{N}^*$ is an automorphism.
□

Before establishing the next lemma[6], we need to recall the key notion of strong cuts, first introduced by Kirby and Paris [KP]:

- Suppose $\mathfrak{N}$ is a model of $I\Delta_0$ and M is a cut of $\mathfrak{N}$. M is a *strong cut* of $\mathfrak{N}$, if for each function f whose graph is coded in $\mathfrak{N}$ (via Ackermann's $\in$) and whose domain includes M, there is some s in N, such that for all $m \in M$,

$$f(m) \notin M \ \text{ iff } \ s < f(m).$$

LEMMA A.3. *M is a strong cut of $\mathfrak{N}$.*

PROOF: We first observe that it suffices to show that $\mathfrak{M}$ is a strong cut of the model $\mathfrak{N}^*$ of Lemma A.3. Recall that by Theorem 2.2, we have access to "bounded" set theoretic reasoning within $\mathfrak{N}^*$. Suppose $\overline{f} \in N^*$ codes the graph of a function f whose domain includes M. It is easy to see that $\overline{f} \notin M$. So if $\overline{g} := j(\overline{f})$, then $\overline{g} \notin M$, and $\overline{f} \neq \overline{g}$. Therefore, if g is the function that is coded by $\overline{g}$, then:

$$\forall m \in M \ [f(m) = g(m) \Longleftrightarrow f(m) \in M].$$

[6]This lemma was inspired by the results of Kaye, Kossak, and Kotlarski [KKK].

We wish to find $s \in N^*$ such that for all $m \in M$, $f(m) \notin M$ iff $s < f(m)$. Without loss of generality there is some $m_0 \in M$ with $f(m_0) \notin M$. Fix $c \in N^*$ such that c_E contains $\overline{f}, \overline{g}$, and every $m \in M$ (recall: c_E is $\{x \in N : xEc\}$, where E is Ackermann's $\in$ in the sense of $\mathfrak{N}$). Consider the function $h(x)$ defined within $\mathfrak{M}$ on the interval $[m_0, c]$ by

$$h(x) := \mu y \leq c\ [\exists z \leq x (y = f(z) \neq g(z)],$$

where $\mu y \leq c$ is the (truncated) least number operator, defined via the equation

$$[z := \mu y \leq c\ \varphi(y)] \text{ iff } [z \text{ is the first solution } y \text{ of } \varphi, \text{ if } y \leq c; \text{ else } z = c].$$

Note that if $m \in M$ with $m_0 \leq m$ then $h(m) \notin M$, and if $m_0 \leq m \leq m'$, then $h(m') \leq h(m)$. Moreover,

(1) the graph of h is defined by a Δ_0-formula $\varphi(x, y)$ with parameters $\overline{f}$ and $\overline{g}$; and
(2) $m < h(m)$ for all $m \in M$ with $m \geq m_0$.

Therefore, (1), (2), and Δ_0-OVERSPILL (Lemma 2.1(c)) within $\mathfrak{N}^*$ together imply that there is some $s \in N^* \backslash M$ such that $s < h(s)$ holds in $\mathfrak{N}^*$. This shows that s is the desired lower bound for elements of the form $f(m)$, where $m \in M$ and $f(m) \notin M$.
□

Kirby and Paris proved that strong cuts of models of PA are themselves models of PA [KP, Proposition 8]. An analysis of their proof reveals the stronger result below[7].

LEMMA A.4. *Let $\mathcal{A} := SSy_M(\mathfrak{N})$ and $\mathcal{L} := \mathcal{L}_A \cup \{S : S \in \mathcal{A}\}$. For every $\mathcal{L}$-formula*

$$\varphi(x_1, \cdots, x_m),$$

with free variables $x_1 \cdots, x_m$, there is some Δ_0-formula

$$\theta_\varphi(x_1, \cdots, x_m, b_1, \cdots, b_n),$$

where $b_1, \cdots, b_n$ is a sequence of parameters from N, such that for all sequences $a_1, \cdots, a_m$ of elements of M,

$$(\mathfrak{M}, S)_{S \in \mathcal{A}} \vDash \varphi(a_1, \cdots, a_m) \ \textit{ iff } \ \mathfrak{N} \vDash \theta_\varphi(a_1, \cdots, a_m, b_1, \cdots, b_n).$$

PROOF: In what follows $\mathfrak{N}^*$ is as in Lemma A.2. θ_φ is built by recursion on the complexity of φ :

- If φ is an atomic formula of the form $S_i(v)$, where v is a term, then choose $b \in N$ such that $b_F \cap M = S_i$, and define $\theta_\varphi := (v \in b)$. For other atomic formulas φ, $\theta_\varphi := \varphi$.
- $\theta_{\neg\delta} := \neg\theta_\delta$;
- $\theta_{\delta_1 \vee \delta_2} := \theta_{\delta_1} \vee \theta_{\delta_2}$;
- If $\varphi = \exists v \delta(v, x_1, \cdots, x_t)$, then fix some $c \in N \backslash M$ and consider the function $f(x_1, \cdots, x_t)$ defined in $\mathfrak{N}^*$ on $[0, c]^t$ by:

$$f(x_1, \cdots, x_t) := \begin{cases} \mu v \leq c \text{ such that } \theta_\delta(v, x_1, \cdots, x_t), \text{ if } \exists v \in c\ \theta_\delta(v, x_1, \cdots, x_t); \\ 0, \text{ otherwise.} \end{cases}$$

Note that the graph of f is defined by a $\Delta_0(\mathcal{L})$-formula within $\mathfrak{N}^*$ and so by Theorem 2.2 f is coded in $\mathfrak{N}^*$ and therefore in $\mathfrak{N}$. Hence, we can use

[7] As noted by one of the referees, this result also appears in Kirby's dissertation [Ki-1].

Lemma A.3 to invoke the strength of M in $\mathfrak{N}$ to find some $s \in N$, such that for all $m \in M$, $f(m) \in M$ iff $f(m) \leq s$. Now define:

$$\theta_\varphi := \exists v \leq s\ \theta_\delta(v, x_1, \cdots, x_t).$$

□

Proof of Theorem A: Let $\mathcal{A}$ and $\mathcal{L}$ be as in Lemma A.4. It is easy to see that every nonempty member of $\mathcal{A}$ has a first element in $\mathfrak{M}$ (since $\mathfrak{N}$ satisfies $I\Delta_0$). To establish the arithmetical comprehension scheme in $(\mathfrak{M}, \mathcal{A})$, consider any $\mathcal{L}$-formula $\varphi(x)$ with precisely one free variable x. We wish to show that

$$\{m \in M : (\mathfrak{M}, \mathcal{A}) \vDash \varphi(m)\} \in \mathcal{A}.$$

Let θ_φ be as in Lemma A.4 and fix some $c \in N \backslash M$. By Theorem 2.2 (part 5), there is an element $d \in N$ that codes $\{x < c : N \vDash \theta_\varphi(x)\}$. Therefore, by Lemma A.4

$$\{m \in M : (\mathfrak{M}, \mathcal{A}) \vDash \varphi(m)\} = d_E \cap M \in \mathcal{A}.$$

□

3.2. Automorphisms from $\mathbf{ACA_0}$

The principal result of this section is Theorem B. We should emphasize that Theorem B follows from Gaifman's work in [G], but we have decided to present a detailed proof here for two reasons. Firstly, this theorem is only implicit in Gaifman's paper, and therefore a detailed presentation of this significant result is of some value. Secondly, the method of *iterated ultrapowers modulo generic ultrafilters* developed here for the proof of Theorem B is also employed in the proof of Theorem C (Section 4.1) and a detailed development in this section allows us to later skip some details in the proof of Theorem C.

Theorem B. *Suppose $(\mathfrak{M}, \mathcal{A})$ is a countable model of ACA_0. There is a proper elementary end extension $\mathfrak{N}$ of $\mathfrak{M}$ which satisfies the following two properties:*

(a) *$\mathfrak{N}$ possesses an automorphism j whose fixed point set is precisely M;*
(b) *$SSy_M(\mathfrak{N}) = \mathcal{A}$.*

The proof of Theorem B is presented at the end of this section once the machinery of generic ultrafilters and iterated ultrapowers have been put into place. However, we can easily describe the high-level strategy of the proof: $\mathfrak{N}$ is obtained by an iterated $\mathfrak{M}$-ultrapower along the linearly ordered set of integers $\mathbb{Z}$ modulo a "generic ultrafilter", and the desired automorphism j of $\mathfrak{N}$ is induced by the automorphism $n \mapsto n + 1$ of $\mathbb{Z}$.

3.2.1. Generic Ultrafilters

Suppose $(\mathfrak{M}, \mathcal{A})$ is a countable model of ACA_0. Clearly $\mathcal{A}$ is a Boolean algebra. Our goal is to construct ultrafilters $\mathcal{U}$ over $\mathcal{A}$ with certain desirable combinatorial properties. We shall employ the conceptual framework of forcing in order to efficiently present the necessary bookkeeping arguments in our construction[8]. Let $\mathbb{P}$ be the poset

$$\{S \in \mathcal{A} : S \text{ is unbounded in } (M, <)\},$$

ordered under inclusion.

[8]A closely related notion of forcing, formulated by Gaifman, was employed in [AH, Sec. 1].

- A subset $\mathcal{D}$ of $\mathbb{P}$ is *dense* if for every $X \in \mathbb{P}$ there is some $Y \in \mathcal{D}$ with $Y \subseteq X$.
- $\mathcal{U} \subseteq \mathbb{P}$ is a *filter* if it is (1) closed under intersections and (2) is upward closed.
- A filter $\mathcal{U} \subseteq \mathbb{P}$ is $\mathcal{A}$-*generic over* $(\mathfrak{M}, \mathcal{A})$ if $\mathcal{U}$ meets every dense subset $\mathcal{D}$ of $\mathbb{P}$ which is parametrically definable in $(\mathfrak{M}, \mathcal{A})$.
- A filter $\mathcal{U} \subseteq \mathbb{P}$ is $(\mathfrak{M}, \mathcal{A})$-*complete* if for every $f : M \to [0, a]^{\mathfrak{M}}$, where $a \in M$ and $f \in \mathcal{A}$, there is some $X \in \mathcal{U}$ such that f is constant on X.

Note that if $\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-complete, then $\mathcal{U}$ is a nonprincipal ultrafilter on $\mathcal{A}$ since for each $Y \in \mathcal{A}$, the characteristic function of Y is constant on some member of $\mathcal{U}$. We therefore refer to $(\mathfrak{M}, \mathcal{A})$-complete filters as *ultrafilters.* Generic ultrafilters have some special combinatorial properties. To discuss them we need the following definitions and theorems.

- Let Γ be a canonical bijection between $M \times M$ and M. Every $g : M \to \{0, 1\}$ codes a sequence $\langle S^g_a : a \in M \rangle$ of subsets of M, where
$$S^g_a := \{b \in M : g(\Gamma(a, b)) = 1\}.$$
- A filter $\mathcal{U} \subseteq \mathbb{P}$ is $(\mathfrak{M}, \mathcal{A})$-*iterable*[9] if $\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-complete, and for every $g \in \mathcal{A}$ and $g : M \to \{0, 1\}$,
$$\{a \in M : S^g_a \in \mathcal{U}\} \in \mathcal{A}.$$
- Given a linearly order set $(M, <)$, $[M]^n$ is the set of *increasing* n-tuples from M.
- Suppose $(M, <)$ is a linear order and $f : [M]^n \to M$. A subset X of M is f-*canonical* if there is some $S \subseteq \{1, \cdots, n\}$ such that for all sequences $s_1 < \cdots < s_n$, and $t_1 < \cdots < t_n$ of elements of X,
$$f(s_1, \cdots, s_n) = f(t_1, \cdots, t_n) \iff \forall i \in S \ (s_i = t_i).$$
Note that if $S = \emptyset$, then f is constant on $[X]^n$, and if $S = \{1, \cdots, n\}$, then f is injective on $[X]^n$.
- A filter $\mathcal{U} \subseteq \mathbb{P}$ is $(\mathfrak{M}, \mathcal{A})$-*canonically Ramsey* if for every $f : [M]^n \to M$, where n is a standard natural number, with $f \in \mathcal{A}$, there is some $X \in \mathcal{U}$ on which f is canonical.
- $\omega \to *(\omega)^n$ is the statement in the language of second order arithmetic which asserts that for every $f : [\omega]^n \to \omega$ there is an unbounded $X \subseteq \omega$ such that X is f-canonical.

Erdös and Rado [ER] proved that $\omega \to *(\omega)^n$ holds for all $n < \omega$. Their proof derives $\omega \to *(\omega)^n$ from $\omega \to (\omega)^{2n}$ and is readily formalizable[10] in ACA_0 for each fixed standard n, i.e.,

THEOREM 3.2. $\forall n \in \omega$, $ACA_0 \vdash \omega \to *(\omega)^n$.

REMARK 3.2.1. If ACA_0 is replaced by Z_2 (or just ACA_0 plus the full schema of induction) then "$\forall n \in \omega$" can be moved to the right hand side of the provability

[9]This terminology is motivated by the fact (discussed in Section 3.2.2) that the formation of ultrapowers modulo iterable ultrafilters is amenable to iteration. Iterable ultrafilters are also referred to as *definable* ultrafilters, e.g., as in [Ki-2], motivated by their intimate link with the model theoretic notion of *definable type*.

[10]The text [GRS] includes a detailed proof of a special case of Theorem 3.2. See also [Ra] and [Mile] for more perspicuous proofs of the full result.

symbol $\vdash$. It is known that $ACA_0 \nvdash \forall n \in \omega\ \omega \to (\omega)^n$. This follows from a theorem of Jockusch [Jo], which states that for each natural number $n \geq 2$ there is a recursive partition P_n of $[\omega]^n$ into two parts such that P_n has no infinite Σ_n-homogeneous subset[11].

The usual proof establishing the existence of filters meeting countably many dense sets shows:

PROPOSITION 3.3. *There is a generic filter $\mathcal{U}$ over every countable model $(\mathfrak{M}, \mathcal{A})$.*

The following result reveals the key properties of generic ultrafilters.

THEOREM 3.4. *If $(\mathfrak{M}, \mathcal{A})$ is a model of ACA_0 and $\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-generic, then*

(a) *$\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-complete;*
(b) *$\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-iterable;*
(c) *$\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-canonically Ramsey.*

PROOF:

(a): Given $f \in \mathcal{A}$ with $f : M \to [0, a]^{\mathfrak{M}}$, let

$$\mathcal{D}_1^f := \{Y \in \mathbb{P} : f \restriction Y \text{ is constant}\}.$$

$\mathcal{D}_1^f$ is dense since for each $X \in \mathbb{P}$, $(\mathfrak{M}, X, f) \vDash PA(X, f)$.

(b): For X and Y in $\mathbb{P}$, let us write $X \subseteq_* Y$ (read: "X is almost contained in Y") if $X \backslash Y$ is bounded in $(M, <)$. Also, let "X *decides* Y" abbreviate

$$\text{"}X \subseteq_* Y \text{ or } X \subseteq_* M \backslash Y\text{"}.$$

Observe that to establish (b) it suffices to show that if $g : M \to \{0, 1\}$, with $g \in \mathcal{A}$, then

$$\mathcal{D}_2^g = \{Y \in \mathbb{P} : \forall a \in M,\ Y \text{ decides } S_a^g\} \text{ is dense.}$$

To show the density of $\mathcal{D}_2^g$ suppose $X \in \mathbb{P}$. We first claim that there is an $\mathcal{A}$-coded sequence $F = \langle F_a : a \in M \rangle$ satisfying the following two properties:

$(*)$ $\forall a \in M$, $F_a = S_a^g \cap X$ or $F_a = X \backslash S_a^g$;

$(**)$ $\forall a \in M \bigcap\limits_{b \leq a} F_b$ is unbounded in X.

Argue within $(\mathfrak{M}, \mathcal{A})$. For each $s : [0, a] \to \{0, 1\}$, define $\langle F_b^s : b \leq a \rangle$ by:

$$F_b^s := \begin{cases} S_b^g \cap X, \text{ if } s(b) = 1; \\ X \backslash S_b^g, \text{ if } s(b) = 0. \end{cases}$$

Consider the subtree τ of $(2^{<\omega})^{\mathfrak{M}}$ consisting of functions $s : [0, a] \to \{0, 1\}$ such that $\bigcap\limits_{b \leq a} F_b^s$ is unbounded in X. It is easy to see that τ has nodes of every rank $b \in M$, because each level of τ gives rise to a partition of X into 2^b pieces, so one of the pieces must be unbounded since X itself is unbounded. By König's lemma, τ has a branch, which yields the desired sequence $\langle F_a : a \in M \rangle$.

We can now define $Y = \{y_a : a \in M\} \in \mathbb{P}$ by induction within $(\mathfrak{M}, \mathcal{A})$ such that Y is almost contained in every F_a as follows:

- y_0 is the first element of F_0;
- y_{a+1} is the least member of $\bigcap\limits_{b \leq a} F_b$ above $\{y_b : b \leq a\}$.

[11] See [W, p.25] for more detail on this matter. Note that ACA_0 is referred to as PPA (predicative Peano arithmetic) in [W].

It is clear that Y decides each S^g_a. Therefore $\mathcal{D}^g_2$ is dense.

(c): Suppose $f : [M]^n \to M$, where n is a standard natural number, and $f \in \mathcal{A}$. Let

$$\mathcal{D}^f_3 := \{Y \in \mathbb{P} : f \text{ is canonical on } Y\}.$$

By Theorem 3.2, $\mathcal{D}^f_3$ is dense.
□

REMARK 3.4.1. By a theorem of Kunen, a Rudin-Keisler minimal ultrafilter on $\mathcal{P}(\omega)$ is already a Ramsey ultrafilter [Jec, Lemma 38.1]. Moreover, the proof of the Erdös-Rado canonical partition theorem can be used to show that a Ramsey ultrafilter on $\mathcal{P}(\omega)$ is also canonically Ramsey. In the context of models of ACA_0, it is known that if $\mathcal{U}$ is 3-Ramsey[12] over $(\mathfrak{M}, \mathcal{A})$, then $\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$-iterable and n-Ramsey for all $n \in \omega$ [Ki-2, Theorem 2.4]. Coupled with [Ki-2, Theorem 6.5] and the aforementioned Erdös-Rado proof, this shows that the following are equivalent for an ultrafilter $\mathcal{U}$ over a model $(\mathfrak{M}, \mathcal{A})$ of ACA_0 :

(*i*) $\mathcal{U}$ *is* 3-Ramsey over $(\mathfrak{M}, \mathcal{A})$;
(*ii*) $\mathcal{U}$ is both iterable and canonically Ramsey over $(\mathfrak{M}, \mathcal{A})$;
(*iii*) $\mathcal{U}$ is a minimal end extension type over $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ in the sense of Gaifman [G] (i.e., $\mathcal{U}$ is an iterable ultrafilter over $(\mathfrak{M}, \mathcal{A})$ and for every function $f \in \mathcal{A}$ with $f : M \to M$, f is one-to-one or constant on a member of $\mathcal{U}$).

It is also worth pointing out that the converse of Theorem 3.4 is false, i.e., "$\mathcal{U}$ is generic over $(\mathfrak{M}, \mathcal{A})$" is stronger than the above three conditions. This is a consequence of the fact that (a) generic ultrafilters are not first order definable in $(\mathfrak{M}, \mathcal{A})$, and (b) there is a Ramsey ultrafilter on $\mathcal{P}^{\mathbf{L}}(\omega)$ (the powerset of ω in the sense of Gödel's constructible universe) that is first order definable within the model $(\omega, +, \cdot, \mathcal{P}^{\mathbf{L}}(\omega))$. (a) follows from a standard forcing argument, and (b) can be established by coupling the fact that there is a well-ordering of $\mathcal{P}^{\mathbf{L}}(\omega)$ that is definable in $(\omega, +, \cdot, \mathcal{P}^{\mathbf{L}}(\omega))$ [Jec, Theorem 97] with the proof of the existence of a Ramsey ultrafilter assuming the continuum hypothesis [Jec, p.478].

3.2.2. Ultrapowers and Iterations

Gaifman [G] refined the MacDowell-Specker Theorem by showing that if $\mathcal{L}$ is a *countable*[13] language extending $\mathcal{L}_A$, $\mathfrak{M}$ is a model of $PA(\mathcal{L})$ of *any cardinality*, and $\mathcal{A}$ is the family of definable subsets of $\mathfrak{M}$, then there is an e.e.e. $\mathfrak{N}$ of $\mathfrak{M}$ such that $\mathcal{A} = SSy_M(\mathfrak{N})$. In the jargon of model theorists of arithmetic, this is rephrased as: if $\mathcal{L}$ is countable, then every model of $PA(\mathcal{L})$ has a *conservative* e.e.e. The first result of this section is an adaptation of Gaifman's result tailormade for our purposes.

LEMMA 3.5. *Suppose* $(\mathfrak{M}, \mathcal{A})$ *is a model of* ACA_0. *The following two conditions are equivalent:*

(a) *There exists a nonprincipal* $(\mathfrak{M}, \mathcal{A})$*-iterable ultrafilter* $\mathcal{U}$ *over* $(\mathfrak{M}, \mathcal{A})$.
(b) $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ *has a proper e.e.e.* $(\mathfrak{N}, S^*)_{S\in\mathcal{A}}$ *such that* $\mathcal{A} = SSy_M(\mathfrak{N})$.

[12]Here $\mathcal{U}$ is *n-Ramsey* over $(\mathfrak{M}, \mathcal{A})$ if for every $f : [M]^n \to \{0, 1\}$ with $f \in \mathcal{A}$, there is some $X \in \mathcal{U}$ on which f is homogeneous.

[13]Mills [Mill] used a forcing construction to show that the countability assumption cannot be dropped from Gaifman's result.

PROOF: To show $(a \Rightarrow b)$, let $(\mathfrak{N}, S^*)_{S\in\mathcal{A}}$ be the ultrapower of $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ modulo $\mathcal{U}$, i.e., the universe N of $\mathfrak{N}$ consists of the $\mathcal{U}$-equivalence classes $[f]$ of functions f from M into M such that f is coded by some element of $\mathcal{A}$, and the operations on N are defined as in the classical theory of ultrapowers, e.g., $+^{\mathfrak{N}}$ is defined by

$$[f] \ +^{\mathfrak{N}} \ [g] = [h] \text{ iff } \{m \in M : f(m) +^{\mathfrak{M}} g(m) = h(m)\} \in \mathcal{U}.$$

Similarly, for each $S \in \mathcal{A}$,

$$[f] \in S^* \text{ iff } \{m \in M : f(m) \in S\} \in \mathcal{U}.$$

The Łoś Theorem for ultrapowers goes through in this limited context, thanks to the fact that every parametrically definable subset of $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ has a $<^{\mathfrak{M}}$-least element (and therefore the model $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ has definable Skolem functions). Consequently, if $\mathcal{U}$ is a non-principal ultrafilter, then $\mathfrak{N}$ is a proper elementary extension of $\mathfrak{M}$ (with the obvious identification of the $\mathcal{U}$-equivalence classes of constant maps with elements of M). It remains to verify (i) and (ii) below:

(i) $\mathfrak{M} \subseteq_e \mathfrak{N}$, and
(ii) $\mathcal{A} = SSy_M(\mathfrak{N})$.

To verify (i), suppose $\mathfrak{N} \vDash [f] \leq m$ for some $m \in M$. Then for some $X \in \mathcal{U}$, $\forall x \in X \ f(x) < m$. Let f^* be the function in $\mathcal{A}$ defined by $f^*(x) = f(x)$ if $x \in X$, and 0 otherwise. By $(\mathfrak{M}, \mathcal{A})$-completeness of $\mathcal{U}$, there is some $m_0 \leq m$ and some $Y \in \mathcal{U}$ such that $\forall x \in Y \ f^*(x) = m_0$. It is now easy to verify that $\mathfrak{N} \vDash [f] = [f^*] = m_0$. To establish (ii), first, note that for each $X \in \mathcal{A}$,

$$(\mathfrak{M}, X) \prec_e (\mathfrak{N}, X^*) \vDash PA(X^*).$$

This shows that $\mathcal{A} \subseteq SSy_M(\mathfrak{N})$ since if $d \in N\backslash M$, there is some $c \in N$ such that c precisely codes those elements of X^* which are less than d. Therefore, $X = c_E \cap M$. To see that $SSy_M(\mathfrak{N}) \subseteq \mathcal{A}$ we need to invoke the assumption of iterability of $\mathcal{U}$. Given an element $[f] \in N$, we wish to show

(1) $\{m \in M : \mathfrak{N} \vDash mE[f]\} \in \mathcal{A}$.

Observe that (1) is equivalent to

(2) $\{m \in M : \{n \in M : \mathfrak{M} \vDash mEf(n)\} \in \mathcal{U}\} \in \mathcal{A}$.

Let $X_m = \{n \in M : \mathfrak{M} \vDash mEf(n)\}$. By the iterability assumption,

(3) $\{m \in M : X_m \in \mathcal{U}\} \in \mathcal{A}$.

Therefore (1) holds. This completes the proof of (ii).

To show $(b \Rightarrow a)$, assume (b) holds and fix $c \in N\backslash M$. Consider $\mathcal{U}$ defined by

$$\mathcal{U} := \{S \in \mathcal{U} : c \in S^*\}.$$

The assumption that $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ is elementarily end extended by $(\mathfrak{N}, S^*)_{S\in\mathcal{A}}$ can now be invoked to verify that $\mathcal{U}$ is $(\mathfrak{M}, \mathcal{A})$ complete, for if $f \in \mathcal{A}$, $a \in M$, and $(\mathfrak{M}, f) \vDash$ "$f : M \to [0, a]$", then $(\mathfrak{N}, f^*) \vDash$ "$f^* : N \to [0, a]$". Note that since $\mathfrak{N}$ is end extended by $\mathfrak{M}$, $f^*(c) \in M$. It is now easy to verify that

$$\{m \in M : f(m) = f^*(c)\}$$

is the desired member of $\mathcal{U}$ on which f is constant. Similarly, by invoking the assumption $\mathcal{A} = SSy_M(\mathfrak{N})$ we can show that $\mathcal{U}$ is also $(\mathfrak{M}, \mathcal{A})$-iterable, since if

$(\mathfrak{M}, g) \vDash$ "$g : M \to \{0,1\}$", where $g \in \mathcal{A}$, then $X_g := \{m \in M : c \in (S^g_m)^*\}$ is a member of $\mathcal{U}$, and therefore

$$\{m \in M : S^g_a \in \mathcal{U}\} = X_g \in \mathcal{A}.$$

□

For an $(\mathfrak{M}, \mathcal{A})$-iterable ultrafilter $\mathcal{U}$, the fact that the $\mathcal{U}$-based ultrapower does not introduce new subsets of $\mathfrak{M}$ allows one to *iterate the ultrapower formation any finite number of times* to obtain the n-fold iterations $Ult_{\mathcal{U},n}(\mathfrak{M}, S)_{S\in\mathcal{A}}$ for each positive natural number n. Indeed, a finite iteration of length n can be obtained in *one step* by defining an ultrafilter $\mathcal{U}^n$ on M^n. To do so, suppose $X \subseteq M^{n+1}$ is coded in $\mathcal{A}$. By definition[14],

(♣) $X \in \mathcal{U}^{n+1}$ iff $\{\alpha_1 : \{\langle \alpha_2, \cdots, \alpha_{n+1}\rangle : \langle \alpha_1, \alpha_2, \cdots, \alpha_{n+1}\rangle \in X\} \in \mathcal{U}^n\} \in \mathcal{U}$.

Remark 3.6. It is easy to see that $\mathcal{U}^n$ concentrates on $[M]^n$. Moreover, if $\mathcal{U}$ is n-Ramsey over $(\mathfrak{M}, \mathcal{A})$ for some $n \in \omega$, then

$$\mathcal{U}^n = \{Y \subseteq M^n : \exists X \in \mathcal{U}\ \ [X]^n \subseteq Y\}.$$

The process of ultrapower formation modulo $\mathcal{U}$ can be iterated *along any linear order* $\mathbb{L}$ to yield the iterated ultrapower $Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}}$. To describe the isomorphism type of $Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}}$ $\mathcal{U}$ one can either use a direct limit construction (as originally formulated by Kunen [Ku], and often used in set theoretic literature) or, equivalently, one can take the following model theoretic route (as in Gaifman [G]). Given an iterable ultrafilter $\mathcal{U}$ we can define, for each positive natural number n, a complete n-type Γ_n over the model $(\mathfrak{M}, S)_{S\in\mathcal{A}}$ by defining $\Gamma_n(x_1, \cdots, x_n)$ as the set of formulas $\varphi(x_1, \cdots, x_n)$ such that

$$\{\langle \alpha_1, \cdots, \alpha_n\rangle : (\mathfrak{M}, S)_{S\in\mathcal{A}} \vDash \varphi(\alpha_1, \cdots, \alpha_n)\} \in \mathcal{U}^n.$$

Here φ is a formula in the language $\mathcal{L} = \mathcal{L}_A \cup \{S : S \in \mathcal{A}\}$ (since for each $m \in M$, $\{m\} \in \mathcal{A}$, for all intents and purposes $\mathcal{L}$ has constant symbols for elements of M as well). Then we augment the language $\mathcal{L}$ with a set of new constant symbols $\{\bar{l} : l \in \mathbb{L}\}$, and define $T_{\mathcal{U},\mathbb{L}}$ to consist of formulas of the form $\varphi(\bar{l}_1, \bar{l}_2, \cdots, \bar{l}_n)$, where $\varphi(x_1, \cdots, x_n) \in \Gamma_n(x_1, \cdots, x_n)$ and $l_1 <_{\mathbb{L}} \cdots <_{\mathbb{L}} l_n$. Since $T_{\mathcal{U},\mathbb{L}}$ is a *complete Skolemized theory*, $Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}}$ can be meaningfully defined as the *prime model* of $T_{\mathcal{U},\mathbb{L}}$.

The following theorem, due to Gaifman [G], summarizes the key properties of iterated ultrapowers[15].

Theorem 3.7. *Suppose $\mathcal{U}$ is an $(\mathfrak{M}, \mathcal{A})$-iterable ultrafilter over a model $(\mathfrak{M}, \mathcal{A})$ of ACA_0, and $\mathbb{L}$ is a linearly ordered set. Let $(\mathfrak{N}, S^*)_{S\in\mathcal{A}} := Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}}$, and $c_l := \left(\bar{l}\right)^{\mathfrak{N}}$.*

(a) *Elements of N are of the form $f^*(c_{l_1}, \cdots, c_{l_n})$, where $f \in \mathcal{A}$, and $l_1 <_{\mathbb{L}} \cdots <_{\mathbb{L}} l_n$;*

(b) *For every $\mathcal{L}$-formula $\varphi(x_1, \cdots, x_n)$, and every increasing sequence $l_1 <_{\mathbb{L}} \cdots <_{\mathbb{L}} l_n$*

$$\mathfrak{N} \vDash \varphi(c_{l_1}, \cdots, c_{l_n}) \text{ iff } \{\langle \alpha_1, \cdots, \alpha_n\rangle \in M^n : \mathfrak{M} \vDash \varphi(\alpha_1, \cdots, \alpha_n)\} \in \mathcal{U}^n;$$

(c) *$\{c_l : l \in \mathbb{L}\}$ is a set of order indiscernibles in $(\mathfrak{N}, S^*)_{S\in\mathcal{A}}$;*

[14]The iterability condition is invoked to ensure that $\mathcal{U}^{n+1}$ is well-defined via (♣).

[15]The analogue of this result for models of set theory with a weakly compact cardinal is due to Kunen [Ku], and fully developed in [Jec] and [Kan].

(d) *Every automorphism h of $\mathbb{L}$ induces an automorphism*

$$j_h : (\mathfrak{N}, S^*)_{S\in\mathcal{A}} \to (\mathfrak{N}, S^*)_{S\in\mathcal{A}}$$

defined by

$$j_h(f^*(c_{l_1}, \cdots, c_{l_n}) = f^*(c_{h(l_1)}, \cdots, c_{h(l_n)}).$$

- If $\mathcal{U}$ is also canonically Ramsey, then Theorem 3.7(d) can be strengthened as follows:

THEOREM 3.8. *Suppose $(\mathfrak{M}, \mathcal{A}) \vDash ACA_0$, and let h is an automorphism of a linearly ordered set $\mathbb{L}$ with no fixed points. If $\mathcal{U}$ is iterable and canonically Ramsey over $(\mathfrak{M}, \mathcal{A})$, then the fixed point set of the automorphism j_h of $Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}}$ is precisely M.*

PROOF: Clearly j_h fixes each $a \in M$ since the constant map $f_a(x) = a$ is in $\mathcal{A}$. To see that j_h fixes no member of $N\backslash M$, suppose that

(1) $f^*(c_{h(l_1)}, \cdots, c_{h(l_n)}) = f^*(c_{l_1}, \cdots, c_{l_n})$

for some $f^*(c_{l_1}, \cdots, c_{l_n}) \in N$. Since $f \in \mathcal{A}$, by Theorem 3.4(c) there is some $X \in \mathcal{U}$, and some $S \subseteq \{1, \cdots, n\}$ such that for all sequences $a_1 < \cdots < a_n$, and $b_1 < \cdots < b_n$ of elements of X,

(2) $f(a_1, \cdots, a_n) = f(b_1, \cdots, b_n) \iff \forall i \in S\ (a_i = b_i)$.

Moreover, since $X^n \in \mathcal{U}^n$,

(3) $Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}} \vDash \langle c_{l_1}, \cdots, c_{l_n}\rangle \in X^n$.

(1), (2), and (3) together imply that $S = \emptyset$, which in turn implies that f must be *constant* on X. Therefore, $f^*(c_{l_1}, \cdots, c_{l_n}) \in M$.
□

PROOF OF THEOREM B: Let $(\mathfrak{M}, \mathcal{A})$ be a model of ACA_0. Fix some $(\mathfrak{M}, \mathcal{A})$-generic ultrafilter $\mathcal{U}$ and let

$$(\mathfrak{N}, S^*)_{S\in\mathcal{A}} := Ult_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S\in\mathcal{A}},$$

where $\mathbb{Z}$ is the ordered set of integers. Consider the automorphism

$$n \longmapsto_h n + 1$$

of $\mathbb{Z}$. By Theorems 3.4 and 3.8 j_h is an automorphism of $(\mathfrak{N}, S^*)_{S\in\mathcal{A}}$ whose fixed point set is precisely M.
□

3.3. An Arithmetical Theory with a Built-in Automorphism

Consider the theory VA formulated in $\mathcal{L}_A \cup \{j\}$, where j is a unary function symbol, obtained by augmenting the axioms of $I\Delta_0$ with a single axiom expressing "j is a nontrivial $\{+, \cdot\}$-automorphism whose fixed-point set is closed downwards".

This theory was formulated by Albert Visser who noted that Corollary 3.1 implies that PA can be interpreted in VA. This led Visser to ask:

- **Visser's Question**: what is the *interpretability*[16] *relationship* between ACA_0 and VA?

In this section we partially answer Visser's question by establishing that ACA_0 can be faithfully interpreted within VA. Since the proofs of Theorem 1.1(b) and Theorem A are both formalizable within ACA_0, and ACA_0 is a conservative extension of PA for arithmetical sentences, the statement "VA is equiconsistent with PA" is provable within PA (see Remark 3.9.3 for a refinement). As we shall see, an analysis of the proof of Theorem A yields an interpretation δ of ACA_0 within VA, and Theorem B will show that δ is indeed a faithful interpretation.

THEOREM 3.9. *There is a faithful interpretation δ of ACA_0 within VA.*

PROOF: Suppose $(\mathfrak{N}, j)$ is a model of VA. Let M be the fixed point set of j, and $\mathcal{A} := SSy_M(\mathfrak{N})$. By Theorem A, all axioms of PA are true in $\mathfrak{M}$. This can be syntactically reformulated by saying that if for each formula φ of $\mathcal{L}_A$, φ^M is the formula in $\mathcal{L}_A \cup \{j\}$ obtained by restricting all the quantifiers of φ to the $(\mathfrak{N}, j)$-definable cut M, then by Theorem 1.1(b), Theorem A, and the completeness theorem for first order logic:

$$\text{For all sentences } \varphi \text{ of } L_A,\ PA \vdash \varphi \text{ iff } VA \vdash \varphi^M.$$

This shows that the map $\varphi \longmapsto \varphi^M$ describes a *faithful* interpretation of PA within VA. In order to interpret ACA_0 within $(\mathfrak{N}, j)$ define an equivalence relation $\equiv$ by

$$a \equiv b \text{ iff } \forall x \forall y (M(x) \wedge M(y) \rightarrow (E(x,a) \leftrightarrow E(x,b)),$$

where $M(x)$ is the formula "$x = j(x)$" and $E(x,y)$ is Ackermann's $\in$. Note that

$$[(\mathfrak{N}, j) \models a \equiv b] \text{ iff } [a_E \cap M = b_E \cap M],$$

which shows that $\equiv$ interprets the equality relation among sets. Therefore, we can interpret the two-sorted model $(\mathfrak{M}, \mathcal{A}, \in, =_{\mathcal{A}})$ within $(\mathfrak{N}, j)$ by interpreting $\mathfrak{M}$ via $I(x)$, $\mathcal{A}$ via $N/\equiv$, and the membership relation $\in$ (between members of M, and members of $\mathcal{A}$), via $E(x,y)$. So, by Theorem A, ACA_0 is uniformly interpretable in every model of VA. In syntactical terms, this idea can be used to show:

PROPOSITION 3.9.1. *For every formula $\psi(v_1, \cdots, v_s, X_1, \cdots, X_t)$ in the language of second order arithmetic, whose first order free variables are $v_1, \cdots, v_s$, and whose second order free variables are $X_1, \cdots, X_t$, there is a formula*

$$\delta_\psi(x_1, \cdots, x_s, x_{s+1}, \cdots, x_{s+t})$$

in the language $\mathcal{L}_A \cup \{j\}$ such that the following are equivalent for all models $(\mathfrak{N}, j)$ of VA, all sequences $a_1, \cdots, a_s$ from M, $b_1, \cdots, b_t$ from N, and $S_1, \cdots, S_t$ from $SSy_M(\mathfrak{N})$ such that $S_i = (b_i)_E \cap M$ for $1 \leq i \leq t$, where M is the fixed point set of j:

(i) $(\mathfrak{M}, SSy_M(\mathfrak{N})) \models \psi(a_1/v_1, \cdots, a_s/v_s, S_1/X_1, \cdots, S_t/X_t)$.

---

[16] See Sections 1 and 2 of Visser's paper [V] in this volume for the precise definition of interpretability. The intuitive idea can be explained as follows: a theory T_1 formulated in a language $\mathcal{L}_1$, is interpretable in a theory T_2 formulated in a language $\mathcal{L}_2$, if there is a "well-behaved" function δ, which translates formulae ψ from $\mathcal{L}_1$ into formulae δ_ψ in $\mathcal{L}_2$ such that for all sentences ψ of $\mathcal{L}_1$,

$$T_1 \vdash \psi \text{ implies } T_2 \vdash \delta_\psi.$$

If, in additon, the converse of the above implication holds for all sentences ψ of $\mathcal{L}_1$, δ is said to be a *faithful* interpretation.

(ii) $(\mathfrak{N}, j) \vDash \delta_\psi(a_1/x_1, \cdots, a_s/x_s, b_1/x_{s+1}, \cdots, b_t/x_{s+t})$.

We can now use Theorem A, Theorem B, Proposition 3.9.1, and the completeness theorem of first order logic together to conclude that for all sentences ψ of second order arithmetic, $ACA_0 \vdash \psi$ iff $VA \vdash \delta_\psi$. Therefore, ACA_0 is *faithfully* interpretable in VA via the interpretation δ.
□

COROLLARY 3.9.2. *VA has superexponential speed-up over PA (assuming the consistency of PA). More specifically, for every natural number k there is a theorem φ_k of PA whose interpretation has a proof of length d_k within VA such that the shortest proof of φ_k within PA is longer than $Superexp(k, d_k)$.*

PROOF: This is a direct consequence of interpretability of ACA_0 within VA and the independently obtained results of Friedman and Pudlák on the speed-up of ACA_0 over PA. More specifically, let us write $T \vdash_{\leq k} \psi$ for "there is a proof of φ from T of length k", and $T \vdash_{>k} \psi$ for "$T \vdash \psi$ and all proofs of φ from T are longer than k". Given a sentence φ in the language of Peano arithmetic, let $\overline{\varphi}$ be the canonical interpretation of φ within ACA_0. As shown by Friedman ([Fr], [Sm]) and Pudlák[17] [Pu-2, Corollary 4.5]:

(1) There is a sequence $\langle \varphi_k : k \in \omega \rangle$ of theorems of PA and an increasing sequence $\langle d_k : k \in \omega \rangle$ of natural numbers such that for all $k \in \omega$:

$$ACA_0 \vdash_{\leq d_k} \overline{\varphi_k}, \text{ but } PA \vdash_{>Superexp(k,d_k)} \varphi_k.$$

On the other hand, ACA_0 is finitely axiomatizable[18] and therefore there is a single theorem τ of ACA_0 with the same set of consequences as ACA_0 itself. Since VA interprets ACA_0 via δ of Theorem 3.9, $VA \vdash_{\leq c} \delta_\tau$ for some c. Therefore, for all sentences ψ in the language of second order arithmetic,

(2) $\tau \vdash_{\leq k} \psi \rightarrow VA \vdash_{\leq c+k} \delta_\psi$.

This is easy to see: if $\langle \varphi_n : 1 \leq n \leq k \rangle$ is a Hilbert-style proof of ψ from τ (so $\varphi_k = \psi$), then we can obtain a proof of δ_ψ from VA of length $k + c$ by first proving δ_τ in c-steps from VA, and then following the resulting proof with $\langle \delta_{\varphi_n} : 1 \leq n \leq k \rangle$. The result now easily follows from coupling (1) and (2).
□

REMARK 3.9.3. The proof of Theorem 3.9 can be used to show that $I\Delta_0 + Exp$ proves $Con(VA) \rightarrow Con(ACA_0)$, and therefore

$$I\Delta_0 + Exp \vdash Con(VA) \rightarrow Con(PA).$$

Coupled with $I\Delta_0 + Exp \nvdash Con(PA) \rightarrow Con(ACA_0)$ ([Pu-2], [Fr]), this shows that

$$I\Delta_0 + Exp \nvdash Con(PA) \rightarrow Con(VA).$$

[17] The exposition in [Pu-2] is geared toward the speed-up of GB (Gödel-Bernays theory of classes) over ZF. It is well-known that the same machinery can be used to show the speed-up of ACA_0 over PA.

[18] See [HP, Ch.III, Sec.1(b)] or [Si-2, Lemma VIII.1.5].

4. AUTOMORPHISMS AND SECOND ORDER ARITHMETIC

In the previous section we saw that there is a close relationship between models of PA and ACA_0 and fixed point sets of automorphisms of models $\mathfrak{N}$ of $I\Delta_0$. In this section we pursue this theme by investigating a minimal condition (M-amenability) under which the fixed point sets of automorphisms of bounded arithmetic give rise to models of *full second order arithmetic* Z_2.

4.1. Amenable Automorphisms from $\mathbf{Z_2}$

The following definition is suggested by the work of Solovay on automorphisms of models of set theory with a weakly compact cardinal [Sol, Section 3.5, Criteria 1 and 2].

- Suppose $\mathfrak{N}$ is a model of $I\Delta_0$, and M is a cut of $\mathfrak{N}$. An automorphism j of $\mathfrak{N}$ is *M-amenable* if the fixed point set of j is precisely M, and for every formula $\varphi(x, j)$ in the language $\mathcal{L}_A \cup \{j\}$, possibly with suppressed parameters from N,

$$\{m \in M : (\mathfrak{N}, j) \vDash \varphi(m, j)\} \in SSy_M(\mathfrak{N}).$$

THEOREM C. *Suppose $(\mathfrak{M}, \mathcal{A})$ is a countable model of Z_2+Π^1_∞-DC. There exists an e.e.e. $\mathfrak{N}$ of $\mathfrak{M}$ that has an M-amenable automorphism j such that $SSy_M(\mathfrak{N}) = \mathcal{A}$.*

PROOF: Before beginning the proof, recall that Π^1_∞-DC is the scheme in the language of second order arithmetic consisting of formulas of the form

$$\forall n \forall X \exists Y \theta(n, X, Y) \rightarrow [\forall X \exists Z (X = (Z)_0 \text{ and } \forall n\ \theta(n, (Z)_n, (Z)_{n+1})],$$

where φ is allowed to have number or set parameters, and $(Z)_n = \{i : \Gamma(i, n) \in Z\}$, where Γ is a canonical pairing function. See [Si-2, Sec.VII.6] for more on choice schemes in second order arithmetic[19].

The proof of Theorem C has two distinct stages. In the first stage, a well-behaved Ramsey ultrafilter $\mathcal{U}$ is constructed by forcing, while in the second stage, an internal iterated ultrapower modulo $\mathcal{U}$ is used to exhibit the desired model $\mathfrak{N}$ and the M-amenable automorphisms j of $\mathfrak{N}$.

Stage 1: Forcing a Ramsey ultrafilter

Forcing was used only as an efficient bookkeeping tool in Section 3.2. In contrast, here it is invoked in an essential manner to adjoin a generic ultrafilter to a model of second order arithmetic[20]. Our notion of forcing $\mathbb{P}$ (and therefore our notion of genericity) is the same as the one used already in Section 3.2, but in this section we shall invoke substantive properties of forcing to show that $\mathbb{P}$-forcing over a countable model $(\mathfrak{M}, \mathcal{A})$ of second order arithmetic with dependent choice produces a generic ultrafilter $\mathcal{U}$ such that the expansion $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$ continues to satisfy the

[19] N.B. the formulation of DC in [Si-2] is slightly different from the above, but equivalent.

[20] I am indebted to one of the referees for suggesting the self-contained approach for this stage. In the original proof of Theorem C, I used a forcing construction of Mostowski [Mo-1] to adjoin a global well-ordering $\lhd$ of $\mathcal{A}$ so that the comprehension scheme of Z_2 continues to hold even for formulas mentioning $\lhd$. It is then routine to define a Ramsey ultrafilter within $(\mathfrak{M}, \mathcal{A}, \lhd)$ by implementing the classical proof of the existence of a Ramsey ultrafilter using CH. Note that in his original paper [Mo-1], Mostowski claimed that his forcing construction works for countable models of Z_2 with the *choice scheme*. However, as observed by Simpson [Si-1], Mostowski's proof relies on the stronger scheme of *dependent choice*. This is acknowledged in [Mo-2].

comprehension schema in the language of second order arithmetic for formulae that refer to $\mathcal{U}$. To verify this, we begin with some definitions.

- Let $\mathcal{L}_2(U)$ be the result of augmenting the language of second order arithmetic $\mathcal{L}_2$ with a new predicate U with the understanding that U is a *predicate of sets*, i.e., models of $\mathcal{L}_2(U)$ are of the form $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$ where $(\mathfrak{M}, \mathcal{A})$ is an $\mathcal{L}_2$-structure, and $\mathcal{U} \subseteq \mathcal{A}$.
- The *forcing language* Φ is obtained by augmenting $\mathcal{L}_2(U)$ with constant symbols for each element of $M \cup \mathcal{A}$.
- Recall from Section 3.2 that $\mathbb{P}$ is $\{X \in \mathcal{A} : X \text{ is unbounded in } M\}$, ordered under inclusion. The forcing relation is inductively defined as follows:
  (1) $X \Vdash (Y \in U)$ iff $X \subseteq Y$ (where $Y \in \mathcal{A}$); for all other atomic formulae φ, $X \Vdash \varphi$ iff φ holds in $(\mathfrak{M}, \mathcal{A})$.
  (2) $X \Vdash (\varphi_1 \vee \varphi_2)$ iff $X \Vdash \varphi_1$ or $X \Vdash \varphi_2$.
  (3) $X \Vdash (\neg\varphi)$ iff $\forall Y \subseteq X (Y \nVdash \varphi)$.
  (4) $X \Vdash (\exists x \varphi(x))$ iff for some $m \in M$ such that $X \Vdash \varphi(m)$.

The following lemma is standard and is stated without proof. Note that it holds for all $\mathcal{L}_2$-structures $(\mathfrak{M}, \mathcal{A})$.

LEMMA C.1.

(1) (Monotonicity) *If $X \Vdash \varphi$ and $Y \subseteq X$, then $Y \Vdash \varphi$.*
(2) (Definability) *For every formula $\varphi(v_1, \cdots, v_s, X_1, \cdots, X_t)$ of $\mathcal{L}_2(U)$, there is a formula $\mathtt{Force}_\varphi(X, v_1, \cdots, v_s, X_1, \cdots, X_t)$ of $\mathcal{L}_2(U)$ such that for every model $(\mathfrak{M}, \mathcal{A})$ of $\mathcal{L}_2$, every $X \in \mathbb{P}$, every $m_1, \cdots, m_s \in M$, and every $S_1, \cdots, S_t \in \mathcal{A}$, $X \Vdash \varphi(m_1, \cdots, m_s, S_1, \cdots, S_t)$ iff $(\mathfrak{M}, \mathcal{A}) \vDash \mathtt{Force}_\varphi(X, m_1, \cdots, m_s, S_1, \cdots, S_t)$.*
(3) (Truth-and-Forcing) *If $\mathcal{U}$ is $\mathbb{P}$-generic over $(\mathfrak{M}, \mathcal{A})$, then for every Φ-sentence φ, $(\mathfrak{M}, \mathcal{A}, \mathcal{U}) \vDash \varphi$ iff $X \Vdash \varphi$ for some $X \in \mathcal{U}$.*

LEMMA C.2. *Suppose X and Y are elements of $\mathbb{P}$ whose symmetric difference $X \Delta Y$ is finite in the sense of $\mathfrak{M}$. For any sentence φ of Φ, $X \Vdash \varphi$ iff $Y \Vdash \varphi$.*

PROOF: Recall from Theorem 3.4(a) that generic filters are $(\mathfrak{M}, \mathcal{A})$-complete. Also note that for any $(\mathfrak{M}, \mathcal{A})$-complete ultrafilter $\mathcal{U}$, $X \in \mathcal{U}$ iff $Y \in \mathcal{U}$. The result now easily follows from Truth-and-Forcing.
□

The next two results unveil the key properties of generic ultrafilters. From here on, we use the abbreviation $X \parallel \varphi$ for "$X \Vdash \varphi$ or $X \Vdash \neg\varphi$".

LEMMA C.3. *If $(\mathfrak{M}, \mathcal{A})$ is a model of $Z_2 + \Pi^1_\infty$-DC, then for any unary formula $\varphi(x)$ of Φ, the following set D_φ is dense in $\mathbb{P}$*

$$D_\varphi := \{Y \in \mathbb{P} : \forall m \in M \ (Y \parallel \varphi(m)\}.$$

PROOF: Let $\theta(X, Y, n)$ be the formula "$X \supseteq Y$ and $Y \parallel \varphi(n)$". It is easy to see that

$$(\mathfrak{M}, \mathcal{A}) \vDash \forall n \forall X \exists Y \theta(n, X, Y).$$

Given any $X \in \mathbb{P}$, by the dependent choice scheme there is some element of $\mathcal{A}$ that codes a sequence $\langle X_0, X_1, X_2, \cdots, X_m, \cdots \rangle_{m \in M}$ of elements of $\mathbb{P}$ such that (1) and (2) below hold in $(\mathfrak{M}, \mathcal{A})$.

(1) $X_0 := X$ and $\forall m \in M \ X_{m+1} \subseteq X_m$;
(2) $\forall n \theta(n, X_n, X_{n+1})$.

Next, construct Y by setting $Y := \{y_m : m \in M\} \in \mathbb{P}$, where y_m is defined within $(\mathfrak{M}, \mathcal{A})$ via the recursion:

- y_0 is the first element of X_0;
- y_{m+1} is the least member of X_m above $\{y_i : i \leq m\}$.

Clearly $Y \subseteq X$ and $Y \backslash X_m$ is $\mathfrak{M}$-finite for all $m \in M$. Therefore, since by Monotonicity, $Y \cap X_m \parallel \varphi(m)$ for all $m \in M$, by Lemma C.2, $Y \in D_\varphi$.
□

LEMMA C.4. *If $(\mathfrak{M}, \mathcal{A})$ is a model of $Z_2 + \Pi^1_\infty$-DC and $\mathcal{U}$ is $\mathbb{P}$-generic over $(\mathfrak{M}, \mathcal{A})$, then for any unary Φ-formula $\varphi(x)$,*

$$S_\varphi := \{m \in M : (\mathfrak{M}, \mathcal{A}, \mathcal{U}) \vDash \varphi(m)\} \in \mathcal{A}.$$

PROOF: By Lemma C.3 there is a condition $Y_0 \in \mathcal{U}$ such that for all $m \in M$, $Y_0 \parallel \varphi(m)$. It is routine to verify (using Truth-and-Forcing) that

$$\{m \in M : (\mathfrak{M}, \mathcal{A}, \mathcal{U}) \vDash \varphi(m)\} = \{m \in M : (\mathfrak{M}, \mathcal{A}, \mathcal{U}) \vDash \text{“}Y_0 \Vdash \varphi(m)\text{”}\}.$$

Therefore S_φ is the solution set of a unary formula $\mathcal{L}_2$-formula (by definability of the forcing relation), and therefore by the comprehension scheme, $S_\varphi \in \mathcal{A}$.
□

Stage 2: Internally building an iterated ultrapower

In this stage of the proof, we employ the machinery of iterated ultrapowers discussed in Section 3.2.2, except that the entire construction is carried out internally within $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$. To see how this works, consider a generic ultrafilter $\mathcal{U}$ over $(\mathfrak{M}, \mathcal{A})$. By Theorem 3.4, $\mathcal{U}$ is $(M, \mathcal{A})$-iterable. Moreover, in light of Remark 3.2.1 it is easy to see that $\mathcal{U}$ is also m-canonically Ramsey[21] over $(M, \mathcal{A})$ for all $m \in M$. Since the construction of the m-type Γ_m uses the ultrafilter $\mathcal{U}^m$, and m might be nonstandard, we need to overcome the following obstacle: $\mathcal{U}^n$ was defined by an *external* induction in Section 3.2.2 via equation (♣) for *standard natural numbers* n. Therefore, to define $\mathcal{U}^m$ for nonstandard m, we seem need to work within *third order* arithmetic in order to carry out the necessary recursion. However, in light of Remark 3.6, there is a way out: since $\mathcal{U}$ is m-Ramsey over $(\mathfrak{M}, \mathcal{A})$, we can use the following recursion-free definition of $\mathcal{U}^m$ within $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$:

$$\mathcal{U}^m := \{Y \subseteq M^m : \exists X \in \mathcal{U}\ [X]^m \subseteq Y\}.$$

Therefore for any linear order $\mathbb{L} \in \mathcal{A}$ we can define the *internally* iterated ultrapower $Ult^*_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S \in \mathcal{A}}$ by carrying out the construction of Section 3.3 entirely within $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$. Note that the key difference between the internal and the external iterated ultrapower is that the external iterated ultrapower can be viewed as a direct limit of models that result from iterating the ultrapower formation process finitely many times, while the internal iteration can be viewed as a direct limit of models that result from iterating the ultrapower formation process $\mathfrak{M}$-*finitely* many times. We can therefore choose $\mathbb{L} \in \mathcal{A}$ such that $\mathbb{L}$ has an automorphism $h \in \mathcal{A}$ with no fixed points (e.g., $\mathbb{L} =$ the ordered set of integers in the sense of $\mathfrak{M}$, and $h(n) = n + 1$). By minor variants of Theorems 3.7 and 3.8, there is an automorphism

$$j^*_h : Ult^*_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S \in \mathcal{A}} \to Ult^*_{\mathcal{U},\mathbb{L}}(\mathfrak{M}, S)_{S \in \mathcal{A}}$$

[21] $\mathcal{U}$ is m-canonically Ramsey over $(M, \mathcal{A})$, where $m \in M$, if for every $f : [M]^m \to M$ with $f \in A$, there is some $X \in \mathcal{U}$ that is f-canonical.

that is definable within $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$, and whose fixed point set is precisely M. Since j_h^* is outright definable in $(\mathfrak{M}, \mathcal{A}, \mathcal{U})$, by Lemma C.4 j is M-amenable. This concludes the proof of Theorem C.
□

4.2. $\mathbf{Z_2}$ from Amenable Automorphisms

We now show that the full strength of second order arithmetic is needed in the proof of Theorem C.

THEOREM D. *If $\mathfrak{M} \vDash I\Delta_0$ and $\mathfrak{N}$ is an end extension of $\mathfrak{M}$ satisfying $I\Delta_0$ such that $\mathfrak{N}$ has an M-amenable automorphism, then $(\mathfrak{M}, SSy_M(\mathfrak{N})) \vDash Z_2$.*

PROOF: By Theorem A, $(\mathfrak{M}, SSy_M(\mathfrak{N})) \vDash ACA_0$. Therefore, we only need to verify the comprehension scheme of Z_2. Recall the mapping $\psi \longmapsto \delta_\psi$ of formulas of second order arithmetic to formulas of $\mathcal{L}_A \cup \{j\}$ of Proposition 3.9.1 in the proof of Theorem 3.9. If $\psi(x)$ is a unary formula of second order arithmetic (possibly with suppressed set or number parameters), then by Proposition 3.9.1

$$\{a \in M : (\mathfrak{M}, SSy_M(\mathfrak{N})) \vDash \psi(a)\} = \{a \in M : (\mathfrak{N}, j) \vDash \delta_\psi(a)\}.$$

Coupling this with the M-amenability of j, it now becomes evident that $(\mathfrak{M}, SSy_M(\mathfrak{N}))$ satisfies the comprehension scheme.
□

Let T^* be the extension of the theory T of Section 3.3 obtained by adding a *scheme* asserting that j is an M-amenable automorphism (where M as usual is the fixed point set of j). The proof of Theorem C, coupled with the well-known fact that the theory $Z_2 + \Pi^1_\infty$-DC can be interpreted within Z_2 via the "ramified analytical hierarchy" [Si-2] shows that T^* can be interpreted within Z_2. Furthermore, Proposition 3.9.1 and Theorem D together show that Z_2 is interpretable within T^*. Hence:

THEOREM 4.1. *The theories Z_2 and T^* are equiconsistent.*

5. FURTHER RESULTS AND OPEN QUESTIONS

5.1. Consequences for NFU

As mentioned in the introduction, the main results of this paper were obtained by the author in the context of the metamathematical study of certain extensions of the theory NFU, where NFU is Jensen's variant [Jen] of Quine's system of set theory *New Foundations* NF [Q]. NFU is obtained from NF by relaxing the extensionality axiom in order to allow urelements. The consistency of NF relative to any ZF-style set theory remains an open problem, but Jensen showed the consistency of NFU relative to a fragment of ZF-set theory. Theorems A, B, C, and D have been used in the joint work of Robert Solovay and the author to establish the results reported in this section. Here we only briefly define the concepts needed to state our results, and refer the reader to [Fo] or [Ho-1] for detailed background information and references.

- X is *Cantorian* if there is a one-to-one correspondence between X and the set of its singletons $\{\{v\} : v \in X\}$;
- X is *strongly Cantorian* if the map sending v to $\{v\}$ (as v varies in X) exists;

- $NFU^{-\infty}$ is NFU plus the axiom "every set is finite";
- $NFUA^{-\infty}$ is $NFU^{-\infty}$ plus the axiom "every Cantorian set is strongly Cantorian"; and
- $NFUB^{-\infty}$ is the extension of $NFUA^{-\infty}$ obtained by adding a scheme asserting that the intersection of any parametrically definable class with the class of Cantorian sets is the result of the intersection of the extension of some element with the class of Cantorian sets.

Of course, in ZF-style set theories every set is strongly Cantorian, but in NF and NFU this is no longer true, e.g., the universal set of a model of NF or NFU is not even Cantorian, and there are models of NFU + "there is an infinite set" + the axiom of choice in which the set of finite cardinals is Cantorian, but not strongly Cantorian. We are now ready to state the ramifications of Theorems A and B for NFU:

THEOREM 5.1. *The following are equivalent for complete theories T in the language $\mathcal{L}_A$ of arithmetic:*

(a) *There is a model of $NFUA^{-\infty}$ whose class of Cantorian cardinals satisfies T.*
(b) *T is an extension of PA.*

COROLLARY 5.1.1. *$NFUA^{-\infty}$ is equiconsistent with PA.*

Furthermore, Theorem 4.1 can be used to show:

THEOREM 5.2. *$NFUB^{-\infty}$ is equiconsistent with Z_2.*

5.2. A Characterization of $I\Delta_0 + B\Sigma_1 + Exp$

In recent work [E-2], the author has established the following characterization of the fragment $I\Delta_0 + B\Sigma_1 + Exp$ of PA in terms of automorphisms. Here $B\Sigma_1$ is the scheme consisting of the universal closure of formulae of the form

$$[\forall x < a \exists y\ \varphi(x,y)] \rightarrow [\exists z \forall x < a \exists y < z\ \varphi(x,y)].$$

In what follows $I_{fix}(j)$ denotes the largest initial segment of a model $\mathfrak{N}$ of $I\Delta_0$ that is pointwise fixed under an automorphism j of $\mathfrak{N}$.

THEOREM 5.3.

(a) *Suppose $\mathfrak{M}$ is a countable model of $I\Delta_0 + B\Sigma_1 + Exp$. $\mathfrak{M}$ has a proper end extension to a model $\mathfrak{N}$ of $I\Delta_0$ such that for some automorphism j of $\mathfrak{N}$, $I_{fix}(j) = M$.*
(b) *If j is a nontrivial automorphism of some model $\mathfrak{N}$ of $I\Delta_0$, then $I_{fix}(j)$ is a model of $I\Delta_0 + B\Sigma_1 + Exp$.*

COROLLARY 5.3.1. *$I\Delta_0 + B\Sigma_1 + Exp$ is the theory of the class of models whose universes are of the form $I_{fix}(j)$ for some nontrivial automorphism j of a model of $I\Delta_0$.*

5.3. Open Questions

- QUESTION 1. Let VA be the theory discussed in Section 3.3. Can VA be interpreted in ACA_0?
- QUESTION 2. Can Theorem D be strengthened by including the clause "$(\mathfrak{M}, SSy_M(\mathfrak{N}))$ satisfies Π^1_∞-DC" in the conclusion?
- QUESTION 3. Besides PA, Z_2, and $I\Delta_0 + B\Sigma_1 + Exp$, are there other arithmetical theories that can be naturally characterized in terms of automorphisms?

REFERENCES

[AH] F. Abramson and L. Harrington, *Models without indiscernibles*, **Journal of Symbolic Logic,** vol. 43 (1978), no. 3, 572–600.

[Av] J. Avigad, *Number Theory and elementary arithmetic*, **Philosophia Mathematica**, vol. 11, issue 3 (2002), pp 257-284.

[Be] J. H. Bennett, **On Spectra,** Ph.D. dissertation, Princeton University, 1962.

[Bu] S. Buss, *Proof theory of arithmetic*, in **Handbook of Proof Theory** (S. Buss, ed.), North-Holland, Amsterdam, 1998.

[CK] C. C. Chang and H. J. Keisler, **Model Theory**, Elsevier North Holland, Amsterdam,1973.

[D] P. D'Aquino, *A sharpened version of McAloon's theorem on initial segments of models of* $I\Delta_0$, **Annals of Pure and Applied Logic**, vol. 61 (1993), pp. 49-62.

[DG] C. Dimitracopoulos and H. Gaifman, *Fragments of Peano's Arithmetic and the MRDP theorem*, in **Logic and Algorithmic**, Monogr. Enseign. Math. University of Geneva, 1982, pp. 187-206.

[E-1] A. Enayat, *Automorphisms, Mahlo cardinals, and NFU*, in **Nonstandard Models of Arithmetic and Set Theory** (A. Enayat and R. Kossak ed.), Contemporary Mathematics Series, vol. 361, American Mathematical Society, 2004.

[E-2] ———, *Automorphisms of models of bounded arithmetic*, to appear.

[E-3] ———, *Weakly compact cardinals and automorphisms*, to appear.

[ER] P. Erdös and R. Rado, *A combinatorial theorem*, **Journal of the London Mathematical Society**, vol. 25 (1950), pp. 249-255.

[Fe] U. Felgner, *Comparisons of the axioms of local and global choice*, **Fundamenta Mathematicae** vol. 71 (1971), pp. 43-62.

[Fo] Forster, **Set Theory with a Universal Set**, Second ed., Oxford Logic Guides, vol. 31, Oxford University Press, 1995.

[Fr] H. Friedman, *Translatability and relative consistency, II.* Ohio State University, unpublished notes, 1979.

[G] H. Gaifman, *Models and types of arithmetic,* **Annals of Mathematical Logic**, vol. 9 (1976), pp. 223-306.

[GRS] R. Graham, B. Rothschild, and J. Spencer, **Ramsey Theory**, Wiley-Interscience publications, New York, 1980.

[HP] P. Hájek and P. Pudlák, **Metamathematics of First Order Arithmetic**, Springer, Heidelberg 1993.

[Ho-1] R. Holmes, *Strong axioms of infinity in NFU,* **Journal of Symbolic Logic**, vol. 66 (2001), pp. 87-116.

[Ho-2] _____, *Errata to "Strong axioms of infinity in NFU",* **Journal of Symbolic Logic**, vol. 66 (2001), pp. 1974.

[Jec] T. Jech, **Set Theory**, Academic Press, New York, 1978.

[Jen] R. B. Jensen, *On the consistency of a slight* (?) *modification of Quine's New Foundations,* **Synthese**, vol. 19 (1969), pp. 250-263.

[Jo] C. Jockusch, *Ramsey's theorem and recursion theory*, **Journal of Symbolic Logic,** vol. 37 (1972), pp.268-280.

[Kan] A. Kanamori, **The Higher Infinite**, Springer-Verlag, Heidelberg (1994).

[Kay] R. Kaye, *Model-theoretic properties characterizing Peano arithmetic*, **Journal of Symbolic Logic**, vol. 56 , pp. 949-963 (1991).

[KKK] R. Kaye, R. Kossak, and H. Kotlarski, *Automorphisms of recursively saturated models of arithmetic*, **Annals of Pure and Applied Logic**, vol. 55 (1991), pp. 67-99.

[KM] R. Kaye and D. MacPherson, **Automorphisms of First-Order Structures**, Oxford University Press, Oxford, 1994.

[Ki-1] L. Kirby, **Initial segments of models of arithmetic**, Ph. D. thesis, University of Manchester, 1977.

[Ki-2] ———, *Ultrafilters and types on models of arithmetic*, **Annals of Pure and Applied Logic**, vol. 27 (1984). pp. 215-252.

[KP] L. Kirby and J. Paris, *Initial segments of models of Peano's axioms*, in **Lecture Notes in Mathematics**, vol. 619, Springer-Verlag, 1977, pp. 211-226.

[Kr] J. Krajíček, **Bounded Arithmetic, Propositional Logic, and Complexity Theory**, Cambridge University Press, Cambridge, 1995.

[Ku] K. Kunen, *Some applications of iterated ultrapowers in set theory*, **Annals of Mathematical Logic** vol. 1 (1970), pp. 179-227.

[Mile] J. Mileti, *The Canonical Ramsey Theorem and Computability Theory,* to appear.

[Mill] G. Mills, *A model with no elementary end extension*, **Journal of Symbolic Logic**, vol. 43 , pp. 563–567 (1978).

[Mo-1] A. Mostowski, *Models of second order arithmetic with definable Skolem fuctions*, **Fundamenta Mathematicae**, vol. 75 (1972), pp. 223-234.

[Mo-2] ———, *Errata to the paper "Models of second order arithmetic with definable Skolem fuctions"*, **Fundamenta Mathematicae**, vol. 84 (1974), pp. 173.

[MS] R. MacDowell and E. Specker, *Modelle der Arithmetik*, In **Infinitistic Methods**, Proc. Symp. Found. Math. (Warsaw, 1959), Pergamon (1961), pp. 257-263.

[Pa] R. Parikh, *Existence and feasibility in arithmetic*, **Journal of Symbolic Logic**, vol. 36 (1971), pp. 494-508.

[Pu-1] P. Pudlák, *A definition of exponentiation by a bounded arithmetical formula*, **Comm. Math. Univ. Carol.** vol. 24 (1983), pp. 667-671.

[Pu-2] ———, *Cuts, consistency statements, and interpretations*, **Journal of Symbolic Logic**, vol. 50 (1985)), pp. 423-441.

[Q] W.V.O. Quine, *New foundations for mathematical logic*, **American Mathematical Monthly**, vol. 44 (1937), pp. 70-80.

[Ra] R. Rado, *Note on canonical partitions*, **Bulletin of London Mathematical Society**, vol. 18 (1986), pp. 123-126.

[Re] J. P. Ressayre, *Nonstandard Universes with strong embeddings*, in **Contemporary Mathematics** (S. Simpson, ed.), vol. 65, American Mathematical Society, 1987.

[Sc] J. Schmerl, *Automorphism groups of models of Peano arithmetic*, **Journal of Symbolic Logic** vol. 67, pp. 1249–1264 (2002).

[Si-1] S. Simpson, Review of [Fe] and [Mo-1], **Journal of Symbolic Logic**, vol. 38, pp.652-653.

[Si-2] ———, **Subsystems of Second Order Arithmetic**, Springer, Heidelberg 1999.

[Sm] C. Smorynski, *Nonstandard models and related developments*, in **Harvey Friedman's Research on the Foundations of Mathematics** (edited by L.A. Harrington et al.), North-Holland, Amsterdam, 1980.

[Sol] R. Solovay, *The consistency strength of NFUB,* preprint available at Front for the Mathematics ArXiv, http://front.math.ucdavis.edu

[Som-1] R. Sommer, **Transfinite Induction and Hierarchies Generated by Transfinite Recursion within Peano Arithmetic**, Ph.D. thesis, University of California, Berkeley, 1990.

[Som-2] ———, *Transfinite induction within Peano arithemtic*, **Annals of Pure and Applied Logic**, vol. 76 (1995), pp. 231-289.

[V] A. Visser, *Categories of theories and interpretations*, THIS VOLUME.

[W] H. Wang, **Popular Lectures on Mathematical Logic**, Dover Publications, Mineola (1993).

Department of Mathematics and Statistics, American University, 4400 Mass. Ave., N.W., Washington, D.C. 20016-8050, e-mail: enayat@american.edu