On groups and initial segments in nonstandard models of Peano Arithmetic

by

John David Allsup

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy in the Faculty of Sciences

School of Mathematics

The University of Birmingham

Edgbaston

Birmingham

B15 2TT

27th February, 2007

Abstract

This thesis concerns $\mathfrak{M}$-finite groups and a notion of discrete measure in models of Peano Arithmetic.

First we look at a measure construction for arbitrary non-$\mathfrak{M}$-finite sets via suprema and infima of appropriate $\mathfrak{M}$-finite sets. The basic properties of the measures are covered, along with non-measurable sets and the use of end-extensions.

Next we look at nonstandard finite permutations, introducing nonstandard symmetric and alternating groups. We show that the standard cut being strong is necessary and sufficient for coding of the cycle shape in the standard system to be equivalent to the cycle being contained within the external normal closure of the nonstandard symmetric group.

Subsequently the normal subgroup structure of nonstandard symmetric and alternating groups is given as a result analogous to the result of Baer, Schreier and Ulam for infinite symmetric groups.

The external structure of nonstandard cyclic groups of prime order is identified as that of infinite dimensional rational vector spaces and the normal subgroup structure of nonstandard projective special linear groups is given for models elementarily extending the standard model.

Finally we discuss some applications of our measure to nonstandard finite groups.

Acknowledgements

First and foremost, I am truly indebted to my supervisor, Dr Richard Kaye, for his time, oversight, and patience as things dragged on to their eventual conclusion.

Secondly, I wish to thank the School of Mathematics for their support, both financial and otherwise, during my postgraduate studies.

Thirdly, I must thank my family for their generous support without which this thesis would not have been possible.

Fourthly, I would like to thank the examiners Professor Dugald Macpherson and Mr Brian Philp for reading this thesis and supplying comments.

Finally, many thanks must go to all the other postgraduates of the school for being around.

Contents

Chapter 1

Preliminaries

1.1 Introduction

Numbers capture the common intuition of size, groups do likewise for symmetry. Nonstandard models of arithmetic provide an interesting setting in which to explore these notions. This thesis presents an attempt to bring these three themes together: measure, groups and nonstandard models.

The main theme is the investigation of the notion of nonstandard finite groups constructed within nonstandard models of Peano arithmetic considered as groups in their own right. Subordinate to this are a number of other themes, some more related to this main theme than others. The first of these looks at a generalisation of the idea of counting finite sets, generalising so as to measure the size of arbitrary subsets of a nonstandard model, following up an idea of Kaye[23, 22]. Next we proceed to look at nonstandard permutations and basic properties of nonstandard groups. We then come to the main chapters of the thesis, looking at the normal subgroup structure of some groups which are simple within their parent model, yet lose simplicity when looked at through the eyes of the metatheory. Finally we collect together other related aspects of the author's research before concluding the work in the traditional way.

The target audience consists of mathematicians and mathematical logicians with a background in both logic and reasonably elementary group theory. The reader will need a basic familiarity with models of Peano arithmetic, for which Kaye's book[21] is more than enough and elementary techniques of logic and model theory, for which the contents of an introductory course in logic together with Chang and Keisler[11] will suffice. Also required are a basic familiarity with permutations, permutation groups and the classical groups.

The main aim is to look at nonstandard analogues of well known families of finite groups as infinite groups in their own right. Of particular interest is the normal subgroup structure of these groups.

We begin our investigation into nonstandard groups with a discussion of nonstandard finite permutations on a nonstandard finite set X and their relation to external permutations on X. With this in hand, the opportunity for a preliminary discussion of symmetric and alternating groups presents itself, and these topics are covered in chapter 4.

Following on from this, in chapter 5, we give a complete classification of the normal subgroup structure of nonstandard symmetric and alternating groups for arbitrary models of Peano arithmetic. The observation that these normal subgroups are linearly ordered in the case of nonstandard alternating groups and nearly linearly ordered for nonstandard symmetric groups gives rise to obvious questions regarding the normal subgroup structure of other nonstandard analogues of well known finite groups. In particular we begin to focus on analogues to the non-alternating nonabelian finite simple groups, namely the families of simple groups of Lie type, for coverage of which the reader may refer to Carter[10]. From the Classification of finite simple groups, we know all the families of finite simple groups, and we do not address the question of whether other nonstandard finite simple groups exist outside of these families. Thus we may approach the question of the normal subgroup structure of these groups by approaching each family of finite simple groups. The cyclic groups of prime order present no difficulty and we proceed to apply results of Lev to classify the normal subgroups of nonstandard projective special linear groups, and give a brief look at the nonstandard projective symplectic case.

Throughout the results concerning these nonstandard groups, initial segments (which here mean convex subsets of a model $\mathfrak{M}$ containing 0) play a central rôle, capturing as they do, a certain intuition of size. Thus in our first nonintroductory chapter we consider a 'discrete measure' of subsets of $\mathfrak{M}$, using definable sets in much the same way that Lebesgue measure uses unions of open intervals. We comment that this measuring concept has much in common with Lebesgue measure on the real line, in particular with the phenomemon of non-measurable and universally non-measurable sets. In a later chapter we show how the ideas introduced may be applied in the context of nonstandard finite groups, in particular to properties such as indices of one group inside another. There are also analogues between these measures and the Loeb measure construction, which in particular allows the Lebesgue measure on the real line to be recovered from a discrete hyperfinite measure, though we leave such matters out of the material under discussion.

Nonstandard models of arithmetic and nonstandard analysis

The study of nonstandard models of arithmetic, and hence the use of nonstandard methods, began in the 1930s with the work of Skolem. The mindset of nonstandard methods gave rise to a markedly different view of the foundations of analysis with the work of Robinson[34], making rigorous the use of actual infinitesimals that was prevalent in Newton and Leibnitz's early development of calculus (see also Hurd and Loeb[20]).

One of the overtones of the themes in this thesis is that of taking the point of view and mindset of nonstandard analysis and applying it within the context of nonstandard models of arithmetic—in other words, considerations of internal and external properties of objects and the use of first order formulae and first order results about the standard model (and structures interpreted within it) in order to transfer some results to other first order models of arithmetic.

Whilst much of the literature of nonstandard models of arithmetic concerns various notions of provability and definability, some of the more recent trends have tended to think about actually *doing* things with and within models of arithmetic. This thesis falls into the latter category and to reiterate, the attitude behind much of what we do is found more in nonstandard *analysis* than in nonstandard models of arithmetic. Thus the tendency is to 'do maths' using nonstandard methods and combinatorial ideas marshalled through them.

As such then, the nomenclature employed at different stages will draw from many camps. For example, in nonstandard analysis we have internal objects in a nonstandard model ${}^*\mathbb{R}$ of the reals that are the $*$-transform of corresponding objects in the standard reals $\mathbb{R}$, other objects are then called external (Lindstrøm[29] is another point of reference for this). Analogous to internal objects are the $\mathfrak{M}$-finite objects in a nonstandard model of arithmetic, which are in some way bounded and definable with parameters (and can be represented by individual elements of the model via some suitable method of coding.) At times, then, it will be convenient to talk in terms of internal vs. external, and in others it will be convenient to talk of $\mathfrak{M}$-finite, coding, etc. We shall try to gravitate towards nonstandard models of Peano arithmetic, but in any case a basic familiarity with the background material should be enough to prevent confusion arising.

Infinite groups

Examples of infinite groups pervade mathematics at almost every level. They range in familiarity from structures such as $\mathbb{Z}$ and $\mathbb{R}$, with which the average mathematics student is well acquaited long before he is ever told what a group is, through to the numerous esoteric examples present in the modern day literature.

To study the totality of infinite groups is impractical, and yields very little in the way of uniformly applicable machinery: the infinite groups are too numerous, and only too willing to conspire in the creation of nasty counterexamples unless various assumptions and restrictions are made in order to rein in their behaviour. As is usual, mathematicians approach this totality with a divide and conquer methodology, defining useful classes of infinite groups, studying these classes, often classifying their constituents further.

Among the classes of infinite groups studied in the literature, many arise from 'finiteness' conditions of one form or another. For example, there are the locally finite groups (a group being locally finite if every element lies in a finite subgroup), finitary permutation groups (whose elements move only a finite number of elements), cofinitary permutation groups (whose nonidentity elements fix finitely many points) and many more.

One of the main subjects of this thesis is the notion of nonstandard finite groups. That is, infinite groups that are finite in the sense of some model $\mathfrak{M}$ of Peano arithmetic. To the reader familiar with the nomenclature of nonstandard analysis, 'hyperfinite' may be a more descriptive term.

The question of which groups occur as nonstandard finite groups is natural to ask, though we do not go into depth on this topic, preferring to look at analogues of well known families of finite groups.

Different perspectives on familiar families of groups

Different perspectives tend to emphasise different aspects of a given area of study, and nonstandard analogues of finite groups are no different in this regard.

Consideration of the normal subgroup structure of nonstandard groups leads to questions regarding covering normal closures of subgroups by conjugacy classes and this in turn gives a concrete realisation of the effects of results concerning such covering.

Along different lines, but with similar tendencies, are ultraproducts of finite groups. These yield models of the theory of finite groups, and by selecting the family of groups constituting the ultraproduct, various aspects of finite groups can be both emphasised and hidden. Of course, a well known construction of a nonstandard model of the reals is via an ultrapower construction—in fact this is often given as *the* construction of the nonstandard reals for those that aren't too interested in logic (*e.g.* as in Lindstrøm[29]). In addition, Skolem's original construction of a nonstandard model is a reduced product that is similar in many ways to an ultrapower construction.

As we have noted earlier, provability and its related notions are not the main themes of this thesis, our interest being mainly in *doing* things inside nonstandard models. That said, from the considerations of nonstandard groups and families of groups we speak of, come good intuitive pictures of what is needed in certain proofs. A good example here is the result of Baer, Schreier and Ulam classifying the normal subgroups of infinite symmetric groups. Furthermore, taking the product of infinite sequences that are finite in the sense of a given model allows us to consider, from the external point of view, closure properties of subgroups of nonstandard groups that are slightly stronger than the usual properties of the groups from an external point of view, whilst not as strong as the usual group closure in the sense of the model.

A note on consistency of notation

From time to time along the duration of this thesis, certain notational devices will be used and reused for different purposes. The author has taken the trouble to ensure that otherwise ambiguous notations are separated by either the context within which they appear, or else are separated geographically and introduced separately in the context in which they are used. Much of this is due to reasons of practicality and the aesthetic desires of the author.

Mathematics is, in a certain sense, an abbreviated and condensed form of written language which has evolved through the ages for the purpose of communicating abstract ideas. In this sense, it has much in common with music, though the ideas communicated are different in nature. Furthermore, the need for aesthetic beauty is not, and should not be diminished by the scientific appearance of proceedings: good mathematics is an art. Thus in our exposition we aim to present interesting, attractive mathematics, resorting to rigid, complex and ultimately unattractive notational entities only where it aids communication or else is rendered necessary by the nature of the material being discussed. The author begs the reader's forgiveness for any lack of artistic talent in this regard.

1.2 Literature survey

There is much in the literature under the heading of models of arithmetic, and much more under the heading of group theory. The literature survey of this section thus aims only to place the current work amongst that which already exists.

A brief history of Peano arithmetic

The subject of models of Peano arithmetic began in 1933 with the discovery by Skolem of nonstandard models of Peano's axioms. Early on, much interest lay in those models which elementarily extended the standard model, so called 'strong' nonstandard models. Later it was found that Peano arithmetic was strong enough as a theory in its own right, in particular being able to prove various coding properties. The reader may refer to Smorynski[38] for a more thorough history of the subject.

Coding of sequences of elements of a model of arithmetic first appeared in the work of Gödel in the form of the following lemma. (The form of the lemma quoted here is taken from Kaye[21].)

Gödel's Lemma. *There is a Δ_0 formula $\beta(x,y,z)$ such that*

$$\mathbb{N} \models \forall xy \exists! z \beta(x,y,z)$$

and for all $k \in \mathbb{N}$ and all $z_0, z_1, \ldots, z_{k-1} \in \mathbb{N}$, there is $x \in \mathbb{N}$ such that

$$\mathbb{N} \models \beta(x,i,z_i)$$

for all $i < k$.

There are many formulae which will suffice to prove this result. Our method, as Gödel's did, relies on the Chinese remainder theorem. Coding of sets follows easily from that of sequences, though this can also be accomplished in other ways. The importance of coding of sets arose with, amongst other things, the significance of the standard system of a model.

Work on initial segments began with Friedman and continued with the work of Paris and Kirby[25]. Their development of indicators lead to the discovery of independence results of 'mathematical interest' such as the celebrated Paris-Harrington theorem[33].

The investigation by Paris and Kirby[25] of analogies between initial segments and cardinals, in particular regular cardinals, led to their development of the notions of semi-regular, regular and strong initial segments of a model. Of these strong initial segments have turned out to be of interest in the modern literature, for example in their connections with arithmetical saturation. It was shown by Paris and Kirby that strong initial segments of models of Peano arithmetic are themselves models of Peano arithmetic, though it is known that the converse of this is false.

More recently, the study of recursively and arithmetically saturated models of Peano arithmetic has gained popularity. Recursive saturation arose with the work of Barwise and Schlipf[5] and, independently, Ressayre. Models of arithmetic were known to have a fair amount of saturation, though recursive saturation adds further richness to the models, especially when one is concerned with the automorphism groups of these structures. Smorynski's paper[37] gives a good overview of the basic results concerning recursively saturated models of arithmetic, namely embeddability criteria and uniqueness results. It is from ideas arising in later work on automorphism groups of recursively saturated models of Peano arithmetic that much of the inspiration for the ideas of measure explored in this thesis are drawn.

Connections between model theory and group theory

There are many different connections between model theory and group theory that have been explored in the literature. We take a brief look at some here.

The study of automorphisms of models provides a rich source of permutation groups, as covered in, for example, the collection *Automorphisms of First Order Structures* edited by Kaye and Macpherson[24].

Among other interesting connections between models and groups is the result (see *e.g.* Cameron[9]) that the theory of a model $\mathfrak{M}$ is ω-categorical iff $\mathrm{Aut}(\mathfrak{M})$ is oligomorphic, that is, iff for each $k \in \mathbb{N}$ there are only finitely many orbits of $\mathrm{Aut}(M)$ on the set of k-tuples of elements of $\mathfrak{M}$.

Closer to our area of study is the work of Kaye, Kossak and Kotlarski[23] on automorphisms of recursively saturated models of Peano arithmetic, in particular their work on the normal subgroup structure of these automorphism groups and their relationship with initial segments of models of Peano arithmetic. It is in this work, and a later follow-up article by Kaye[22] that the idea of initial segments as measure that we explore in chapter 3 emerges.

A notion drawn up in analogy to the pseudofinite *fields* introduced by Ax[3] is the idea of pseudofinite *groups*. Pseudofinite groups being infinite models of the theory of finite groups, an interesting question arises as to where and how much this notion coincides with the notion of $\mathfrak{M}$-finite groups explored in this thesis. Drawing on the classification and the theory of Chevalley groups, Wilson[41] shows that the simple pseudofinite groups are all elementarily equivalent to Chevalley groups over pseudofinite fields, leaving the question of whether this can be strengthened to isomorphism unanswered.

Finally, the work of Lascar on automorphisms of strongly minimal structures[27] is employed by Gardener[15] to give results pertaining to the normal subgroup structure of classical groups of countably infinite dimension. This thesis looks at the situation when the dimension is nonstandard rather than countably infinite in the sense of set theory.

Chapter 2

Background

The setting for this thesis is nonstandard models of Peano Arithmetic, with an emphasis on algebraic and combinatorial constructions within them. In this chapter we exposit and develop the necessary background material for the rest of the thesis. We assume that the reader is familiar with the requisite logic and set theory: only the basics are essential.

As for notational convention, logical symbols such as formal variables (such as v_0), predicates (for example Perm) and theories (like PA) will be set in $\mathsf{sans\ serif}$. Models will be set in gothic capitals. All else, such as informal variables, will be set in a serif typeface.

2.1 Model Theory

For a good overall view of the subject of Model Theory we refer the reader to Chang and Keisler[11] or Hodges[18]. Also of note is Hodges'[17] game theoretic slant on the construction of models.

We fix our precise choice of notation and definitions here. The notion of a model is standard within the literature and the reader may supply his own references, though our precise choice of notation is approximately that of Kaye and Macpherson[24].

Definition 2.1.1. A *model* with domain M is a tuple

$$\mathfrak{M} = (M\,;\ldots R_i\ldots\,;\ldots F_j\ldots\,;\ldots c_k\ldots)_{i\in I, j\in J, k\in K}$$

where $\mathscr{R} = \{R_i : i \in I\}$ is a family of n_i-ary relations on M, $\mathscr{F} = \{F_j : j \in J\}$ is a family of m_j-ary functions on M and $\mathscr{C} = \{c_k : k \in K\}$ is a family of constants $c_k \in M$.

We denote the cardinality of a model $\mathfrak{M}$ (in the sense of our metatheory or metauniverse) by $\operatorname{card}\mathfrak{M}$ and try to follow standard notation hereon. Each model is associated with a *formal language* that describes it. This language has symbols for

1. variables $\mathsf{v}_0, \mathsf{v}_1, \ldots$;
2. logical operations $=, \vee, \wedge, \rightarrow, \neg, \leftrightarrow$;
3. relation symbols R_i for each $i \in I$ representing an m_i-ary relation;
4. function symbols F_i for each $j \in J$ representing a n_j-ary function; and
5. constant symbols c_k for each $k \in K$.

The tuple (I, J, K, α, β), where $\alpha\colon i \mapsto m_i$ and $\beta\colon j \mapsto n_j$, is called the *similarity type* of the language, which we denote by

$$\mathscr{L}(\ldots\mathsf{R_i}\ldots;\ldots\mathsf{F_j}\ldots;\ldots\mathsf{c_k}\ldots)_{i\in I, j\in J, k\in K}$$

though in practice this convention is far less cumbersome than we end up writing here, for example the statement of the language of arithmetic becomes, for example, $\mathscr{L}_{\mathsf{PA}} = \mathscr{L}(<;\mathsf{S},+,\cdot;\mathsf{0},\mathsf{1})$.

Having stated this formally, we tend to relax things a great deal, for example blurring the distinction between relation symbols and the relations they represent. The reader should be able to work out what is meant, and formalise without difficulty to his or her heart's content.

2.2 Peano arithmetic

The subject of Peano Arithmetic is all about first order models for the language $\mathscr{L}_{\mathsf{PA}} = \mathscr{L}(<;\mathsf{S},+,\cdot;0,1)$. This is covered extensively in Kaye[21]. The natural numbers $\mathbb{N}$ provide us with the standard model for this language, and the subject investigates how it relates to other models for this language. The word 'true' in context of models of Peano Arithmetic will be taken to mean 'true in $\mathbb{N}$', this $\mathbb{N}$ capturing our intuitive notion of what the counting numbers are (with the proviso that 0 is one of them.) The reader may prefer to substitute this intuitive $\mathbb{N}$ with a concrete construction, it makes little difference to the content of this thesis.

2.3 Initial segments and ordering of models of Peano arithmetic

Models of PA have a natural ordering $<$ which is usually given explicitly as part of the structure. This immediately yields the notion of an initial segment of a model, which here are taken to be convex subsets containing 0. Initial segments of models of PA will figure greatly in this thesis, giving as they do a notion of size allowing one to divide the elements of a model into the 'small' ones and the 'large' ones. It is in this rôle that they figure the most, with the deeper relations with structure of a model taking a back seat in the following proceedings.

We use this order to define the following relation between subsets of a model $\mathfrak{M}$ of PA:

$$A \subseteq_e B \text{ iff } (\mathfrak{M}, A, B) \vDash \forall x{\in}A\, \forall y{\in}B\, (y < x \to y \in A)\,.$$

When $A \subseteq_e M$ we say that A is an *initial segment* of $\mathfrak{M}$. An initial segment of $\mathfrak{M}$ closed under successor is called a *cut*. When $\mathfrak{M}$ and $\mathfrak{N}$ are models of PA and $M \subseteq_e N$ then we say that $\mathfrak{N}$ is an *end-extension* of $\mathfrak{M}$ and write $\mathfrak{M} \subseteq_e \mathfrak{N}$. The careful reader should note that we effectively have two relations, one between subsets of a model and one between models, usually the one meant in this thesis is the relation between subsets. We shall often, especially during the material on $\mathfrak{M}$-measures, identify a with $\{x \in M \,:\, (\mathfrak{M}, a) \vDash x < a\}$, the latter being an initial segment of $\mathfrak{M}$. When we wish to name the set $\{x \in M \,:\, (\mathfrak{M}, a) \vDash x < a\}$ explicitly, we shall use the notation ${}_{<}a$.

The supremum and infimum of a subset of a model are central to our definitions of measure in the sequel. For these we have the definitions

$$\sup A = \{x \in M \,:\, (\mathfrak{M}, A) \vDash \exists y{\in}A\, x < y\} \text{ and}$$
$$\inf A = \{x \in M \,:\, (\mathfrak{M}, A) \vDash \forall y{\in}A\, x < y\}$$

following those of Kaye *et al.*[23], though we use $<$ rather than $\leq$ for consistency with the identification $a = \{x \in M \,:\, x < a\}$ assumed above. The reader should note that $\sup\varnothing = \varnothing$ and $\inf\varnothing = M$ with the above definitions.

Three interesting properties of initial segments, roughly analogous to the property of a cardinal being regular, were introduced by Kirby and Paris[25] and are the following. We say that a cut I of a model $\mathfrak{M}$ is *semi-regular* if for every $a \in M$ and every $\mathfrak{M}$-

finite[1] function $f\colon a \to M$, $f({<}a) \cap I$ is bounded in I; we call I *regular* if every $\mathfrak{M}$-finite function f whose domain includes I is constant on a cofinal subset of I; and we say that I is *strong* if given an $\mathfrak{M}$-finite function f, there is $a \in M \setminus I$ such that for all $x \in I$, $f(x) > I$ if, and only if, $f(x) > a$.

The need to express the notions: 'the largest initial segment closed under addition not containing a' and 'the smallest initial segment closed under addition containing a' arise often in this thesis. To this end we define

$$a \cdot \mathbb{N} = \sup\{an \,:\, n \in \mathbb{N}\},$$

that is, the smallest initial segment closed under addition containing a, and

$$a/\mathbb{N} = \inf\{a/n \,:\, n \in \mathbb{N}\} = \inf\{b \,:\, bn < a, n \in \mathbb{N}\},$$

which is the largest initial segment closed under addition not containing a. Various notational defices have been defined from time to time to express similar notions, for example the $M(a)$ and $M[a]$ of Kaye, Kossak and Kotlarski[23], though none is standard in the literature. Thus we introduce here the most convenient notational mechanism for our purposes. A more general and more verbatim notation system for such initial segments was defined in the author's MPhil. Qual. thesis.[1]

2.4 Coding sequences, sets and structures

Central to the work of this thesis is the notion of a number *coding* a finite set or a finite sequence. For the general discussion of this topic we refer the reader to the literature, for example Kaye[21], though we give a brief development of the basic machinery here.

[1] Note that our definition of $\mathfrak{M}$-finite will be defined later in this chapter.

The reader should note that there are many ways in which the following may be accomplished: we pick one for the purpose of being definite.

Coding sequences

Following chapter 5 of Kaye's book[21], where the reader may refer for a more detailed introduction to our coding mechanism, we recall the pairing function

$$\langle , \rangle \; : \; (x,y) \longmapsto \frac{(x+y)(x+y+1)}{2} + y$$

and extend this to arbitrary n-tuples via the recursive definition

$$\langle x_1, x_2, \ldots, x_n \rangle = \langle x_1, \langle x_2, \ldots, x_n \rangle \rangle \, .$$

We define $\mathrm{hd}(\langle x,y \rangle) = x$, which we call[2] the *head of the pair* $\langle x,y \rangle$, and $\mathrm{tl}(\langle x,y \rangle) = y$ which we call the *tail of the pair*. Note that in the case of n-tuples we have

$$\mathrm{tl}(\langle x_1, x_2, \ldots, x_n \rangle) = \langle x_2, \ldots, x_n \rangle \, .$$

We may now define our coding mechanism via

$$(x)_y = \mathrm{rem}(\mathrm{hd}(x), 1 + (y+1)\mathrm{tl}(x))$$

where $\mathrm{rem}(x,y)$ denotes the remainder of x upon division by y. (The reader may refer to Kaye chapter 5 for a more detailed introduction of this coding system.) We thus have

[2]The names 'head' and 'tail' come from the modern literature of functional programming among other places. This author believes this, along with the logic programming literature, to be a useful source of organisational methods for the long strings of logical symbols that one must deal with when formalising the work that occurs later in this thesis. Bird[8] is but one good introductory text in the area of functional programming.

Δ_0-formulae for "z is the yth element of the sequence coded by x" and "y is the length of the sequence coded by x" such that PA proves the following:[3]

$$\forall x, y \, \exists! z \, (x)_y = z, \tag{2.4.1}$$

$$\forall x, y \, (x)_y \leq x, \tag{2.4.2}$$

$$\forall x \, \exists y \, (y)_0 = x, \text{ and} \tag{2.4.3}$$

$$\forall x, y, z \, \exists w \, (\forall i{<}y (x)_i = (w)_i \, \wedge \, (w)_y = z) \, . \tag{2.4.4}$$

We wish our coding mechanism to uniformly encode the length of a sequence, so we follow the development of chapter 9 of Kaye[21] in defining

$$\mathrm{len}(x) = (x)_0 \tag{2.4.5}$$

$$[x]_y = \begin{cases} (x)_{y+1} & \text{for } y < \mathrm{len}(x)\text{; and} \\ 0 & \text{otherwise.} \end{cases} \tag{2.4.6}$$

We call such sequences $\mathfrak{M}$-finite sequences and think of the canonical code for a given sequence to be the least element of M (with respect to the ordering of $\mathfrak{M}$) that codes the sequence. The following defines the *empty sequence*, *singleton sequence* and *concatenation of sequences*:

$$[] = \mu x \, \mathrm{len}(x) = 0, \tag{2.4.7}$$

$$[x] = \mu y \, (\mathrm{len}(y) = 1 \wedge [y]_0 = x) \, , \tag{2.4.8}$$

$$\begin{aligned} x ^\frown y = \mu z \, (&\mathrm{len}(x) + \mathrm{len}(y) = \mathrm{len}(z) \wedge \\ &\forall w{<}\mathrm{len}(x)([x]_w = [z]_w) \wedge \\ &\forall w{<}\mathrm{len}(y)([y]_w = [z]_{w+\mathrm{len}(x)}\big) \, . \end{aligned} \tag{2.4.9}$$

[3]We wish to refer back to these properties from chapter 5, hence the numbering of the expressions.

The well known 'is an element of' symbol $\in$ may be defined between elements of M via

$$x \in y \text{ iff } \exists z{<}\mathrm{len}(y)\, [y]_z = x$$

and the complementary relation $\notin$ is defined in the obvious way.

Now we add three useful tools for extracting subsequences from a given sequence. The first is the *restriction* of a sequence, the second is a general tool known as *replacement* and the third is known as *comprehension* (the latter two echoing the terminology from set theory.) The first two are given by

$$x\restriction_y = \mu z\,(\mathrm{len}(z) = y \,\wedge\, \forall w{<}y\,[z]_i = [x]_i)$$

$$[f(r,\bar{a})\colon\, r \leftarrow x] = \mu z\,(\mathrm{len}(x) = \mathrm{len}(z) \,\wedge\, \forall w{<}\mathrm{len}(x)\,[z]_w = f([x]_w, \bar{a}))$$

where f denotes some function $f(x,\bar{y})$ and $\bar{a}$ is some tuple of parameters. In the above, r is a formal variable. Essentially $[f(r,\bar{a})\colon\, r \leftarrow x]$ refers to the algorithmic process: 'Take an element from the start of the input list and call it r; append $f(r,\bar{a})$ to the output list.'

The third is defined inductively by

$$[r\colon\, r \leftarrow c, \phi(r,\bar{a})] = \psi_\phi(c,\bar{a},x)$$

where

$$\psi_\phi(c,\bar{a},0) = []$$

and

$$\psi_\phi(c,\bar{a},x) = \begin{cases} \psi_\phi(c,\bar{a},x-1) & \text{if } \neg\phi(x-1,\bar{a})\text{; and} \\ \psi_\phi(c,\bar{a},x-1)^\frown[x-1] & \text{if } \phi(x-1,\bar{a}) \end{cases}$$

so that $[r\colon r \leftarrow c, \phi(r,\bar{a})]$ is the subsequence of the sequence coded by c consisting of all those $x \in M$ that satisfy $\phi(x,\bar{a})$ *in the order that they appear in the sequence coded by c, including any repeats.* We then unify sequence comprehension and replacement notation by defining

$$[f(r,\bar{a})\colon r \leftarrow x,\, \phi(r,\bar{b})] = [f(r,\bar{a})\colon r \leftarrow [s \leftarrow x\colon \phi(s,\bar{b})]].$$

The reader will notice that we use '$\leftarrow$' where traditional notation would use '$\in$'. The reasons for this are twofold: firstly we wish to define $\in$ in the context of $\mathfrak{M}$-finite sets and secondly, there is the question of the sequence from which the order of the new sequence is inherited. For example, for different sequences coded by c and d respectively, the sequences $[x\colon x \leftarrow c\colon x \in d]$ and $[x\colon x \leftarrow d\colon x \in c]$ are not equal in general, yet the notation $[x\colon x \in c\colon x \in d]$ could be construed to suggest this. Thus we use $\leftarrow$ to show where the order of the new sequence is inherited from. For a final piece of shorthand, we permit writing $[x \leftarrow c\colon x \in d]$ where it is formally meant $[x\colon x \leftarrow c\colon x \in d]$.

Set-like operations on sequences

We need a notion of coding sets, and the material immediately preceding this gives us just the machinery we need to implement such a notion. Much of the machinery applies readily to the sequences defined above, and so we introduce it in this context first, showing how we formalise the notion of an $\mathfrak{M}$-finite set in the next subsection. The definition of $\in$ was given earlier, and we turn our attention to the other symbols in the language employed in set theory.

We identify union with concatenation and define intersection and set difference using comprehension thus:

$$x \cup y = x^\frown y,$$

$$x \cap y = [r \leftarrow x \colon r \in y], \text{ and}$$

$$x \smallsetminus y = [r \leftarrow x \colon r \notin y].$$

The reader should be aware from the discussion at the end of the previous subsection that $x \cap y$ is not in general the same as $y \cap x$ when talking about sequences. We use primitive recursion to define the 'union' and 'intersection' of a sequence of sequences. Informally, we define

$$\bigcup[x_0, x_1, x_2, \ldots] = \bigcup[x_0 \cup x_1, x_2, \ldots]$$
$$\bigcap[x_0, x_1, x_2, \ldots] = \bigcap[x_0 \cap x_1, x_2, \ldots]$$

and for a more formal definition we consider $\mathfrak{M}$-finite sequences such as

$$[x_0 \cup x_1, (x_0 \cup x_1) \cup x_2, \ldots]$$

which are defined using primitive recursion. This process of applying a binary function to a sequence of nonstandard length, associating to the left or right as desired may be readily formalised, but we leave the details of this to the reader.

Coding sets and structures

It is now opportune to fix our method and notation for the notion of $\mathfrak{M}$-finite sets. We also introduce a few shorthand mechanisms to simplify what will otherwise become ungainly notation.

Firstly, we identify $\mathfrak{M}$-finite *sets* with minimal codes for strictly increasing $\mathfrak{M}$-finite *sequences*. The reader may use this as an effective definition of a set being $\mathfrak{M}$-finite.

We use $[\![c]\!]$ to denote the set coded by c. Strictly, $[\![c]\!]$ is defined to be the unique $c' \in M$ satisfying

$$\begin{aligned}(\mathfrak{M}, c, c') \vDash c' = \mu x{<}c\ (&\forall y{<}\mathrm{len}(c) \exists z{<}\mathrm{len}(x)\, [c]_y = [x]_z \wedge \\ &\forall y{<}\mathrm{len}(x) \exists z{<}\mathrm{len}(c)\, [x]_y = [c]_z \wedge \\ &\forall y{<}(\mathrm{len}(x) - 1)\, [x]_y < [x]_{y+1}\ )\end{aligned}$$

We then define the cardinality of an $\mathfrak{M}$-finite set (in the sense of $\mathfrak{M}$) by $|c| = \mathrm{len}([\![c]\!])$. The notation $\mathrm{card} X$ is used for cardinals in the set theoretic sense so as to keep these two notions distinct. We use

$$[\![f(r, \bar{a}) \colon r \leftarrow c, \phi(r, \bar{b})]\!]$$

as shorthand for

$$[\![[f(r, \bar{a}) \colon r \leftarrow c, \phi(r, \bar{b})]]\!].$$

(The reader should note that the parameters to the function f need not be the same as the parameters to the formula ϕ.) This is in turn used to formalise the usual informal set notation, which we use where convenient, for example we shall happily talk of 'the $\mathfrak{M}$-finite set $\{1, 2, 3\}$.'

We say that a subset $X \subseteq M$ is $\mathfrak{M}$-finite if there is $c \in M$ with $[\![c]\!] = X$ and write $|X| = |c|$ when this is the case. We use $\mathscr{P}$ to denote the $\mathfrak{M}$-finite power set of a given $\mathfrak{M}$-finite set X and $\mathscr{P}_n(X)$ to denote the (necessarily $\mathfrak{M}$-finite) set of $\mathfrak{M}$-finite n-subsets of X.

There is an alternative approach to coding sets that is to be found in the literature. This revolves around defining $x \in y$ to be 'the xth digit in the binary expansion of y is 1.' It is easy to verify that this gives the same set of $\mathfrak{M}$-finite sets, and that both give a model of set theory with the negation of the axiom of infinity.

We thus have a robust notion for coding sets and sequences. Our choice of system for coding sets makes these two notions reasonably interchangable: we can use $\mathfrak{M}$-finite sequences in lieu of ordered sets. We consider $\mathfrak{M}$-finite functions and relations to be coded sets of tuples and define predicates $\mathsf{IsFunction}$ and $\mathsf{IsRelation}$ with obvious intended meanings.

Finally, $\mathfrak{M}$-finite languages and structures may be defined by interpreting the usual definitions via the $\mathfrak{M}$-finite sets, functions and relations introduced above.

Chapter 3

Measuring subsets of a model of arithmetic

The ideas expounded here have their roots in the work of Kaye *et al.*[23], in particular the utilisation of the infimum of the number of elements satisfying the first n elements of a recursive type, and showing that the initial segment formed is fixed pointwise by an automorphism of the model $\mathfrak{M}$ if, and only if, the set of elements satisfying the type is also fixed pointwise. This showed a nontrivial relation between initial segments of a model and more subtle structures in a model. In a certain sense this echoed the earlier developments in the area of models of Peano arithmetic, where connections were sought between initial segments of models and other interesting aspects of the models.

In the development here, however, we focus on initial segments as a generalisation of the notion of counting rather than any particular application. Actual applications are not addressed in this chapter, this being a place to introduce potential new avenues for thought, rather than specific discoveries in current areas of thinking.

Returning to the roots of this chapter in the work of Kaye, the set of elements satisfying a type is either definable, or else very near to being definable. There are, however, many subsets of M for which we are not so lucky. We wish to generalise this notion of counting to arbitrary subsets of a model of Peano arithmetic. Thus, we begin with a reasonably natural approach to the 'counting' of such subsets. This approach was suggested by the author's supervisor at the outset of, and investigated in, the author's M. Phil. Qual. thesis[1]. Our exposition here is an abridged version of that work, though we endeavour to include the important points.

3.1 The $\mathfrak{M}$-measure of a subset of a model of PA

For the purposes of this section, $\mathfrak{M}$ will denote some arbitrary model of PA with domain M. Let X be some subset of M and let us consider the question of how the structure of $\mathfrak{M}$ may be used to measure the size of X. To this end we begin by defining the following two complementary notions of measure.

Definition 3.1.1. Let X and $\mathfrak{M}$ be as in the above paragraph. If X is $\mathfrak{M}$-finite then we define $\mu(X) = \nu(X) = |X|$. Otherwise we define

$$\mu(X) = \sup\{|x| \ : \ x \in M,\ (\mathfrak{M}, X, x) \models [\![x]\!] \subseteq X\} \text{ and}$$

$$\nu(X) = \inf\{|x| \ : \ x \in M,\ (\mathfrak{M}, X, x) \models X \subseteq [\![x]\!]\}.$$

We call $\mu(X)$ and $\nu(X)$ the *inner* and *outer* $\mathfrak{M}$-*measures* of X respectively. We say that X is $\mathfrak{M}$-*measurable* precisely when we have $\mu(X) = \nu(X)$.

We will look to generalisations of these definitions of measure in later sections. It is intuitively clear that μ will always underestimate and ν will always overestimate (in the non-strict sense), but by how much? We will begin to address this question after covering the basic facts concerning μ and ν.

Lemma 3.1.2. *Let X be a subset of M. If $a \in \mu(X)$ then there is $c \in M$ with $[\![c]\!] \subseteq X$ and $|c| = a$. If $a \geq \nu(X)$ then there is $c \in M$ with $[\![c]\!] \supseteq X$ and $|c| = a$.*

Proof. Suppose that $a \in \mu(X)$. Then there is $c' \in M$ with $a \leq |c'|$ and $[\![c]\!] \subseteq X$. But we can obtain the required $c \in M$ by deleting the last $|c'| - a$ elements from the sequence coded by c'. Suppose now that $a \geq \nu(X)$. Then there is $c' \in M$ with $|c'| \leq a$ and $X \subseteq [\![c']\!]$. But then the set

$$[\![c']\!] \cup \left\{c', \ldots, c' + (a - \left|c'\right| - 1)\right\}$$

is clearly $\mathfrak{M}$-finite and we need only take a code c for this set to complete the proof. $\square$

Lemma 3.1.3. *For any set $X \subseteq M$ we have $\mu(X) \subseteq \nu(X)$.*

Proof. Take any $d \in M$ such that $X \subseteq [\![d]\!]$ and take any $a \in \mu(X)$. For any $c \in M$ with $|c| = a$ and $[\![c]\!] \subseteq X$ (there are such c by the previous lemma) we have $[\![c]\!] \subseteq [\![d]\!]$ and so immediately we have $a \leq |d|$. Since a was arbitrary we have $\mu(X) \subseteq |d|$ and since d was arbitrary subject to $X \subseteq [\![d]\!]$ the result follows. $\square$

Lemma 3.1.4. *For all initial segments I we have $\mu(I) = \nu(I) = I$.*

Proof. For $a \in I$ we have $_{<}a \subseteq I$ and hence $a \in \mu(I)$. When $a \notin I$ we have $I \subseteq {}_{<}a$ and hence $\mu(I) \leq a$. The result then follows by lemma 3.1.3. $\square$

It is not always most convenient to talk of $\mathfrak{M}$-finite sets in terms of codes for them. In the next lemma we deal with $\mathfrak{M}$-finite sets more directly.

Lemma 3.1.5. *Let X be an $\mathfrak{M}$-measurable subset of M and let Y be an $\mathfrak{M}$-finite set disjoint from X. Then $X \cup Y$ is measurable and*

$$\mu(X) + |Y| = \mu(X \cup Y) \text{ and } \nu(X \cup Y) = \nu(X) + |Y|.$$

Proof. Let Z be an $\mathfrak{M}$-finite subset of X. Then $Y \cap Z = \varnothing$ and $X \cup Z$ is an $\mathfrak{M}$-finite subset of $X \cup Y$. If Z is an $\mathfrak{M}$-finite superset of X then so is $Z \setminus Y$, so assume without loss that $Z \cap Y = \varnothing$. Then $Z \cup Y$ is an $\mathfrak{M}$-finite superset of $X \cup Y$ with $|Z \cup Y| = |Z| + |Y|$. Thus for $a \in \mu(X) = \nu(X)$ we have $a + |Y| \in \mu(X \cup Y)$ and for $a \geq \mu(X)$ we have $a + |Y| \geq \nu(X \cup Y)$. The result then follows. □

Lemma 3.1.6. *Let X and Y be disjoint. Then*

$$\mu(X) + \mu(Y) \subseteq \mu(X \cup Y) \text{ and } \nu(X \cup Y) \subseteq \nu(X) + \nu(Y).$$

Proof. First we show the result for μ. Take some $a \in \mu(X) + \mu(Y)$ and write $a = a_X + a_Y$ where $a_X \in \mu(X)$ and $a_Y \in \mu(Y)$. Then there are c_X and c_Y in M with $\operatorname{len}(c_X) = a_X$, $[\![c_X]\!] \subseteq X$, $\operatorname{len}(c_Y) = a_Y$ and $[\![c_Y]\!] \subseteq Y$. But then we may take $[\![c]\!] = [\![c_X \cup c_Y]\!]$ with $\operatorname{len}(c) = a$ and this suffices to show that $a \in \mu(X \cup Y)$.

We prove the result for ν in a similar fashion. Take any $a \geq \nu(X) + \nu(Y)$ and write $a = a_X + a_Y$ where $a_X \geq \nu(X)$ and $a_Y \geq \nu(Y)$. Then there are c_X and c_Y in M with $\operatorname{len}(c_X) = a_X$, $X \subseteq [\![c_X]\!]$, $\operatorname{len}(c_Y) = a_Y$ and $Y \subseteq [\![c_Y]\!]$. But then $\operatorname{len}(c_X \cup c_Y) \leq \operatorname{len}(c_X) + \operatorname{len}(c_Y)$ and so $a \geq \nu(X \cup Y)$ as required. □

There is a significant corollary to the above.

Corollary 3.1.7. *Let X and Y be disjoint $\mathfrak{M}$-measurable sets. Then $X \cup Y$ is $\mathfrak{M}$-measurable.*

It should be noted that the later result, Theorem 3.2.9, giving the existence of non-$\mathfrak{M}$-measurable sets may be used to show that we cannot do much better than the above. As a generalisation of this lemma we have the following.

Lemma 3.1.8. *Let X and Y be bounded subsets (not necessarily $\mathfrak{M}$-measurable) of M such that there is a formula $\phi(x,\bar{y})$ and $\bar{a} \in M$ such that $M \vDash \phi(x,\bar{a})$ if $x \in X$ and $M \vDash \neg\phi(x,\bar{a})$ if $x \in Y$. Then $\mu(X)+\mu(Y) = \mu(X \cup Y)$ and $\nu(X \cup Y) = \nu(X)+\nu(Y)$.*

Proof. First we have $\mu(X)+\mu(Y) \subseteq \mu(X \cup Y)$ from Lemma 3.1.6. Then take any $b \in \mu(X \cup Y)$ and some $\mathfrak{M}$-finite $C \subseteq X \cup Y$ with $|C| = b$. It follows that

$$A = \{x \in C \,:\, (\mathfrak{M},\bar{a},x) \vDash \phi(x,\bar{a})\}$$

$$B = \{x \in C \,:\, (\mathfrak{M},\bar{a},x) \vDash \neg\phi(x,\bar{a})\}$$

are two disjoint $\mathfrak{M}$-finite subsets of $X \cup Y$ with $A \subseteq X$, $B \subseteq Y$ and $|A|+|B| = |C|$. But then $|A|+|B| \in \mu(X)+\mu(Y)$ and it follows that $\mu(X \cup Y) \subseteq \mu(X)+\mu(Y)$ as required.

We also have $\nu(X \cup Y) \subseteq \nu(X)+\nu(Y)$ from Lemma 3.1.6. Let C be any $\mathfrak{M}$-finite superset of $X \cup Y$ and let A and B be as before. Then $|C| = |A|+|B| \geq \nu(X)+\nu(Y)$ and thus we have $\nu(X \cup Y) \supseteq \nu(X)+\nu(Y)$ as required to complete the proof. □

Of more importance and interest is the question of whether the $\mathfrak{M}$-measurable sets form an algebra of sets. That is, if X and Y are $\mathfrak{M}$-measurable then what about $X \cup Y$ and $X \cap Y$. This author suspects that the notion of $\mathfrak{M}$-measurability looked at here may well be too weak for this, but cannot come up with a better refinement of this notion.

3.2 Non-$\mathfrak{M}$-measurable sets of a model of PA

We begin with the following observation. Essentially, a set has infinite measure if, and only if, it is unbounded. In contradistinction, on the real line it is possible for an unbounded set to have zero measure (for example the rational numbers.)

Lemma 3.2.1. *For any subset $X \subseteq M$, we have $\nu(X) = M$ if, and only if X is a cofinal subset of M.*

Proof. Since any element of the set coded by c, say, is bounded by c, it immediately follows that no $\mathfrak{M}$-finite set can contain a cofinal subset of M.

Conversely, if a subset of $X \subseteq M$ is bounded by a, say, then clearly

$$X \subseteq \{x \in M \ : \ x < a\}$$

and so $\nu(X) \leq a$. $\square$

In a model of cofinality $\aleph_0$, a cofinal ω-sequence then yields an example of how much μ and ν may differ, μ giving the smallest possible value for a non-$\mathfrak{M}$-finite set, namely $\mathbb{N}$ and ν giving the maximum, namely M. In this particular case it is clearly ν that is held at fault, its definition having to default to M for lack of codes to choose from. We approach this problem with generalisations of the measures introduced in a later section.

As a more interesting example (or family of examples), and as an introduction to the systematic construction of non-$\mathfrak{M}$-measurable sets we give the following.

Question 3.2.2. Let J be a proper cut of $\mathfrak{M}$. There is a natural order isomorphism

$$f\colon J \to \mathbb{N} + \mathbb{Z} \cdot Q$$

for some dense linear order Q with no initial endpoint. Let

$$X = f^{-1}(\mathbb{N} + \mathbb{N} \cdot Q).$$

Then what are $\mu(X)$ and $\nu(X)$?

A partial answer was presented in the author's M. Phil. Qual. thesis, but is included here more as food for thought. The answer, as it turns out, really is 'it depends.' In this case, the answer depends upon the precise choice of order isomorphism, and in particular what gets mapped to the zeroes of the $\mathbb{Z}$'s.

We now look briefly at a couple of constructions of non-$\mathfrak{M}$-measurable sets. For this first example, which we present as the following lemma, $\lfloor J+ \rfloor$ will be used to denote the largest initial segment $I \subseteq J$ such that J is closed under addition with I, that is $J + I \subseteq J$. Essentially this picks up on the point that for an initial segment J not closed under addition, there are many $x \in J$ such that $x + J = J$: we use $\lfloor J+ \rfloor$, for which there is no standard notation, to denote the initial segment consisting of all such x.

Lemma 3.2.3. *Let J be a proper cut in $\mathfrak{M}$. Then for any ω-sequence $X = (x_i)$ cofinal in J we have $\mu(X) = \mathbb{N}$ and $\nu(X) \subseteq \lfloor J+ \rfloor$.*

Proof. In case J is not closed under addition take $a, b \in J$ with $a + b > J$. Let $n \in \mathbb{N}$ so that $\{x_0, \ldots, x_n\}$ is the set containing the first $n+1$ elements of X and observe that

$$X \subseteq \{x_0, \ldots, x_n\} \cup \{x \in X \ : \ a \leq x < a + b - (n+1)\}.$$

Now $\{x_0, \ldots, x_n\}$ is actually finite and $\{x \in X \ : \ a \leq x < a + b - (n+1)\}$ is contained in an $\mathfrak{M}$-finite set with cardinality $b - n - 1$ in the sense of $\mathfrak{M}$. The proof then follows from this observation. □

We leave as an exercise in diagonalisation, similar to the proof of the next lemma, that this bound can be attained. It is trivial to show that the bound needn't be attained. We next show how μ and ν can be systematically moved apart in a countable model. The difficulties obtained by trying to generalise this result to uncountable models are not addressed here. In the proof of the following results we use the notion of a $\mathbb{Z}$-block from chapter 6 of Kaye[21]. Specifically we mean the following.

Definition 3.2.4. Let $\mathfrak{M}$ be a nonstandard model of PA and let $a \in M$. By the *$\mathbb{Z}$-block around a* we mean the set $\{x \in M \ : \ |x - a| \text{ is finite}\}$ and we denote this set by $[a]$.

Recall that this notion arises naturally from consideration of the order type of $\mathfrak{M}$, which is $\omega + (\omega^* + \omega) \cdot Q$ for some dense linear order Q without endpoints. The theorem that follows is broken down into a short series of lemmas which elucidate the various cases.

Lemma 3.2.5. *Let I and J be proper cuts of a countable model $\mathfrak{M} \models \mathsf{PA}$ such that $J \setminus I$ is order isomorphic to $\mathbb{Z}$. Then there is a subset X of M such that $\mu(X) = I$ and $\nu(X) = J$.*

Proof. Suppose that $J \setminus I$ is order isomorphic to $\mathbb{Z}$, that is, it is a single $\mathbb{Z}$-block. Let $\{d_n \,:\, n \in \mathbb{N}\}$ enumerate the set $\{a \in M \,:\, (\mathfrak{M}, I, J, a) \models |a| \in J \setminus I\}$.

We construct X inductively, beginning with $X_0 = I$ and x_0 some arbitary element of $J \setminus I$. At stage n we assume that X_n has been constructed in accordance with this procedure. If $X_n \not\subseteq [\![d_n]\!]$ then we simply take $x_{n+1} = x_n + 1$. If not, then observe that since $I \subseteq X_n \subseteq [\![d_n]\!]$ then by overspill we have that $(\mathfrak{M}, d_n, a) \models \forall x{<}a\, x \in [\![d_n]\!]$ for some $a \in J \setminus I$. But then $|d_n| - a$ is finite, so $[\![d_n]\!]$ has a maximum m_n in $J \setminus I$. In this case we take $x_{n+1} = \max(x_n + 1, m_n + 1)$. In either case we then define $X_{n+1} = X_n \cup \{x_{n+1}\}$. Finally we take $X = \bigcup_{n \in \mathbb{N}} X_n$.

It is evident that the resulting X is the union of I and a strictly increasing ω-sequence contained within $J \setminus I$. We claim that $\mu(X) = I$. To see this, let us write $X = U \cup X'$ where X' and I are disjoint. Then $\mu(I) = \nu(I) = I$ and any $\mathfrak{M}$-finite subset of X' is actually finite. But I is closed under successor and so it follows that $\mu(I) \leq \mu(X) \leq \mu(I) + \mathbb{N} = \mu(I)$ which proves the claim.

Furthermore, for every $d \in M$ with $|d| \in J$ either $I \not\subseteq [\![d]\!]$ whence $X \not\subseteq [\![d]\!]$ or else $|d| \in J \setminus I$ and the construction then guarantees that $X \not\subseteq [\![d]\!]$. Since this is true for *every* $d \in M$ with $|d| \in J$ we cannot possibly have $d > \nu(X)$. Thus $\nu(X) \subseteq J$ and the proof is complete. □

Lemma 3.2.6. *Let I and J be proper cuts of a countable model $\mathfrak{M} \models \mathsf{PA}$. Suppose that for all $a \in J$ there is $b \in J$ with $b > a$ and $b - a$ nonstandard. Then there is a subset X of M such that $\mu(X) = I$, $\nu(X) = J$, $\mu(J \smallsetminus X) \subseteq I$ and $\nu(J \smallsetminus X) = J$. In fact, there are $2^{\aleph_0}$ such subsets.*

Proof. Suppose that for all $a \in J$ there is $b \in J$ with $b > a$ and $b - a$ nonstandard. Then there are countably many $\mathbb{Z}$-blocks in $J \smallsetminus I$ and no topmost $\mathbb{Z}$-block.

Since the model is countable, we may let $\{d_n : n \in \mathbb{N}\}$ enumerate the set

$$\{a \in M : (\mathfrak{M}, I, a) \models |a| \notin I \text{ and } (\mathfrak{M}, J, a) \models |a| \in J\}.$$

We proceed by constructing a descending chain $\{X_n\}$ of subsets of J such that $X = \bigcap_{n \in \mathbb{N}} X_n$ is the required subset. Before we begin, fix some $t \in J \smallsetminus I$ so that we may use the elements of the $\mathbb{Z}$-block of t to get the $2^{\aleph_0}$ subsets desired.

Now, let $X_0 = J$ and proceed inductively. At stage n we assume that X_n has been constructed in accordance with this procedure. Pick, and consider touched, four distinct $\mathbb{Z}$-blocks, say $[r_n]$, $[r_n']$, $[s_n]$ and $[s_n']$ in $J \smallsetminus I$ which have not been touched at a previous stage (and by this we mean to include $[t]$), such that:

1. there is some $b_n \in [r_n] \smallsetminus [\![d_n]\!]$, which is possible since $|d_n| \in J$ and J has no topmost $\mathbb{Z}$-block;
2. there is some $b_n' \in [\![d_n]\!] \smallsetminus [r_n']$, which is possible since $|d_n| \notin I$;
3. there is some $c_n \in [s_n] \smallsetminus [\![d_n]\!]$; and
4. there is some $c_n' \in [\![d_n]\!] \smallsetminus [s_n']$.

One can see that this process never runs into difficulties, for given any $d \in M$ with $|d| \in J \smallsetminus I$, it follows that $[\![d]\!] \smallsetminus I$ intersects countably many $\mathbb{Z}$-blocks, as does $J \smallsetminus [\![d]\!]$.

Now we construct X_{n+1} by defining $X_{n+1} = X_n \setminus ([b'_n] \cup [c_n])$ and observe that:

1. $X_{n+1} \not\subseteq [\![d_n]\!]$ since $b_n \in X_{n+1} \setminus [\![d_n]\!]$;
2. $[\![d_n]\!] \not\subseteq X_{n+1}$ since $b'_n \in [\![d_n]\!] \setminus X_{n+1}$;
3. $J \setminus X_{n+1} \not\subseteq [\![d_n]\!]$ since $c_n \in (J \setminus X_{n+1}) \setminus [\![d_n]\!]$;
4. $[\![d_n]\!] \not\subseteq J \setminus X_{n+1}$ since $c'_n \in [\![d_n]\!] \setminus (J \setminus X_{n+1})$.

Do this procedure for each $n \in \mathbb{N}$ and define $X = \bigcap_{n\in\mathbb{N}} X_n$. This resulting X clearly satisfies the statement of the theorem. Furthermore, we can replace the $\mathbb{Z}$-block $[t]$ with any subset of $[t]$ and the rest of the argument runs through essentially unchanged. Thus there are $2^{\aleph_0}$ subsets of J satisfying the requirements of the theorem. □

Corollary 3.2.7. *Let I, J and $\mathfrak{M}$ be as in the statement of Lemma 3.2.6. Suppose that I is closed under addition. Then X may be taken such that $\mu(J \setminus X) \subseteq I$. As before, there are $2^{\aleph_0}$ such subsets.*

Proof. Take X' to be the X obtained above and define $X = X' \setminus \{2x : x \in I\}$. □

Lemma 3.2.8. *Let I and J be proper cuts of a countable model $\mathfrak{M} \models \mathsf{PA}$. Suppose that there is $a \in J$ such that for no $b \in J$ with $b > a$ is it the case that $b - a$ nonstandard, but that $J \setminus I$ is not order isomorphic to $\mathbb{Z}$, that is, J has an uppermost $\mathbb{Z}$-block but I does not so that $J \setminus I$ is not a single $\mathbb{Z}$-block. Let $J' = \inf\{a - n : n \in \mathbb{N}\}$. Then there is a subset X of M such that $\mu(X) = I$, $\nu(X) = J$, $\mu(J \setminus X) \subseteq I$ and $J' \subseteq \nu(J \setminus X)$. In fact, there are $2^{\aleph_0}$ such subsets.*

Proof. Apply Lemma 3.2.6 to I and J' to get a subset X' of J' (in fact, to get $2^{\aleph_0}$ subsets). For any particular X' among these $2^{\aleph_0}$ subsets we may do the following.

Let $\{d_n : n \in \mathbb{N}\}$ enumerate the set $\{a \in \mathfrak{M} : (\mathfrak{M}, J', J, a) \models |a| \in J \setminus J'\}$. Let x_0 be some arbitrary element of $J \setminus J'$. At stage n we assume that X_n has been constructed in accordance with this procedure. If $X_n \not\subseteq [\![d_n]\!]$ then take $x_{n+1} = x_n + 1$ and we are done. If not then consider $[\![d_n]\!] \cap [0, x_n]$. Since this intersection contains X' we have

$$|[\![d_n]\!] \cap [0, x_n]| \notin J'$$

from the construction of X' and so

$$|[\![d_n]\!] \cap [0, x_n]| \in J \setminus J'.$$

But then $[\![d_n]\!] \setminus [0, x_n]$ is finite. Thus we may take x_{n+1} to be some element of $J \setminus [0, x_n]$ not contained in $[\![d_n]\!]$. Now take $X = X' \cup \{x_n : n \in \mathbb{N}\}$.

Observe that any $\mathfrak{M}$-finite superset of X must have outer $\mathfrak{M}$-measure greater than any element of J since such a superset contains X' and cannot be any of the d_n from above. But $X \subseteq J$ and so $\nu(X) = J$. Also, $\nu(J \setminus X)$ contains J' since we constructed X' using Lemma 3.2.6.

Now we may apply Lemma 3.1.8 with $\phi(x, x_0) = x < x_0$ to get

$$\mu(X) = \mu(X') + \mu(X'') = I + \mathbb{N} = I$$

where $X'' = \{x_n : n \in \mathbb{N}\}$. Finally, $\mu(J \setminus I)$ follows from the application of Lemma 3.2.6. $\square$

Collecting the above together, we see that we can prove the following.

Theorem 3.2.9. *Let $\mathfrak{M}$ be a countable nonstandard model of* PA *and let I and J be proper cuts of $\mathfrak{M}$ with $I \subseteq J$. Then the following are true.*

1. *There is a subset X of M such that $\mu(X) = I$ and $\nu(X) = J$.*

2. *Further, if $J \setminus I$ is not order isomorphic to $\mathbb{Z}$ then there are $2^{\aleph_0}$ subsets X of M such that $\mu(X) = I$ and $\nu(X) = J$.*

3. *Further still, if for all $a \in J$ there is $b \in J$ with $b > a$ and $b - a$ nonstandard then X can be taken such that $\mu(J \setminus X) \subseteq I$ and $\nu(J \setminus X) = J$.*

4. *Finally, if we also have that I is closed under addition then X can be taken such that $\mu(J \setminus X) = I$.*

It is instructive to compare this measure with the Lebesgue construction on the real line. Recall that the Lebesgue measure may be constructed as a standard part of a hyperfinite measure, as explained in Lindstrøm[29]. There exist non-Lebesgue-measurable subsets of the real line (for example a set of coset representatives of the rationals in the reals taken between 0 and 1.) There also exist universally nonmeasurable subsets of the interval $\{0 \leq x \leq 1\}$ as explained in Fremlin[14] and Ciesielski[12]. (The reader may recall that a subset X of the real interval $[0, 1]$ is universally nonmeasurable if, and only if the inner Lebesgue measure of both X and its complement in $[0, 1]$ is 0 whilst both outer Lebesgue measures are 1.)

As such, then, the pathological examples exist similarly for both these measures on nonstandard models and the Lebesgue measure on the real line. It is known that to show the existence of non-measurable sets requires the axiom of choice and it would be an interesting question as to what happens in this case, though an investigation of this goes beyond the scope of this thesis.

3.3 Measuring using end-extensions

We saw in the previous section that our outer measure falls into difficulty in the case of cofinal subsets of the model. The purpose of this section is to show how this matter may be rectified by temporarily end-extending the model.

We begin by recalling some relevant definitions. The reader may refer to Kossak and Paris[26] for an overview of what are termed extensional sets. Note that our notion of 'coded in' is a slight generalisation of that usually found in the literature.

Definition 3.3.1. Let $\mathfrak{M}$ be a model of PA and let V be a subset of M. A subset X of V is said to be *coded in* a subset Y of M if there is $c \in Y$ such that

$$(\mathfrak{M}, V, X, c) \vDash \forall x\ (x \in X \ \leftrightarrow\ x \in V \wedge x \in [\![c]\!])$$

that is, such that $X = V \cap [\![c]\!]$. Usually V is taken to be some initial segment, often a model of PA as will be the case in the sequel.

Definition 3.3.2. Let $\mathfrak{M}$ be a model of PA and let X be a subset of M. We say that X is *extensional* if X is coded in some end-extension $\mathfrak{N}$ of $\mathfrak{M}$ and *elementarily extensional* if $\mathfrak{N}$ can be taken such that $\mathfrak{M} \prec_e \mathfrak{N}$.

The reader may verify that in the above definition, nothing is lost by insisting that $\operatorname{card}\mathfrak{N} = \operatorname{card}\mathfrak{M}$ (essentially by considering the extension of $\mathfrak{M}$ within $\mathfrak{N}$ 'generated' by the code for X.)

The ease with which $[\![c]\!] \cap M$ may be measured for $c \in N$ motivates the following definition.

Definition 3.3.3. Let $\mathfrak{M}$ and $\mathfrak{N}$ be models of PA, $\mathfrak{N}$ an end-extension of $\mathfrak{M}$. Let $\nu^{\mathfrak{N}}$ denote the outer measure of the previous section, taken in the sense of $\mathfrak{N}$. We define the $\mathfrak{N}$-*extended outer measure on* $\mathfrak{M}$ by

$$\nu_{\mathfrak{N}}^{\mathfrak{M}}(X) = M \cap \nu^{\mathfrak{N}}(X).$$

We say that a subset X of M is $\mathfrak{N}$-*measurable in* $\mathfrak{M}$ if $\mu^{\mathfrak{M}}(X) = \nu_{\mathfrak{N}}^{\mathfrak{M}}$ where $\mu^{\mathfrak{M}}$ denotes the inner measure on $\mathfrak{M}$ introduced in the previous section. Where context allows we omit annotations on μ and ν, preferring to write just μ, ν and $\nu_{\mathfrak{N}}$.

The measures just introduced were motivated by the problems caused to the outer measure defined earlier in the case of cofinal subsets of the model. Moving to end extensions will only help this case, as the next lemma illustrates. The proof is straightforward.

Lemma 3.3.4. *Let $\mathfrak{M}$ be a model of* PA *and let X be a bounded subset of M. Then for any end-extension $\mathfrak{N}$ of $\mathfrak{M}$ we have $\nu_{\mathfrak{N}} = \nu_{\mathfrak{M}}(X)$.*

It is evident from the following lemma that extensional sets are the right generalisation of $\mathfrak{M}$-finite sets.

Lemma 3.3.5. *Let X be coded in an end-extension $\mathfrak{N}$ of $\mathfrak{M}$. Then $\mu(X) = \nu_{\mathfrak{N}}(X)$.*

Thus, given a cofinal subset X of M, one would be inclined to look for extensional supersets Y of X and end extensions $\mathfrak{N}$ of $\mathfrak{M}$ coding them. We conclude this section with a result on constructions of non-extended-measurable sets.

Lemma 3.3.6. *Let I and J be cuts of M with $I \subseteq_e J$. Let $\mathfrak{N}$ be an end-extension of $\mathfrak{M}$ coding a cofinal ω-sequence. Then there is a subset X of M such that $\mu(X) = I$, $\nu(X) = M$ and $\nu_{\mathfrak{N}}(X) = J$.*

Proof. Apply Theorem 3.2.9 to get $Y \subseteq M$ with $\mu(Y) = I$ and $\nu(Y) = J$. Let W be the cofinal ω-sequence coded in $\mathfrak{N}$ and take $X = W \cup Y$. $\square$

This last result can be generalised a little without much effort, but we leave that to the interested reader.

3.4 Summation and approximation of external functions

The purpose of this section is to introduce a possible area of application of the preceding material. As such, we outline the basic idea and leave things there. One is motivated by the question of how well one can approximate a given function on a given subset of a model by an $\mathfrak{M}$-finite function or a system of functions. The need to evaluate accuracy leads to the need to be able to sum the values of an external function over a possibly non-$\mathfrak{M}$-finite set and this requires the ideas of measure explored earlier in this chapter. Just as the measure constructions of before had similarities with Lebesgue measure, these summation ideas have similarities with Lebesgue and Riemann integration. This motivates the following definition.

Definition 3.4.1. Let $f\colon M \to M$ be defined on a set $X \subseteq M$. We define the lower and upper integrals of $f(x)$ over X as follows.

$$\int_{x \in X} f(x) \bar{d}\mu = \sup\left\{ \sum_{x \in [\![c]\!]} [d]_x \ : \ (\mathfrak{M}, X, f, c, d) \vDash [\![c]\!] \subseteq X \ \wedge \ \forall x {\in} [\![c]\!] \ ([d]_x \leq f(x)) \right\}$$

$$\int_{x \in X} f(x) \underline{d}\nu = \inf\left\{ \sum_{x \in [\![c]\!]} [d]_x \ : \ (\mathfrak{M}, X, f, c, d) \vDash X \subseteq [\![c]\!] \ \wedge \ \forall x {\in} [\![c]\!] \ ([d]_x \geq f(x)) \right\}$$

Note that the $\underline{d}\mu$ and $\bar{d}\nu$ is used to indicate how X is to be measured and the location of the underline or overline indicates how $f(x)$ is to be approximated. When these integrals agree we may omit the bar and just write ‘dx’. In fact, $\int_{x\in X} f(x)dx$ will also be used to represent the integral or sum that we wish to approximate and measure.

Importantly, given an external function $f\colon M \to M$ we may consider the integral

$$\int_{x\in X} [f(x) - [c]_x]^2$$

and use this as a means to determine when one approximation is more exact than another. The reader should also note that minimising the square of the error is not the only way of finding a best approximation, and analogies to other statistical methods may be worthwhile investigating.

Chapter 4

Nonstandard finite permutations

4.1 Nonstandard permutations in models of PA

Any permutation on a finite subset of M is necessarily an $\mathfrak{M}$-finite bijection. It follows immediately that there are nonstandard $\mathfrak{M}$-finite permutations acting on infinite $\mathfrak{M}$-finite sets. To simplify the definitions, we identify $\mathfrak{M}$-finite permutations with their images coded as $\mathfrak{M}$-finite sequences (*e.g.* the permutation of $[0,n]$ represented in cycle notation as $(1,2,3)$ will be identified with the sequence $[0,2,3,1,4,5,\ldots,n]$). We shall write Ω_n for the set $[0,n-1]$.

We begin by fixing some formula $\mathsf{Perm}(n,g)$ with the intended meaning: g is a minimal code for an $\mathfrak{M}$-finite bijection $\Omega_n \to \Omega_n$. We define the nth symmetric group S_n to be the set $\{x \in M \,:\, \mathsf{Perm}(n,x)\}$ together with the obvious definition for the group operation. We denote the identity element by 1_n. This group should be distinguished from the full symmetric group on the subset $[0,n-1]$ of the domain of $\mathfrak{M}$. We write this full symmetric group as Sym_n (though strictly it may be more correct to use notation such as $\mathrm{Sym}(X)$ where $X = \{x \in M \,:\, x < n\}$).

We adopt standard notation for the action of S_n upon Ω_n, writing x^g for the image of x under the action of g, and defining composition of permutations in the natural way so that $(x^g)^h = x^{gh}$.

Permutations have *cycle shapes* and may be written in *disjoint cycle notation* and we need to briefly explain what this means for nonstandard permutations; essentially we are aiming to capture what happens with finite permutations.

Definition 4.1.1. Let $g \in \mathrm{S}_n$. We define κ_g to be the sequence of elements of S_n corresponding to the disjoint cycles, ordered by the least point moved for each respective cycle. We define κ_g^* to be the sequence of sequences representing those cycles. That is, if $g = (1,2,3)(4,5,6)$ then $\kappa_g = [(1,2,3),(4,5,6)]$ and $\kappa_g^* = [[1,2,3],[4,5,6]]$.

By $\sigma_g(i)$ we mean the number of i-cycles in the disjoint cycle form of g. Thus σ_g is a function that is definable (with parameters) for each $g \in \mathrm{S}_n$ and which satisfies $\sum_{i \in M} i\sigma_g(i) = n$.

Now identify S_n with its natural embedding in Sym_n and let $g \in \mathrm{Sym}_n$. By $\tau_g(i)$ we mean the number of i-cycles in g where i is a finite cardinal or $\aleph_0$ and $\tau_g(i)$ is a (possibly finite) cardinal.

To clarify the last of the above definitions with an explicit example, let

$$g = (0,1,2)(3,4)(5,6)(7,8)\ldots(2a-1,2a)$$

where $a \in M$ is nonstandard with $\mathrm{card}\,[0,2a] = \aleph_0$. Then with the above definition, $\tau_g(2) = \aleph_0$, $\tau_g(3) = 1$ and $\tau_g(i) = 0$ for all other values of i. Note that by $\tau_g(\aleph_0)$ we mean the number of cycles with order type $\omega^* + \omega$.

Two notions central to the work of the next chapter are the support and degree of a permutation.

Definition 4.1.2. Let $g \in \mathrm{S}_n$. We define the *support* and *degree* of g via

$$\mathrm{supp}(g) = \{x \in [0, n-1] \,:\, x \text{ is moved by } g\}$$

and

$$\deg(g) = |\mathrm{supp}(g)|$$

respectively.

Like finite permutations, and unlike permutations of infinite support in the usual case of infinite permutation groups, $\mathfrak{M}$-finite permutations are either odd or even. There are a number of ways of defining this, such as according to whether the product

$$\prod_{i \neq j} \frac{i^g - j^g}{i - j}$$

is $+1$ or -1. Probably the most convenient in terms of arithmetical concepts required is to say that g is even if, and only if the sum $\sum_{2 \leq i < n} (i-1)\sigma_g(i)$ is even, which essentially amounts to counting transpositions. This allows us to define the nth alternating group in the obvious way, via $\mathrm{A}_n = \{g \in \mathrm{S}_n \,:\, g \text{ is even}\}$.

One should be aware that a permutation on Ω_n is never a permutation on a different set (since with our definitions, the set acted on may be recovered directly from the permutation.) As would be expected, given a subset X of Ω_n, the set of all permutations on X is isomorphic to a subgroup of S_n, and we denote the image of this isomorphism as $\mathrm{S}_n(X)$, though the reader may be aware that this image is simply the pointwise stabiliser $(\mathrm{S}_n)_{(\Omega_n \smallsetminus X)}$ of the complement of X in Ω_n. In practice, this issue causes little difficulty and is easily addressed.

4.2 Conjugates of $\mathfrak{M}$-finite permutations inside Sym_n

It is clear that, as abstract groups, $\mathrm{Sym}_n \cong \mathrm{Sym}(\kappa)$ where $\kappa = \mathrm{card}\,\{x \in M\ :\ x < n\}$. Of course there are many isomorphisms between these two groups, and the image of $\mathrm{S}_n \leq \mathrm{Sym}_n$ will vary from isomorphism to isomorphism. The union of all such images is the image of the normal closure of S_n within Sym_n. In this section we offer a characterisation of this normal closure for certain models of PA and a partial characterisation for arbitrary models of PA. Fix and arbitary model $\mathfrak{M}$ of PA, and choose some fixed nonstandard $n \in M$.

Lemma 4.2.1. *Let $g \in \mathrm{Sym}_n$ be conjugate in Sym_n to some $h \in \mathrm{S}_n$. Then either $\tau_g(\aleph_0)$ is infinite or else the size of the cycles in g has a finite bound.*

Proof. Suppose $\tau_g(\aleph_0) \neq 0$. Then g has a cycle of nonstandard order. Let this cycle be $(a_0, a_1, \ldots, a_k)$ and consider the order type of the initial segment $[0, k]$. This must be $\omega + (\omega^* + \omega) \cdot Q + \omega^*$ where * indicates reverse ordering and Q is a densely linearly ordered set. From here it is clear that there must be infinitely many $\mathbb{Z}$-cycles — one $\mathbb{Z}$-cycle per $\mathbb{Z}$-block. That the order of cycles must have a finite bound in case $\tau_g(\aleph_0) = 0$ follows from an easy overspill argument: if the order of cycles does not have a finite bound, there will be a cycle of nonstandard length by overspill. □

Now for simplicity, we let $\mathfrak{M}$ be a countable model. There are a number of ways of trying to code the external cycle shape of a permutation in Sym_n, though it is not difficult to see that the notion of a cycle shape having a code does not depend in a sensitive way upon the definition. We use the following.

Definition 4.2.2. We say that τ_g is *coded* if, and only if there is some $c \in M$ such that

$$[c]_0 = \begin{cases} 0 & \tau_g(\aleph_0) = 0 \\ 1 & \tau_g(\aleph_0) = \aleph_0 \end{cases}$$

$$[c]_i = \begin{cases} 0 & \tau_g(i) = \aleph_0 \\ \tau_g(i) + 1 & \text{otherwise} \end{cases}$$

for all $i > 0$.

Lemma 4.2.3. *Suppose that $n \in M$, $g \in \mathrm{Sym}_n$ and either there is a nonstandard cycle in g or else there is a finite bound to the lengths of cycles in g. Suppose further that τ_g is coded. Then there is $h \in \mathrm{S}_n$ such that h is conjugate to g in Sym_n.*

Proof. It suffices to construct $h \in \mathrm{S}_n$ with $\tau_g = \tau_h$. This is straightforward given a code for τ_g and we leave the details to the reader. $\square$

The reader will recall from chapter 2 that an initial segment I of a model $\mathfrak{M}$ is *strong* if, and only if given an $\mathfrak{M}$-finite function f, there is $a \in M$ such that for all $x \in I$, $f(x) > I$ if, and only if $f(x) > a$.

Theorem 4.2.4. *A necessary and sufficient condition for $\mathbb{N}$ to be strong in M is the following: $g \in \mathrm{S}_n^{\mathrm{Sym}_n}$ only if τ_g is coded.*

Proof. First we prove necessity. Take any $g \in \mathrm{S}_n^{\mathrm{Sym}_n}$ and assume without loss that in fact $g \in \mathrm{S}_n$. Since $\mathbb{N}$ is strong, there is $k_g \in M$ such that for all finite i, either $\tau_g(i) \in \mathbb{N}$ or $\tau_g(i) > k_g$. Now choose some sufficiently small nonstandard $a \in M$ and choose c such that for all $i > 0$ we have $[c]_i = \sigma_g(i)$ when $\sigma_g(i) < k_g$ and $[c]_i = a$ otherwise. We can use the fact that $\mathbb{N}$ is strong in a similar manner to determine whether the order of g is finite and set $[c]_0 = 0$ or $[c]_0 = 1$ accordingly. Then τ_g is coded.

Now we prove sufficiency. Let $f: M \to M$ be a definable function. Choose $a \in M \smallsetminus \mathbb{N}$ such that $\sum_{i=1}^{a} i f(i) < n$. Let $g \in \mathrm{S}_n$ have internal cycle type σ_g where $\sigma_g(i) = f(i)$ for all $i < a$ and $\sigma_g(i) = 0$ for all $i \geq a$. By hypothesis, τ_g is coded, so

$$m\colon r \mapsto \min_{i<r} \left\{ \sigma_g(i) \ : \ \tau_g(i) = \aleph_0 \right\}$$

is definable. Then $\{m(r) \ : \ r \in \mathbb{N}\}$ is a coded decreasing ω-sequence which cannot converge to $\mathbb{N}$ by overspill. Thus there is some $k_f \in M$ such that

$$\mathbb{N} < k_f < \inf\{f(i) \ : \ f(i) \notin \mathbb{N}\}.$$

This suffices to complete the proof. $\square$

Chapter 5

Normal subgroups of some nonstandard permutation groups

The classification of normal subgroups of finite symmetric and alternating groups is a well known undergraduate exercise. The corresponding classification for countable symmetric groups was obtained by Schreier and Ulam[36] and improved to a classification for infinite groups by Baer[4], the latter often being referred to as the theorem of Baer, Schreier and Ulam, a convention that we shall echo in the sequel. The main subject of this chapter is the corresponding results for the group structures obtained from nonstandard finite symmetric groups in models of Peano arithmetic.

5.1 Normal subgroups of nonstandard symmetric groups

The purpose of this section is to show the analogue of the following theorem(s) of Baer, Schreier and Ulam for nonstandard symmetric groups. Later we will show the necessary modifications to obtain a similar result for nonstandard alternating groups.

The statement of the theorem is taken from Bhattarcharjee *et al.*[6], the notation

$$\mathrm{BS}(X,\lambda) = \{g \in \mathrm{Sym}(X) \,:\, \deg(g) < \lambda\}$$

is common in the modern literature, though we shall introduce and use a more general and flexible notational mechanism for this shortly.

Theorem 5.1.1. *Let Ω be a set of cardinality κ and let H be a normal subgroup of* $\mathrm{Sym}(\Omega)$. *Then H is one of the groups in the chain*

$$1 \lhd \mathrm{Alt}(\kappa) \lhd \mathrm{BS}(\Omega, \aleph_0) \lhd \ldots \lhd \mathrm{BS}(\Omega, \kappa) \lhd \mathrm{Sym}(\Omega).$$

Furthermore, all of these normal subgroups do occur.

Analogous to the subgroups of bounded support that appear in the above theorem, we define the following. (The different notation is motivated by the need, in this chapter and the next, to be able to talk of bounded subgroups of different classes of group: not merely bounded subgroups of a symmetric group.)

Definition 5.1.2. Let G be an arbitrary $\mathfrak{M}$-finite permutation group acting on an $\mathfrak{M}$-finite set Ω and let I be an arbitrary initial segment of $\mathfrak{M}$. We define

$$G^{[I]} = \{g \in G \,:\, \deg(g) \in I\}\,.$$

One should note that the proof, found for example in Stewart[40], of the fact that A_n is simple for all $n \geq 5$ is easily translated into the language of PA, so that there are no nontrivial $\mathfrak{M}$-finite normal subgroups of A_n: the normal subgroups we are interested in are the non-$\mathfrak{M}$-finite ones. The analogue of Baer, Schreier and Ulam's theorem that we will prove is the following.

Theorem 5.1.3. *Let $n \in M$ be nonstandard and let I be a cut of n closed under addition. Then $\mathrm{S}_n^{[I]} \lhd \mathrm{S}_n$, $\mathrm{A}_n^{[I]} \lhd \mathrm{S}_n$ and together with A_n these comprise all the normal subgroups of S_n. Furthermore, distinct cuts closed under addition yield distinct normal subgroups and all of these possible subgroups do occur.*

We will prove the analogue for alternating groups in the next section. We begin the demonstration of this result with a few lemmas.

Lemma 5.1.4. *Let $g \in \mathrm{S}_n$ be a k-cycle with $k \geq 3$ and let $t \leq \lfloor k/3 \rfloor$. Then there is $h \in \mathrm{S}_n$ such that $\mathrm{supp}(h) \subseteq \mathrm{supp}(g)$, $\deg(gg^h) = 3t$ and gg^h has order 3.*

Proof. Let a code the cycle so that

$$g = ([a]_0, [a]_1, \ldots, [a]_{k-1}).$$

Let

$$h' = ([a]_1, [a]_{k-1}) \ldots ([a]_r, [a]_{k-r})$$

where $2r \in \{k-2, k-3\}$. We know that such h' exists in S_n by the following inductive construction. Construct the sequence

$$[([a]_1, [a]_{k-1}), ([a]_2, [a]_{k-2}), \ldots, ([a]_r, [a]_{k-r-1})]$$

using induction and basic properties of coding, that is, applying sentence 2.4.3 to start the sequence and sentence 2.4.4 to add elements to the sequence at each step of the induction. One can then form the sequence

$$[([a]_1, [a]_{k-1}), ([a]_1, [a]_{k-1})([a]_2, [a]_{k-2}), \ldots, ([a]_1, [a]_{k-1}) \ldots ([a]_r, [a]_{k-r-1})],$$

again using induction and sentences 2.4.3 and 2.4.4, and the desired element of S_n is the $(r-1)$st element of this sequence. We leave subsequent applications of such arguments to the reader.

Now let

$$h'' = ([a]_{k-1}, [a]_{k-2})([a]_{k-4}, [a]_{k-5}) \dots ([a]_{k-3t+2}[a]_{k-3t+1}),$$

employing a similar construction to that just applied, and observe that $h = h'h''$ is the required element of S_n, for

$$\begin{aligned} gg^h &= ([a]_0, [a]_1, \dots, [a]_{k-1})([a]_0, [a]_1, \dots, [a]_{k-1})^{h'h''} \\ &= ([a]_0, \dots, [a]_{k-1})([a]_0, [a]_{k-1}, [a]_{k-2}, \dots, [a]_1)^{h''} \\ &= ([a]_0, \dots, [a]_{k-1})([a]_0, [a]_{k-2}, [a]_{k-1}, \dots, [a]_1)^{h''} \\ &= ([a]_{k-3t}, [a]_{k-3t+2}, [a]_{k-3t+1}) \dots ([a]_{k-3}, [a]_{k-1}, [a]_{k-2}) \end{aligned}$$

and this has $\mathrm{supp}(gg^h) \subseteq \mathrm{supp}(g)$ and $\deg(gg^h) = 3t$ as required. $\square$

To illustrate with finite permutations: if

$$g = (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$$

then (with h' as given in the above proof)

$$g^{h'} = g^{-1} = (0, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1).$$

For $t = 1$, we take $h'' = (11, 12)$ so that

$$g^{h'h''} = (0, 11, 12, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1).$$

and

$$gg^{h'h''} = (10, 12, 11).$$

On the other hand, for $t = 3$ we take

$$h'' = (11, 12)(8, 9)(5, 6)$$

whence

$$g^{h'h''} = (0,11,12,10,8,9,7,5,6,4,3,2,1).$$

and

$$gg^{h'h''} = (4,6,5)(7,9,8)(10,12,11).$$

We now proceed by applying the above result to each cycle in the disjoint cycle notation of an element $g \in \mathrm{S}_n$. First we define notation for the maximum number of 3-cycles possible in the resulting element of S_n.

Definition 5.1.5. We define the *3-cycle capacity* $\chi(g)$ of an element g of S_n via

$$\chi(g) = \sum_{i=3}^{n} \lfloor i/3 \rfloor \, \sigma_i(g).$$

Lemma 5.1.6. *Let* $g \in \mathrm{S}_n \smallsetminus \{1_n\}$ *satisfy* $\mathrm{supp}(g) = \mathrm{supp}(g^2)$ *and let* $t \leq \chi(g)$. *Then there is* $h \in \mathrm{S}_n$ *such that* $\mathrm{supp}(h) \subseteq \mathrm{supp}(g)$, gg^h *has order* 3 *and* $\deg(gg^h) = 3t$.

Proof. The first step is to write $g = g_1 g_2 g_3$ where g_2 is a single cycle and g_1 and g_3 are possibly the identity, such that $\chi(g_1) < t$ but $t \leq \chi(g_1 g_2)$. We can easily prove

$$\mathsf{PA} \vdash \forall n \left(n < 2 \vee \forall g {\in} \mathrm{S}_n \, \exists x \left(|x| = \left|\kappa_g\right| \wedge [x]_0 = 1_n \wedge \forall y < \left|\kappa_g\right| \left([x]_{y+1} = [x]_y \cdot [\kappa_g]_y\right)\right)\right)$$

by induction so that

$$(\mathfrak{M}, g, n) \vDash \exists x \left(|x| = \left|\kappa_g\right| \wedge [x]_0 = 1_n \wedge \forall y < \left|\kappa_g\right| \left([x]_{y+1} = [x]_y \cdot [\kappa_g]_y\right)\right).$$

Let $c \in M$ be a witness to this existential statement. We can then take $g_1 = [c]_a$ where

$$a + 1 = (\mu x < n) \left(\chi\left([c]_x\right) \geq t\right).$$

Then we take $g_2 = [\kappa_g]_{a+1}$ and take $g_3 = g_2^{-1} g_1^{-1} g$.

The second step is to process g_1. Form the sequence of length a, coded by d, such that for each $i < a$, $[d]_i$ is the 'h' obtained by applying Lemma 5.1.4 to $[\kappa_g]_i$ such that, where $k = [\kappa_g]_i$ and $l = [d]_i$, $\deg(kk^l) = 3\chi(k)$. (Each 'h' exists by Lemma 5.1.4 and sentences 2.1.3 and 2.1.4 permit us to form the sequence via an easy inductive argument.) Let h_1 be the product of the elements of the ($\mathfrak{M}$-finite) sequence coded by d.

Next we apply Lemma 5.1.4 to g_2 to get h_2 such that

$$\deg(g_2 g_2^{h_2}) = 3\,(t - \chi(g_1))\,.$$

The reader should take note that $t - \chi(g_1)$ is the remaining number of 3-cycles required after g_1 has been processed.

Now let h_3 take g_3 to g_3^{-1} and be chosen such that $\mathrm{supp}(h_3) \subseteq \mathrm{supp}(g_3)$ and finally observe that $h = h_1 h_2 h_3$ is the required element of S_n. $\square$

Lemma 5.1.7. *Let $g \in \mathrm{S}_n$ have order 3 and suppose $4\deg(g) \leq 3n$. Let X be any $\mathfrak{M}$-finite subset of $[0, n-1]$ disjoint from $\mathrm{supp}(g)$ such that $3\,|X| = \deg(g)$. Then there is $h \in \mathrm{S}_n$ such that gg^h has order 2, $3\deg(gg^h) = 4\deg(g)$, $\mathrm{supp}(h) \subseteq \mathrm{supp}(g) \cup X$ and $\mathrm{supp}(gg^h) = \mathrm{supp}(g) \cup X$.*

Proof. Let x be a code for X. Observe that for any $a, b, c, d \in \mathrm{S}_n$ pairwise distinct we have $(a,b,c)(a,b,d) = (a,d)(b,c)$. Let $\{(a_i, b_i, c_i) \,:\, 0 \leq i < \deg(g)/3\}$ be an $\mathfrak{M}$-finite enumeration of the 3-cycles of g. Take

$$h = (c_0, [x]_0)(c_1, [x]_1) \cdots (c_r, [x]_r)$$

where $r = \deg(g)/3$. Then $\mathrm{supp}(h) \subseteq \mathrm{supp}(g) \cup X$ and

$$gg^h = (a_0, [x]_0)(b_0, c_0)(a_1, [x]_1)(b_1, c_1) \ldots (a_r, [x]_r)(b_r, c_r)$$

whereupon gg^h has order 2 and

$$\mathrm{supp}(gg^h) = \mathrm{supp}(g) \cup X$$

which completes the proof. □

Lemma 5.1.8. *Let $g \in \mathrm{S}_n$ have order 2. Then for every $t \leq \lfloor \deg(g)/4 \rfloor$ there is $h \in \mathrm{S}_n$ with $\mathrm{supp}(h) \subseteq \mathrm{supp}(g)$ such that gg^h is of order 2 with $\deg(gg^h) = 4t$ and $\mathrm{supp}(gg^h) \subseteq \mathrm{supp}(g)$.*

Proof. First recall that

$$(a,b)(c,d)\,((a,b)(c,d))^{(b,c)} = (a,d)(b,c)$$

whereas $((a,b)(c,d))^2 = 1$. We use this fact inductively.

Let $t \leq \lfloor \deg(g)/4 \rfloor$ and let $\{(a_i, b_i) : 2i \leq \deg(g)\}$ be an enumeration of the 2-cycles in g. But then

$$h = (b_0, a_1)(b_2, a_3) \cdots (b_{2t}, a_{2t+1})$$

is the required element of S_n. □

These lemmas are then brought together in the main lemma of this section. It is this lemma that does the work for the main theorem, which we will come to shortly. (The '$g_2 g_2'$' notation utilised in the following is roughly analogous to the g_p, g_p' notation for the p and p-prime parts of an element of a group, a notational mechanism that is reasonably common in group theory. Here g_2 represents the 'involution part', that is the collection of 2-cycles in the disjoint cycle form of g and g_2' represents everything else.)

Lemma 5.1.9. *Let $g \in \mathrm{S}_n$ and let $g = g_2 g_2'$ where g_2 has order 2, g_2 and g_2' have disjoint support and* $\mathrm{supp}(g_2') = \mathrm{supp}({g_2'}^2)$. *Then there are $h_2, h_2', k_2, k_2' \in \mathrm{S}_n$ such that, where $l_2 = g_2 g_2^{h_2}$ and $l_2' = g_2' {g_2'}^{h_2'}$, we have that:*

1. $\mathrm{supp}(h_2), \mathrm{supp}(k_2) \subseteq \mathrm{supp}(g_2)$; $\mathrm{supp}(h_2') \subseteq \mathrm{supp}(g_2')$; $\mathrm{supp}(k_2') \cap \mathrm{supp}(g_2) = \varnothing$;
2. $\deg(l_2 l_2^{k_2}) = 4\lfloor \deg(g_2)/4 \rfloor$;
3. $\deg(l_2' {l_2'}^{k_2'}) = \min\left(4\left\lfloor \frac{n-\deg(g_2)}{4} \right\rfloor, 4\chi(g_2')\right)$; *and*
4. $l_2 l_2^{k_2}$ *and* $l_2' {l_2'}^{k_2'}$ *are involutions.*

Thus, where $h = h_2 h_2'$, $k = k_2 k_2'$ and $l = l_2 l_2'$, $ll^k = (gg^h)(gg^h)^k$ is an involution with

$$\deg(ll^k) = 4\left\lfloor \frac{\deg(g_2)}{4} \right\rfloor + \min\left(4\left\lfloor \frac{n-\deg(g_2)}{4} \right\rfloor, 4\chi(g_2')\right).$$

Proof. Apply Lemma 5.1.8 to g_2 to obtain h_2 and then to $g_2 g_2^{h_2}$ to obtain k_2 such that $\deg(l_2) = 4\lfloor \deg(g_2)/4 \rfloor$. (The importance of this step is to maximise the degree of the resulting permutation t by preventing all but possibly one of the 2-cycles from being cancelled when taking the product of the conjugates of g.)

Now apply Lemma 5.1.6 to g_2' to obtain h_2' such that l_2' has order 3 and $\deg(l_2') = 3t$ where t is maximal subject to $4t \leq n - \deg(g_2)$. This last condition allows us to choose an $\mathfrak{M}$-finite subset X of $[0, n-1]$ of the appropriate size which is disjoint from both $\mathrm{supp}(g_2)$ and $\mathrm{supp}(l_2')$.

Then apply Lemma 5.1.7, with the X chosen as in the previous paragraph, to l_2' to get k_2'. One may then check that taking $h = h_2 h_2'$ and $k = k_2 k_2'$ suffices to prove the result. That

$$\deg(ll^k) = 4\left\lfloor \frac{\deg(g_2)}{4} \right\rfloor + \min\left(4\left\lfloor \frac{n-\deg(g_2)}{4} \right\rfloor, 4\chi(g_2')\right)$$

follows from the choice of t in the application of Lemma 5.1.7. $\square$

It should be noted that any permutation resulting from the above lemma is necessarily an even permutation. Now we come to the proof of the main theorem of this section.

Proof of Theorem 5.1.3. That $\mathrm{A}_n^{[I]} \lhd \mathrm{S}_n$ and $\mathrm{S}_n^{[I]} \lhd \mathrm{S}_n$ is due to the fact that acting upon a set does not change its cardinality. Thus if $\deg(g) \in I$ then $\deg(g^h) \in I$ for all $h \in H$.

It is clear that every normal subgroup listed does indeed occur: for consider a k-cycle of the appropriate size. Thus it remains to show that there can be no other normal subgroups.

Let $H \lhd \mathrm{S}_n$ be nontrivial and let $g \in H$. By Lemma 5.1.9 there is an even involution $t \in H$. Furthermore, if $\deg(g)$ is nonstandard then we necessarily have

$$4\lfloor \deg(g)/5 \rfloor - 2 \leq \deg(t).$$

The bound in the above inequality is attained precisely in case g has cycle type $5^a \cdot 2$ for some $a \in M$.

By conjugating and multiplying t and applying Lemma 5.1.8 as necessary we see that H contains all *even* involutions $u \in \mathrm{S}_n$ subject to $\deg(u) \in \deg(g) \cdot \mathbb{N}$. Now any even permutation in S_n is a product of at worst 3 even involutions: every element and hence every even element is a product of two involutions and if these two involutions happen to be odd then we write

$$g = (g_1(a,b))\,((a,b)(c,d))\,((c,d)g_2)$$

where either $a,b,c,d \notin \mathrm{supp}(g_1) \cup \mathrm{supp}(g_2)$ or else (a,b) is a cycle in g_1 and (c,d) is a cycle in g_2. Thus

$$\mathrm{A}_n^{[\deg(g)\cdot\mathbb{N}]} \subseteq H$$

and if H contains any odd permutations then also

$$\mathrm{S}_n^{[\deg(g)\cdot\mathbb{N}]} \subseteq H.$$

But then

$$H = \bigcup_{h \in H} \mathrm{S}_n^{[\deg(h)\cdot\mathbb{N}]}$$

or else

$$H = \bigcup_{h \in H} \mathrm{A}_n^{[\deg(h)\cdot\mathbb{N}]}$$

according to whether H contains odd permutations or not and this suffices to complete the proof. $\square$

5.2 Normal subgroups of nonstandard alternating groups

The purpose of this section is to show how the arguments of the previous section may be modified to produce an analogous result for nonstandard alternating groups. The analogous theorem is as follows.

Theorem 5.2.1. *Let $n \in M$ be nonstandard and let I be a cut of $[0, n-1]$ closed under addition. Then $\mathrm{A}_n^{[I]} \lhd \mathrm{A}_n$ and these comprise all the normal subgroups of A_n. Furthermore, distinct cuts closed under addition yield distinct normal subgroups and all of these possible subgroups do occur.*

The main difficulty in adapting the lemmas is dealing with the parity of the various conjugating elements. As we shall see, these can be fixed reasonably systematically.

Lemma 5.2.2. *Let $g \in \mathrm{A}_n$ be a nonidentity element such that $\mathrm{supp}(g) = \mathrm{supp}(g^2)$ and let $t \leq \chi(g)$. Then there is $h \in \mathrm{A}_n$ and $\varepsilon \in \{0, 1\}$ such that $\mathrm{supp}(h) \subseteq \mathrm{supp}(g)$, gg^h has order 3 and $\mathrm{supp}(gg^h) = 3(t - \varepsilon)$.*

Proof. Recall the proof of Lemma 5.1.6. We construct involutions h' and h'' such that $h = h'h''$ has the required property. If the resulting h is odd, simply remove one of the 2-cycles from h'' and this will suffice. $\square$

The following lemmas are modifications of the ones in the previous section. The inequality in the first of what follows has been replaced with $4\deg(g) \leq 3n-2$ so as to ensure that we have two 'spare' points a and b so as to be able to use the transposition (a,b) to fix any parity problems.

Lemma 5.2.3. *Let $g \in \mathrm{A}_n$ have order 3 and suppose that $4\deg(g)/3 \leq n-2$. Let X be any $\mathfrak{M}$-finite subset of $[0,n-1]$ disjoint from $\mathrm{supp}(g)$ such that $3\,|X| = \deg(g)$ and let Y be any 2-element subset of $[0,n-1]$ disjoint from $\mathrm{supp}(g) \cup X$. Then there is $h \in \mathrm{A}_n$ such that gg^h has order 2, $\deg(gg^h) = 4\deg(g)/3$, $\mathrm{supp}(gg^h) \subseteq \mathrm{supp}(g) \cup X$ and $\mathrm{supp}(h) \subseteq \mathrm{supp}(g) \cup X \cup Y$.*

Proof. Apply Lemma 5.1.7 to get some possibly odd $h' \in \mathrm{S}_n$ which, apart from being possibly odd, has the required property. If h' is even we take $h = h'$ and we are done. If not, let the two elements of Y be c and d and take $h = h'(c,d)$. The reader may check that this suffices. $\square$

Lemma 5.2.4. *Let $g \in \mathrm{A}_n$ have order 2. Then for every $t \leq \lfloor \deg(g)/8 \rfloor$ there is $h \in \mathrm{A}_n$ with $\mathrm{supp}(h) \subseteq \mathrm{supp}(g)$ such that gg^h is of order 2 with $\deg(g) = 8t$.*

Proof. Again use the fact that $((a,b)(c,d))^2 = 1_n$ whereas

$$(a,b)(c,d)\,((a,b)(c,d))^{(b,c)} = (a,d)(b,c).$$

This time, however, we must preserve 2-cycles four at a time to ensure that h is even. $\square$

The main lemma of the previous section then becomes the following.

Lemma 5.2.5. *Let $g \in \mathrm{A}_n$ and let $g = g_2 g_2'$ where g_2 has order 2 and g_2' has no 2-cycles. Then there are permutations $h_2, h_2', k_2, k_2' \in \mathrm{A}_n$ such that, where $l_2 = g_2 g_2^{h_2}$ and $l_2' = g_2' {g_2'}^{h_2'}$, we have:*

1. $\mathrm{supp}(h_2), \mathrm{supp}(k_2) \subseteq \mathrm{supp}(g_2)$; $\mathrm{supp}(h_2'), \subseteq \mathrm{supp}(g_2')$; *and* $\mathrm{supp}(k_2') \cap \mathrm{supp}(g_2) = \varnothing$;

2. $\deg(l_2 {l_2}^{k_2}) = 8 \lfloor \deg(g_2)/8 \rfloor$;

3. $\deg(l_2' {l_2'}^{k_2'}) = \min\left(8 \left\lfloor \frac{n - \deg(g_2)}{8} \right\rfloor, 8 \left\lfloor \frac{\chi(g_2')}{2} \right\rfloor\right)$; *and*

4. $l_2 l_2^{k_2}$ *and* $l_2' {l_2'}^{k_2'}$ *are involutions with disjoint support.*

Thus, where $h = h_2 h_2'$, $k = k_2 k_2'$ and $l = l_2 l_2'$, $ll^k = (gg^h)(gg^h)^k$ is an involution with

$$\deg(ll^k) = 8 \left\lfloor \frac{\deg(g_2)}{8} \right\rfloor + \min\left(8 \left\lfloor \frac{n - \deg(g_2)}{8} \right\rfloor, 8 \left\lfloor \frac{\chi(g_2')}{2} \right\rfloor\right).$$

Proof. Apply Lemma 5.2.4 to g_2 to get h_2 and again to $g_2 g_2^{h_2}$ to get k_2. Apply Lemma 5.2.2 to g_2' within $\mathrm{S}_n(\Omega_n \setminus \mathrm{supp}(g_2))$ to obtain h_2', choosing t maximal subject to $12t \leq 3n - \deg(g_2) - 2$. Then apply Lemma 5.2.3 to the resulting $g_2' {g_2'}^{h_2'}$ to obtain k_2'. One can then check that this suffices. □

Now we come to the proof of the theorem. As the reader will observe, the required modifications to the proof of the theorem of the previous section are minimal.

Proof of Theorem 5.2.1. We know that $\mathrm{A}_n^{[I]} \lhd \mathrm{A}_n$ and that all these normal subgroups are distinct. To see that there are no others, let H be a nontrivial normal subgroup of A_n. By Lemma 5.2.5 there is an even involution $t \in H$ such that

$$4 \left\lfloor \frac{\deg(g)}{5} \right\rfloor - 2 \leq \deg(t).$$

Importantly, $\deg(t)\cdot\mathbb{N} = \deg(g)\cdot\mathbb{N}$. By conjugating, multiplying and applying Lemma 5.2.4 as necessary we see that H contains all even involutions $u \in \mathrm{A}_n$ subject only to $\deg(u) \in \deg(g)\cdot\mathbb{N}$. But every element of A_n is a product of at most 3 even involutions, whence

$$\mathrm{A}_n^{[\deg(h)\cdot\mathbb{N}]} \leq H$$

and thus we may write

$$H = \bigcup_{h\in H} \mathrm{A}_n^{[\deg(h)\cdot\mathbb{N}]}$$

which suffices to complete the proof. □

5.3 Supplements of normal subgroups of nonstandard symmetric groups

Supplements of bounded subgroups of infinite symmetric groups $\mathrm{Sym}(\kappa)$ were studied by Macpherson and Neumann[30] as part of a study of the subgroup structure of infinite symmetric groups. In the paper just cited they claim a proof of the following result (though as Bigelow[7] points out, Macpherson and Neumann's proof makes certain tacit assumptions regarding cardinals.) We denote by G^Δ the permutation group on Δ induced by $G_{\{\Delta\}}$ and for $g \in G_{\{\Delta\}}$ we denote by g^Δ the permutation induced on Δ by g. For the purposes of this section let Ω be a set of cardinality κ where $\kappa \geq \aleph_0$ and let $S = \mathrm{Sym}(\Omega)$. We adopt the notation for bounded subgroups introduced earlier in this chapter, adapting it to full symmetric groups by writing $\mathrm{Sym}^{[\lambda]}(X)$ for the group of all permutations on X with support of cardinality less than λ and $S^{[\lambda]}$ when $X = \Omega$.

Theorem 5.3.1 (Theorem 1.2 of Macpherson and Neumann[30]). *Suppose that* $\aleph_0 \leq \lambda \leq \kappa$ *and that* G *is a subgroup of* S. *Then* $S^{[\lambda]}G = S$ *if, and only if there exists* $\Delta \leq \Omega$ *such that* $\operatorname{card}\Delta < \lambda$ *and*

$$G^{(\Omega\setminus\Delta)} = \operatorname{Sym}(\Omega\setminus\Delta).$$

Their proof relies on infinite combinatorics and does not readily adapt to the non-standard case and it is the purpose of this section to see what may be salvaged for non-standard symmetric groups. As such, we shall break down the proof into steps, looking at adapting each step in turn. The proof relies on the following two lemmas.

Lemma 5.3.2 (Theorem 1.1 of Macpherson and Neumann[30]). *If* $(H_j)_{j\in J}$ *is a chain of proper subgroups of* S *such that*

$$\bigcup_{j\in J} H_j = S$$

then $\operatorname{card}(J) > \kappa$.

Lemma 5.3.3 (Lemma 2.1 of the same paper). *If* $\Gamma_1, \Gamma_2 \subseteq \Omega$ *and* $\operatorname{card}(\Gamma_1 \cap \Gamma_2) = \min\{\operatorname{card}(\Gamma_1), \operatorname{card}(\Gamma_2)\}$ *then*

$$\operatorname{Sym}(\Gamma_1 \cup \Gamma_2) = \langle \operatorname{Sym}(\Gamma_1), \operatorname{Sym}(\Gamma_2)\rangle.$$

Whether or not (and how) S_n can be expressed as a union of a chain of proper subgroups is a good question. In any case, one should not hold out hope of an analogous result.

The 'if' part of the proof of Theorem 5.3.1 is the easy direction and is covered by the first paragraph of Macpherson and Neumann's proof. The proof of the hard direction is given in steps and we shall dissect these steps later. (The steps are taken from the version of the proof given in lecture notes on infinite permutation groups taken by my supervisor from lectures given by P. M. Neumann.) We denote $B = S^{[\lambda]}$.

Proof of Theorem 5.3.1 in easy direction. Suppose that Δ exists as in the assertion. Let $x \in S$. Since $\operatorname{card}(\Delta) < \lambda$ there exists $g_1 \in B$ such that $g_1 \restriction \Delta = x \restriction \Delta$. Then xg_1^{-1} fixes Δ pointwise. By assumption there exists $g_2 \in G_{\{\Delta\}}$ such that

$$g_2 \restriction (\Omega \setminus \Delta) = xg_1^{-1} \restriction (\Omega \setminus \Delta).$$

Then $xg_1^{-1}g_2^{-1}$ fixes $\Omega \setminus \Delta$ pointwise so $g_3 = xg_1^{-1}g_2^{-1} \in B$ and

$$x = g_3 g_2 g_1 \in BGB = BG$$

since B is a normal subgroup of S. □

Note that the final step of the above proof is where we require that B is closed under composition and hence that B is a subgroup. We do not require the assumption that G is actually a subgroup and this should be borne in mind: it may be possible to prove a related result for some class of subsets in place of G and then deduce facts about subgroups subsequently.

Definition 5.3.4. An infinite subset X of Y is a *moiety* iff both X and $Y \setminus X$ have the same cardinality as Y itself.

Proof in hard direction. Suppose that $S = BG$.

Step 1. There exists a moiety Σ_0 on Ω such that

$$G^{\Sigma_0} = \mathrm{Sym}(\Sigma_0).$$

Proof of step 1. Let $(\Sigma_j)_{j \in J}$ be a family of κ pairwise disjoint moieties of Ω. Suppose, seeking a contradiction, that

$$G^{\Sigma_j} \neq \mathrm{Sym}(\Sigma_j) \quad \text{for all } j.$$

Choose $z_j \in \mathrm{Sym}(\Sigma_j)$ not induced by G. Let $z = \prod_j z_j$. Since $S = BG$ there exists $x \in B$ and $y \in G$ such that $z = xy$. But $\deg(x) < \lambda$ and so $\Sigma_j \subseteq \mathrm{fix}(x)$ for all but at most $\deg(x)$ members of J. Choose $j_0 \in J$ such that $\Sigma_{j_0} \subseteq \mathrm{fix}(x)$. Then

$$z^{\Sigma_{j_0}} = y^{\Sigma_{j_0}} = z_{j_0} \in G^{\Sigma_{j_0}}$$

contradicting the fact that z_{j_0} is not induced by G.

Comments on proof of step 1. The proof above does not actually require that Ω (or some subset of it) is partitioned into as many as $n = \mathrm{card}\,(\Omega)$ moieties, only that there are more than $\deg(x) < \lambda$ moieties.

It is important that we can both choose $z_j \in \mathrm{Sym}(\Sigma_j)$ not induced by G *and* that we can then multiply the z_j together. In any adaptation to the nonstandard case we would require that the sequence of z_j be $\mathfrak{M}$-finite if any progress is to be made. This places certain conditions upon G further to $BG = S$ in order for this proof strategy to proceed. Note that we do not require that G is actually a group here, so one could potentially replace G by some $\mathfrak{M}$-finite superset, or even some $\mathfrak{M}$-finite subset provided that $S = BG$.

Step 2. There exists a moiety Σ_1 of Ω and a subgroup $H \leq G_{\{\Sigma_1\}}$ such that

$$H^{\Sigma_1} = \mathrm{Sym}(\Sigma_1)$$

and

$$H^{(\Omega \smallsetminus \Sigma_1)} \leq \mathrm{Sym}^{[\lambda]}(\Omega \smallsetminus \Sigma_1).$$

Proof of step 2. Let Σ_0 be as in step 1. Let $t \in \mathrm{Sym}(\Sigma_0) \leq S$ with t having cycle type 2^κ on Σ_0. Since $S = BG$ there exists $x \in B$ and $y \in G$ such that $t = xy$. Let

$$\Gamma_0 = \mathrm{supp}(x) \cup (\mathrm{supp}(x) \cdot t).$$

Then $\operatorname{card}(\Gamma_0) \leq 2\deg(x) < \kappa$ and $\Gamma_0 t = \Gamma_0$ since t is an involution.

Define $\Sigma_1 = \Sigma_0 \setminus \Gamma_0$. Then Σ_1 is a moiety and $\Sigma_1 \subseteq \Sigma_0$ so

$$G^{\Sigma_1} = \operatorname{Sym}(\Sigma_1)$$

since $G^{\Sigma_0} = \operatorname{Sym}(\Sigma_0)$. Also x fixes Σ_1 pointwise, since $\operatorname{supp}(x)$ is disjoint from Σ_1 by the choice of Σ_1. Thus $y \in G_{\{\Sigma_1\}}$ and

$$y^{\Sigma_1} = x^{\Sigma_1} y^{\Sigma_1} = t^{\Sigma_1}.$$

So y^{Σ_1} is of cycle type 2^κ (in $\operatorname{Sym}(\Sigma_1)$) and so the conjugates of y^{Σ_1} under $\operatorname{Sym}(\Sigma_1)$ generate $\operatorname{Sym}(\Sigma_1)$.

Let

$$H = \langle y \rangle^{G_{\{\Sigma_1\}}}.$$

Then $H^{\Sigma_1} = \operatorname{Sym}(\Sigma_1)$ since $G^{\Sigma_1}_{\{\Sigma_1\}} = \operatorname{Sym}(\Sigma_1)$ and the involutions of $\operatorname{Sym}(\Sigma_1)$ of cycle type 2^κ generate $\operatorname{Sym}(\Sigma_1)$. But $H^{(\Omega \setminus \Sigma_1)}$ is generated by conjugates of $y^{(\Omega \setminus \Sigma_1)}$ and so

$$H^{(\Omega \setminus \Sigma_1)} \leq \operatorname{Sym}^{[\lambda]}(\Omega \setminus \Sigma_1)$$

as required.

Comments on the proof of step 2. The cycle type of t needs to be 2^κ and not that of any smaller involution to ensure that its conjugates generate the full symmetric group on Σ_1 rather than some bounded subgroup; the correct cycle type for the nonstandard case is unclear. The definition of Γ_0 in terms of the support of x ensures that Γ_0 is small with respect to the bound on the degree of elements of B.

The fact that conjugates of y^{Σ_1}, being involutions of maximal degree, generate the full symmetric on Σ_1 is crucial in the proof. In an adaption to the nonstandard setting we must bear in mind the implications of this, namely that there are such generators for the group analogous to $\mathrm{Sym}(\Sigma_1)$. This is equivalent to the condition that there is $a \in M$ such that $\mu(\Sigma_1) = a \cdot \mathbb{N}$, that is, the inner measure of Σ_1 is closed under nothing stronger than addition. This requirement is trivially satisfied by the requiring that Σ_1 is $\mathfrak{M}$-finite. Finally note that the fact that

$$H^{(\Omega \smallsetminus \Sigma_1)} \leq \mathrm{Sym}^{[\lambda]}(\Omega \smallsetminus \Sigma_1)$$

follows because

$$\deg(y^{(\Omega \smallsetminus \Sigma_1)}) \leq \mathrm{card}\,(\Gamma_0) + \deg(x) \leq 3\deg(x) < \lambda.$$

If the boundedness condition is replaced by $\deg(x) \in I$ as would happen in the nonstandard setting, the closure of I under addition, necessary to make $\mathrm{S}_n^{[I]}$ into a subgroup, is sufficient to ensure that

$$\deg(y^{(\Omega \smallsetminus \Sigma_1)}) \in I.$$

Thus far, if we assume that $B = \mathrm{S}_n^{[I]}$ and that G is an $\mathfrak{M}$-finite *set* such that $BG = \mathrm{S}_n$ then these first two steps may be readily adapted. It is at this stage, however, that things will get a little more difficult so far as adapting the current proof is concerned. The main result of Macpherson and Neumann's paper is used at this step and, as such, it represents the point that the nature of infinite combinatorics draws away from the behaviour in the nonstandard case. Let us proceed.

Step 3. There exists a moiety Σ_1 of Ω and a subgroup $K \leq G_{\{\Sigma_1\}}$ such that

$$K^{\Sigma_1} = \mathrm{Sym}(\Sigma_1)$$

and

$$\mathrm{card}\,((\Omega \smallsetminus \Sigma_1) \cap \mathrm{supp}(K)) < \lambda.$$

Proof of step 3. Let Σ_1 be as in step 2. Consider subsets Γ of $\Omega \setminus \Sigma_1$ such that there exists $K \leq G_{\{\Sigma_1\}}$, with K depending on Γ, with

$$\begin{aligned} K^{\Sigma_1} &= \mathrm{Sym}(\Sigma_1) \\ \mathrm{supp}(K) &\leq \Sigma_1 \cup \Gamma \\ K^{\Gamma} &\leq \mathrm{Sym}^{[\lambda]}(\Gamma). \end{aligned}$$

Note that $\Gamma = \Omega \setminus \Sigma_1$ is such a subset, so this class of subsets is not vacuous and hence we may choose Γ of least cardinality μ. Suppose, seeking a contradiction, that $\mu \geq \lambda$. Identify μ with the appropriate initial ordinal and express Γ as a union

$$\bigcup_{\xi<\mu} \Gamma_\xi$$

where

$$\mathrm{card}\,\Gamma_\xi < \mu$$

and

$$\Gamma_{\xi_1} \subseteq \Gamma_{\xi_2} \text{ iff } \xi_1 \leq \xi_2.$$

Define $K_\xi = \left\{ f \in K \,:\, \mathrm{supp}(f) \subseteq \Sigma_1 \cup \Gamma_\xi \right\}$. Then $(K_\xi)_{\xi<\mu}$ is a chain of subgroups of K. Also we have $K = \bigcup_{\xi<\mu} K_\xi$. Therefore $K^{\Sigma_1} = \bigcup_{\xi<\mu}(K_\xi^{\Sigma_1})$. By Lemma 5.3.2 there exists $\eta < \mu$ such that

$$K_\eta^{\Sigma_1} = \mathrm{Sym}(\Sigma_1).$$

Then Γ_η contradicts the minimality of μ. Hence $\mu < \lambda$ and so

$$\mathrm{card}\,((\Omega \setminus \Sigma_1) \cap \mathrm{supp}(K)) = \mathrm{card}\left(\mathrm{supp}(K^{\Gamma})\right) \leq \mu < \lambda$$

as required.

Comments on the proof of step 3. Here we find the first major obstacle to an adaptation of the proof. In the original proof, as we see, this step sets up the application of Lemma 5.3.2, a lemma which does not carry through to the nonstandard setting in any reasonably straightforward way. Thus, if this were the only major sticking point, we would need to supply a fresh approach to proving a corresponding result in the nonstandard setting. The precise nature of the adapted step would be determined by the requirements of subsequent steps and, as we shall shortly see, these latter steps raise problems with the overall strategy so far as a result for nonstandard symmetric groups is concerned.

Step 4. If Σ' is *any* moiety then there exists Σ such that Σ has the property described in step 3, $\Sigma \subseteq \Sigma'$ and $\operatorname{card}(\Sigma' \setminus \Sigma) < \lambda$.

Proof of step 4. Choose $z \in S$ such that $\Sigma_1 z = \Sigma'$. Then $z = xy$ for suitable $x \in B$ and $y \in G$. If $\Sigma = \Sigma' \setminus \operatorname{supp}(x)$ then $\Sigma = \Sigma'' y$ for some $\Sigma'' \subseteq \Sigma_1$ and $\operatorname{card}(\Sigma_1 \setminus \Sigma'') < \lambda$. Then $y^{-1} K_{\{\Sigma''\}} y$ is a suitable 'K' for Σ.

Comments on the proof of step 4 This step relies on the fact that there are permutations in S moving any moiety onto any other. As such, moieties provide a useful class of subsets of Ω to which there is no analogous class of subsets of a nonstandard set Ω_n, for consider the necessary conditions for two nonstandard sets $X, Y \subseteq \Omega_n$ to have an $\mathfrak{M}$-finite bijection between them.

Proof of theorem. Let Σ_1' and Σ_2' be moieties of Ω such that $\Sigma_1' \cap \Sigma_2'$ is still a moiety and $\operatorname{card}(\Omega \setminus (\Sigma_1' \cup \Sigma_2')) < \lambda$.

Choose

$$\Sigma_1'' \subseteq \Sigma_1'$$

$$\Sigma_2'' \subseteq \Sigma_2'$$

such that

$$\operatorname{card}(\Sigma_i' \setminus \Sigma_i'') < \lambda \text{ for } i \in \{1,2\}$$

and there are groups K_1 and K_2 as in step 3.

Choose

$$\Sigma_1 = \Sigma_1' \setminus (\operatorname{supp}(K_2) \cap (\Omega \setminus \Sigma_2))$$

$$\Sigma_2 = \Sigma_2' \setminus (\operatorname{supp}(K_1) \cap (\Omega \setminus \Sigma_1))$$

so that, by taking $L_i = (K_i)_{(\Sigma_i' \setminus \Sigma_i)}$ for $i \in \{1,2\}$, there are subgroups L_1 and L_2 such that

$$L_1 \text{ fixes } \Sigma_2 \setminus \Sigma_1 \text{ pointwise}$$

$$L_2 \text{ fixes } \Sigma_1 \setminus \Sigma_2 \text{ pointwise}$$

$$L_1^{\Sigma_1} = \operatorname{Sym}(\Sigma_1)$$

$$L_2^{\Sigma_2} = \operatorname{Sym}(\Sigma_2).$$

By Lemma 5.3.3 we have that

$$\langle L_1, L_2 \rangle \restriction (\Sigma_1 \cup \Sigma_2) = \operatorname{Sym}(\Sigma_1 \cup \Sigma_2)$$

and

$$\operatorname{card}(\Omega \setminus (\Sigma_1 \cup \Sigma_2)) < \lambda$$

as required.

Comments on the proof of the theorem. This final step is where, in the context of a desired result for nonstandard symmetric groups, the strategy employed here falls apart. If X and Y are *any* subsets of Ω_n such that $|\Omega_n \setminus (X \cup Y)| \in I$ for an initial segment $I < n$ closed under addition then at least one of X and Y is too large: step 1 requires the partition of Ω into a large number of moieties whereas here we require that the union of just two moieties covers most of Ω and the application of step 4 requires that there are bijections between these subsets of apparently disparate sizes (as indeed there are in the infinite case). As such this incompatibility between the various steps indicates that a different strategy is probably required for an analogous result in the nonstandard setting. $\square$

We shall see in due course what results in the nonstandard case can be lifted from the proof above. We introduce a little more notation first.

Definition 5.3.5. Let Ω be an $\mathfrak{M}$-finite set with $n = |\Omega|$. By S_Ω we mean the set of all $\mathfrak{M}$-finite bijections from Ω to itself. Let Δ be some subset of Ω. By $\mathrm{S}_\Omega(\Delta)$ we mean the subset of S_Ω that fixes $\Omega \setminus \Delta$ pointwise. The notation $\mathrm{S}_\Omega^{[I]}(\Delta)$ is defined to be $\{g \in \mathrm{S}_\Omega(\Delta) \ : \ \deg(g) \in I\}$.

In what follows, I will be an initial segment of $\mathfrak{M}$ closed under addition, n will be some arbitrary nonstandard element of M and Ω will be some arbitrary subset of M with $|\Omega| = n$.

Proposition 5.3.6. *Let $S = \mathrm{S}_\Omega$, $G \subseteq \mathrm{S}_\Omega$ and $B = \mathrm{S}_\Omega^{[I]}$. Suppose that there exists an $\mathfrak{M}$-finite subset $\Delta \subseteq \Omega$ with $|\Delta| \in I$ and*

$$G^{(\Omega \setminus \Delta)} = \mathrm{S}_\Omega(\Omega \setminus \Delta).$$

Then $S = BG$.

This is a straightforward adaptation of the proof from the infinite case, but we give it nonetheless—it illustrates the first thing we look to in adapting a proof to the nonstandard case: replacing cardinal bounds by suitably closed initial segments.

Proof. Let $x \in S$. Then since $|\Delta| \in I$ there exists $g_1 \in B$ such that $g_1 \restriction \Delta = x \restriction \Delta$. Then xg_1^{-1} fixes Δ pointwise. By assumption there exists $g_2 \in G_\Delta$ such that

$$g_2 \restriction (\Omega \smallsetminus \Delta) = xg_1^{-1} \restriction (\Omega \smallsetminus \Delta).$$

Then $xg_1^{-1}g_2^{-1}$ fixes $\Omega \smallsetminus \Delta$ pointwise, so is an element of B. Let $g_3 = xg_1^{-1}g_2^{-1}$ and observe that

$$x = g_3 g_2 g_1 \in BGB = BBG = BG$$

since B is a normal subgroup of S. $\square$

Now we have the adaptation of step 1. The reader should observe that it is in the statement of the proposition that most of the adaptation takes place, things then running through much like the earlier proof.

Proposition 5.3.7. *Let S, B and G be as in the previous proposition. Suppose that $S = BG$ and that G is $\mathfrak{M}$-finite. Let a be any element of M such that $ab < n$ for some $b \in M \smallsetminus I$. Then there is an $\mathfrak{M}$-finite subset $\Sigma_0 \subseteq \Omega$ such that $|\Sigma_0| = a$ and*

$$G^{\Sigma_0} = \mathrm{S}_\Omega(\Sigma_0).$$

Proof. Partition Ω, $\mathfrak{M}$-finitely, into b a-subsets and one other subset. Call these a-subsets Σ_j with $J = [0, b-1]$ as the index set. Suppose, seeking a contradiction, that

$$G^{\Sigma_j} \neq \mathrm{S}_\Omega(\Sigma_j)$$

for all $j \in J$. Choose $z_j \in \mathrm{S}_\Omega(\Sigma_j)$ not induced by G for each i such that the sequence $(z_j)_{j \in J}$ is $\mathfrak{M}$-finite. This is possible since the partition is $\mathfrak{M}$-finite, as is G. Let

$$z = \prod_{j \in J} z_j$$

which is necessarily an element of S since the sequence of multiplicands is $\mathfrak{M}$-finite. (We only need that the sequence is $\mathfrak{M}$-finite to take this product.) Since $S = BG$ there is $x \in B$ and $y \in G$ with $z = xy$. But $\deg(x) \in I$ whence not all Σ_j have elements moved by x since there are at least $b \notin I$ subsets Σ_j in the partition. Let Σ_{j_0} be fixed pointwise by x. Then $y \in G$ induces z_{j_0} on Σ_{j_0}, the desired contradiction. $\square$

Note that, in the above proof, we do not necessarily have to partition the entirety of Ω, merely a subset of cardinality ab. Requiring that G is $\mathfrak{M}$-finite is necessary in the above for the sequence to be $\mathfrak{M}$-finite. This alone prevents a result such as Theorem 5.3.1 from being proved using this strategy, but we shall continue a little further.

In the adaptation of step 2 we do not need the $\mathfrak{M}$-finiteness assumption, so we just require that G be any subset of S subject to $S = BG$.

In what follows we wish to show that if the Σ_0 found by the above proposition is large with respect to I then we can refine it to another subset $\Sigma_1 \subseteq \Omega$ that is also large with respect to I and has another useful property.

Proposition 5.3.8. *Let S and B be as in the previous proposition and suppose that G is some subset of S such that $S = BG$ and there is an $\mathfrak{M}$-finite subset Σ_0 of Ω with $|\Sigma_0| \notin I$ and $G^{\Sigma_0} = \mathrm{S}_\Omega(\Sigma_0)$. Then there is an $\mathfrak{M}$-finite subset Σ_1 of Ω with $|\Sigma_1| \notin I$ and a subgroup $H \subseteq G_{\{\Sigma_1\}}$ such that $H^{\Sigma_1} = \mathrm{S}_\Omega(\Sigma_1)$ and $H^{\Omega \setminus \Sigma_1} \leq \mathrm{S}_\Omega^{[I]}(\Omega \setminus \Sigma_1)$.*

Proof. Let $t \in \mathrm{S}_\Omega(\Sigma_0)$ be an involution of maximal degree. (Such t exists because $\mathrm{S}_\Omega(\Sigma_0)$ is $\mathfrak{M}$-finite.) Since $S = BG$ there are $x \in B$ and $y \in G$ such that $t = xy$. Let $\Gamma_0 = \mathrm{supp}(x) \cup (\mathrm{supp}(x) \cdot t)$. Then Γ_0 is $\mathfrak{M}$-finite and $|\Gamma_0| \in I$ since $\deg(x) \in I$, I is closed under addition and $\Gamma_0 t = \Gamma_0$ since t is an involution. Define

$$\Sigma_1 = \Sigma_0 \setminus \Gamma_0.$$

Then Σ_1 is also $\mathfrak{M}$-finite and $\Sigma_1 \subseteq \Sigma_0$. Also $|\Sigma_1| \notin I$ since I is closed under addition (otherwise if $|\Sigma_1| \in I$ then $|\Sigma_0| = |\Sigma_1| + |\Gamma_0| \in I$.) Furthermore, since

$$G^{\Sigma_0} = \mathrm{S}_\Omega(\Sigma_0)$$

it follows that

$$G^{\Sigma_1} = \mathrm{S}_\Omega(\Sigma_1).$$

Now x fixes Σ_1 pointwise, so $y \in G_{\{\Sigma_1\}}$ and

$$y^{\Sigma_1} = t^{\Sigma_1}.$$

So y acts as an involution with $|\Sigma_1 - \mathrm{supp}(y)| \leq \deg(x)$, since t was an involution of maximal degree acting on Σ_0. Since $\deg(x) \in I$ and I is closed under addition, the conjugates of y^{Σ_1} in $\mathrm{S}_\Omega(\Sigma_1)$ still generate $\mathrm{S}_\Omega(\Sigma_1)$ (for observe that $|\Sigma_1|$ is contained in the smallest initial segment of $\mathfrak{M}$ containing $\deg(y^{\Sigma_1})$ and closed under addition—the issue here is that we need to ensure that $\deg(y^{\Sigma_1})$ isn't too small.) Let

$$H = \langle y \rangle^{G_{\{\Sigma_1\}}}.$$

Then

$$H^{\Sigma_1} = \mathrm{S}_\Omega(\Sigma_1)$$

and since $\deg(y^{\Omega \smallsetminus \Sigma_1}) \leq 3\deg(x) \in I$ we see that $H^{\Omega \smallsetminus \Sigma_1}$ is generated by conjugates of elements of degree contained in I, whence

$$H^{\Omega \smallsetminus \Sigma_1} \leq \mathrm{S}_\Omega^{[I]}(\Omega \smallsetminus \Sigma_1).$$

$\square$

Note that the H obtained above need not be $\mathfrak{M}$-finite, and most probably will not be. At issue here is the fact that H is generated externally rather than internally, so is unlikely to be $\mathfrak{M}$-finite. The final proposition that can be lifted from the earlier proof is an adaptation of step 4.

Proposition 5.3.9. *Suppose that S and B are as before and that G is any subset of S such that $S = BG$ and that there is an $\mathfrak{M}$-finite subset $\Sigma_1 \subseteq \Omega$ such that $|\Sigma_1| \in M \smallsetminus I$ and that there is a ($\mathfrak{M}$-finite) subset/subgroup K_1 of $G_{\{\Sigma_1\}}$ such that*

$$K_1^{\Sigma_1} = \mathrm{S}_{\Omega}(\Sigma_1)$$

and

$$\nu\left((\Omega \smallsetminus \Sigma_1) \cap \mathrm{supp}(K_1)\right) \in I.$$

Suppose that Σ' is any $\mathfrak{M}$-finite subset of Ω such that $|\Sigma'| \leq |\Sigma_1|$ and $|\Sigma_1| - |\Sigma'| \in I$. Then there exists a subset Σ of Σ' such that there is a ($\mathfrak{M}$-finite) subset/subgroup K of $G_{\{\Sigma_1\}}$ such that

$$K^{\Sigma} = \mathrm{S}_{\Omega}(\Sigma)$$

and

$$\nu\left((\Omega \smallsetminus \Sigma) \cap \mathrm{supp}(K)\right) \in I.$$

Proof. Since $|\Sigma'| \leq |\Sigma_1|$ we can take $z \in \mathrm{S}_{\Omega}$ such that $\Sigma' \subseteq \Sigma_1 z$. Then since $S = BG$ we can find $x \in B$ and $y \in G$ such that $z = xy$. Let

$$\Sigma = \Sigma' \smallsetminus \mathrm{supp}(x).$$

Then

$$\Sigma = \Sigma'' y \quad \text{for some } \Sigma'' \subseteq \Sigma_1$$

and

$$\left|\Sigma_1 \smallsetminus \Sigma''\right| \in I$$

since $\deg(x) \in I$. Then $K = y^{-1}K_{1\{\Sigma''\}}y$ suffices, for clearly we have

$$((K_1)_{\Sigma''})^{\Sigma''} = \mathrm{S}_\Omega(\Sigma'')$$

whence

$$\left(y^{-1}(K_1)_{\Sigma''}y\right)^{\Sigma} = \mathrm{S}_\Omega(\Sigma)$$

and

$$\nu\left((\Omega \smallsetminus \Sigma) \cap \mathrm{supp}(K)\right) \leq \nu\left(\Omega \smallsetminus \Sigma_1\right) \cap \mathrm{supp}(K_1)) + |\Sigma_1 \smallsetminus \Sigma| \in I$$

as required. □

As such, we have now reached the limits of what can be readily adapted from the proof reproduced earlier.

5.4 The nature of S_n as an extension of its bounded subgroups

This section raises a question that is, in part, motivated by the result looked at in the previous section. As such there are two interrelated questions concerning normal subgroups of nonstandard symmetric groups that we wish to consider here. As in the previous section, let $B = \mathrm{S}_n^{[I]}$ and $S = \mathrm{S}_n$.

1. Is S a split extension of B? That is to say, does there exist $G \leq S$ such that $BG = S$ and $B \cap S = 1$?

2. If $BG = S$, what can be said about G?

As a beginning observation, it seems unlikely that S_n is actually a split extension of B because of the number of involutions necessarily present in G. To see this suppose that

$B \cap G = 1$ and let $t \in S$ be some involution with $\deg(t) \notin I$. We can write $t = xy$ where $x \in B$ and $y \in G$. Now

$$y^2 = x^{-1}tx^{-1}t = x^{-1}x't^2 = x^{-1}x'$$

for some $x' \in B$ since B is a normal subgroup. Thus $y^2 \in B \cap G$ and so y is an involution. But $\deg(t) \notin I$, and so $\deg(y) \notin I$. Furthermore, $yt^{-1} = x^{-1} \in B$ so that y has the same action as t almost everywhere. We see immediately that G must contain a large number of involutions. Bearing in mind that every element of S is a product of two involutions it seems likely that a nonidentity element of B will turn up in G.

Recently, Kaye[2] has shown that if $BG = S$ and $B \cap G = 1$ then G cannot be a monotone limit (a limit of a definable increasing or decreasing sequence of subsets of M.) This result precludes the possiblity of G being $\mathfrak{M}$-finite.

Chapter 6

Normal subgroups of some other nonstandard finite simple groups

The previous chapter covered the normal subgroup structure of nonstandard symmetric and alternating groups. We saw that the initial segment structure of a model $\mathfrak{M}$ of Peano arithmetic underlies the external normal subgroup structure of nonstandard symmetric and alternating groups constructed within $\mathfrak{M}$ in a clearly explicable way. Of particular note was the structure of nonstandard alternating groups, the normal subgroups of which were linearly ordered by inclusion.

These results inspired a search for similar patterns among other nonstandard groups and our search began with nonstandard analogues of other finite simple groups. After elucidating the structure of nonstandard cyclic groups of prime order we proceed to look at the structure of nonstandard projective special linear groups and nonstandard projective symplectic groups. The linear algebra literature enables us to give some useful results for nonstandard special linear and projective linear groups and we conclude this chapter with a brief note on the case of nonstandard projective symplectic groups.

Recall that, in terms of matrix groups, the special linear group of dimension n over a field K is the set of all $n \times n$ matrices over K of determinant 1. The space $V = K^n$ may be equipped with a *symplectic form* $(,) : V^2 \to K$, which is a bilinear alternating form, making V into a *symplectic space*. Recall that this space necessarily has even dimension, so in context of symplectic spaces we write $2n$ rather than n. The symplectic group $\mathrm{Sp}_{2n}(K)$ is the subgroup of $\mathrm{GL}_{2n}(K)$, or equivalently of $\mathrm{SL}_{2n}(K)$, that preserves this form. The reader should be aware that the two parameters n and K determine these groups up to isomorphism.

For more details about special linear groups, the reader may refer to O'Meara[31], for example, and for coverage of the symplectic groups, O'Meara's excellent book[32] will more than suffice.

For the reader unacquainted with the importance of these groups, we offer a brief explanation. The Classification of finite simple groups states that every finite simple group is one of the following:

1. a cyclic group of prime order;

2. an alternating group of degree ≥ 5;

3. a classical group $\mathrm{PSL}_n(K)$, $\mathrm{PSp}_{2n}(K)$, $PSU_n(K)$ or $P\Omega_n^{\varepsilon}(K)$;

4. an exceptional twisted group of Lie type; or

5. a sporadic simple group, of which there are 26.

The reader may refer to Solomon[39] for an overview of the result. Now, of the above: (1) is a straightforward case, the cyclic groups of nonstandard prime order being infinite dimensional rational vector spaces (2) has been dealt with by the authors; for actually finite dimension, the groups in (3) are simple; the groups in (4) are all simple, always having finite dimension; and one can reasonably assume, as we do, that no nonstandard sporadic groups exist. Thus, after the nonstandard cyclic groups of prime order, we are left with the classical groups of nonstandard dimension. The main thrust of this chapter is to begin work on this latter class of nonstandard groups.

6.1 Nonstandard cyclic groups of prime order

The internally simple nonstandard cyclic groups are amongst the easiest to classify structurally. As this brief section notes, externally they are simply rational vector spaces.

Proposition 6.1.1. *Let p be a nonstandard prime. Then C_p has the structure of a vector space over $\mathbb{Q}$ of dimension* $\operatorname{card}({}_{<}p)$.

Proof. For any $n \in \mathbb{N}$ we have $(p,n) = 1$. Thus by the well known result of Euclid there are x and y in M with either $xp - yn = 1$ or else $yn - xp = 1$. We assume that $xp - yn = 1$, the other case being treated in a virtually identical way. In this case we have $yn = xp - 1$ and hence $n(rp - y) \equiv 1 \pmod{p}$ for the r such that $0 \leq rp - y \leq p$. Then n has a multiplicative inverse mod p and it follows that C_p is a divisible abelian group. Divisibility ensures that C_p admits a scalar multiplication by elements of $\mathbb{Q}$ and thus C_p has the structure of a rational vector space.

Now we need to show that the dimension of such a group cannot be finite. Without loss identify C_p with the initial segment of $\mathfrak{M}$ with top element $p - 1$. Suppose that $\{x_1, x_2, \ldots, x_k\}$ is a finite set of generators. Then every element of C_p is of the form $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k$ and thus for every nonstandard l (since the coefficients must be standard rationals) we have

$$\mathfrak{M} \models \forall y{<}p\, \exists a_1, b_1, \ldots, a_k, b_k{<}l\; y = \frac{a_1}{b_1}x_1 + \frac{a_2}{b_2}x_2 + \cdots + \frac{a_k}{b_k}x_k$$

and by underspill there is $l \in \mathbb{N}$ such that

$$\mathfrak{M} \models \forall y{<}p\, \exists a_1, b_1, \ldots, a_k, b_k{<}l\; y = \frac{a_1}{b_1}x_1 + \frac{a_2}{b_2}x_2 + \cdots + \frac{a_k}{b_k}x_k$$

but there are only finitely many $a_i, b_i < l$ and hence only finitely many linear combinations possible given the bound on the numerators and denominators, which is the required contradiction.

Finally, observe that for any $X \subseteq \mathrm{C}_p$ with $\aleph_0 \leq \lambda = \mathrm{card}\,(X)$ we have

$$\lambda \leq \mathrm{card}\,(\mathrm{span}_{\mathbb{Q}} X) \leq \lambda \cdot \aleph_0 = \lambda.$$

Thus $\mathrm{card}\, X = \mathrm{card}\,(\mathrm{span}_{\mathbb{Q}} X)$ and so $\dim_{\mathbb{Q}}(\mathrm{C}_p) = \mathrm{card}\,(\mathrm{C}_p) = \mathrm{card}\,({<}p)$. □

It should be observed that we used the fact that $(p, n) = 1$ for all $n \in \mathbb{N}$ in the above proof, and not the fact that p is a nonstandard prime. Thus we can immediately generalise the previous result to the following, utilising essentially the same proof.

Proposition 6.1.2. *Let q be an element of M with no standard prime divisors. Then, externally, C_q has the structure of a vector space over $\mathbb{Q}$ of dimension* $\mathrm{card}_{<}q$.

6.2 Normal subgroups of nonstandard projective special linear groups

The normal subgroups of A_n were precisely those of bounded support. We shall show in this section that analogous normal subgroups of PSL_n and SL_n exist. Indeed such normal subgroups must arise by considering SL_n as a group of permutations on K^n.

We assume we have fixed a method of coding $\mathfrak{M}$-finite fields and vector spaces, essentially considering vector spaces as K^n with standard bases. There will be various assertions that things can be stated in the language of PA, for which we shall only give occasional hints as to how the actual formulae and sentences in question are formed. In practice these depend on the precise way in which we interpret groups, fields and vector spaces.

For the remainder of this section, fix a possibly nonstandard prime power $q \in M$. We will mention explicitly if any other conditions are required to be placed upon q. Let $n \in M$ be nonstandard, let K be the $\mathfrak{M}$-finite field of order q and let V be an $\mathfrak{M}$-finite vector space of dimension n over K. Denote by $K^\times$ the multiplicative group of non-zero elements of K.

Let $\mathrm{GL}(V)$, $\mathrm{PSL}(V)$ and $\mathrm{SL}(V)$ denote the general linear group, projective special linear group and special linear group respectively, of dimension n over K. Identify these with the matrix groups $\mathrm{GL}_n(K)$, $\mathrm{PSL}_n(K)$ and $\mathrm{SL}_n(K)$ and again fix some coding apparatus to make everything work (*i.e.* so that the domains, composition and inversion functions are all $\mathfrak{M}$-finite.) Since the field is fixed and only the dimension will need to vary here we abbreviate these to GL_n, PSL_n and SL_n respectively.

A transvection is an element of GL_n that is as close as possible to being the identity.

Definition 6.2.1. A *transvection* is an element of GL_n that is conjugate to

$$B_{1,2}(1) = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

An *elementary transvection* (with respect to the standard basis) is an element of GL_n of the form

$$B_{1,2}(\lambda) = \begin{pmatrix} 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

where $\lambda \in K^\times$.

It is well known that SL_n is generated by transvections (see *e.g.* Rotman[35].) The proof there involves nothing more than Gaussian elimination, and can be proved by induction on n. In fact, for $g \in \mathrm{SL}_n$, g is a product of $\dim\{g(x) - x \,:\, x \in V\}$ transvections (see Hahn and O'Meara[16], 2.1.15 for this.) Thus we see that

$$\mathfrak{M} \models \text{“}\mathrm{SL}_n \text{ is generated by transvections”}$$

but it is not the case that (externally), SL_n is generated by transvections. To see this observe that

$$\mathfrak{M} \models \text{“the product of 2 transvections fixes} \\ \text{a subspace of codimension} \leq 2\text{”}$$

and by an obvious induction we see that

$$\mathfrak{M} \models \text{“the product of } n \text{ transvections fixes} \\ \text{a subspace of codimension} \leq n\text{”}$$

so the product of an actually finite number of transvections fixes a subspace of actually finite codimension. Furthermore, by considering the natural embedding of finite dimensional special linear groups in SL_n and its conjugates, every element of SL_n that fixes a subspace of finite codimension is expressible as a finite product of transvections. It is easy to see that if $g \in \mathrm{SL}_n$ fixes a subspace of codimension m, then so does hgh^{-1} for every $h \in \mathrm{SL}_n$. Thus we have an example of a proper nontrivial normal subgroup, namely the set of elements of SL_n expressible as an actually finite product of transvections. Following the obvious generalisation of 'finite codimension' to 'small codimension' we see the following.

Proposition 6.2.2. *Let $I \subseteq_e n$ be closed under addition and define*

$$\mathrm{SL}_n{}^{(I)} = \Big\{ g \in \mathrm{SL}_n \ : g \textit{ fixes a subspace of } K^n \\ \textit{of codimension } m \in I \Big\}.$$

Then $\mathrm{SL}_n^{(I)} \lhd \mathrm{SL}_n$.

Proof. If $g \in \mathrm{SL}_n$ fixes X pointwise then hgh^{-1} fixes hX pointwise. Thus $g \in \mathrm{SL}_n^{(I)}$ if, and only if, $hgh^{-1} \in \mathrm{SL}_n^{(I)}$ for all $h \in \mathrm{SL}_n$ and so we have closure under conjugation.

Now let $g, h \in \mathrm{SL}_n^{(I)}$ and observe that h^{-1} has the same fixed space as h and thus it suffices to check that $gh \in \mathrm{SL}_n^{(I)}$. But

$$\begin{aligned} \dim(\mathrm{fix}(gh)) &\geq \dim(\mathrm{fix}(g) \cap \mathrm{fix}(h)) \\ &\geq \dim(\mathrm{fix}(g)) + \dim(\mathrm{fix}(h)) - \dim(\mathrm{fix}(g) + \mathrm{fix}(h)) \\ &\geq n - (\mathrm{codim}(\mathrm{fix}(g)) + \mathrm{codim}(\mathrm{fix}(h))) \end{aligned}$$

(where $\mathrm{fix}(\cdot)$ denotes the fixed space.) Thus

$$\begin{aligned} \mathrm{codim}(\mathrm{fix}(gh)) &= n - \dim(\mathrm{fix}(gh)) \\ &\leq \mathrm{codim}(\mathrm{fix}(g)) + \mathrm{codim}(\mathrm{fix}(h)) \in I \end{aligned}$$

so $gh \in \mathrm{SL}_n^{(I)}$ as required. □

We see, then, that we have a chain,

$$1 \lhd \mathrm{SL}_n{}^{(\mathbb{N})} \lhd \cdots \lhd \mathrm{SL}_n{}^{(I)} \lhd \cdots \lhd \mathrm{SL}_n{}^{(n/\mathbb{N})} \lhd \mathrm{SL}_n,$$

of normal subgroups of SL_n where I represents an arbitrary initial segment of n closed under addition and, as defined in section 2.3, $n/\mathbb{N}$ is the greatest initial segment of $[0,n]$ closed under addition. Factoring by the centre, SZ_n, of SL_n we obtain a similar chain of normal subgroups of PSL_n.

Definition 6.2.3. We denote the map from subgroups and elements of SL_n to the corresponding elements of PSL_n by $\mathcal{P}$ and, for an arbitrary initial segment I of $\mathfrak{M}$ containing n, we define

$$\mathrm{PSL}_n^{(I)} = \mathcal{P}(\mathrm{SL}_n{}^{(I)}).$$

It is conceivable, for example when $|K|-1$ divides n, that these are all the normal subgroups of PSL_n, but this is generally not the case. The motivation here is the result of Gardener[15] that for a countably infinite dimensional symplectic space V the following is a composition series for $\mathrm{Aut}(V)$:

$$1 \lhd \mathcal{F} \lhd \mathcal{E} \lhd \mathrm{Aut}(V), \tag{6.2.1}$$

where $\mathcal{F}$ is the subgroup of elements possessing a fixed space of finite codimension and $\mathcal{E}$ is the subgroup of all elements possessing an *eigenspace* of finite codimension. As with the nonstandard symmetric and alternating groups from earlier, we find normal subgroups of nonstandard special linear groups by generalising finite codimension to small codimension, small being meant in the sense that the codimension is contained in some fixed initial segment I of $\mathfrak{M}$.

Definition 6.2.4. Let I be an arbitrary initial segment of n and let $K' \leq K^{\times}$. We define

$$\mathrm{SL}_n^{(I,K')} = \Big\{ g \in \mathrm{SL}_n : g \text{ has an eigenspace } E_g \\ \text{of codimension } m_g \in I \\ \text{and eigenvalue } \lambda_g \in K' \Big\}.$$

If K' is generated by a single element α we write $\mathrm{SL}_n^{(I,\alpha)}$ for $\mathrm{SL}_n^{(I,\langle\alpha\rangle)}$

Lemma 6.2.5. *Suppose that $I \subseteq_e [0,n]$ is closed under addition and $K' \leq K^{\times}$. Then $\mathrm{SL}_n^{(I,K')} \lhd \mathrm{SL}_n$.*

Proof. Suppose that $g \in \mathrm{SL}_n^{(I,K')}$ and $h \in \mathrm{SL}_n$. Then hE_g is the required eigenspace for hgh^{-1}, so $\mathrm{SL}_n^{(I,K')}$ is closed under conjugation in SL_n.

To see that $\mathrm{SL}_n^{(I,K')}$ is also closed under composition take any $g, h \in \mathrm{SL}_n^{(I,K')}$. Then $\lambda_g\lambda_h$ is an eigenvalue for gh with $E_g \cap E_h$ contained in its eigenspace. But

$$\begin{aligned} \dim(E_g \cap E_h) &= \dim(E_g) + \dim(E_h) - \dim(E_g + E_h) \\ &\geq n - (\mathrm{codim}(E_g) - \mathrm{codim}(E_h)), \end{aligned}$$

so $\mathrm{codim}(E_g \cap E_h) \in I$ as needed.

Since closure under inverses is trivial to show, our proof is complete. □

Definition 6.2.6. Where $g \in \mathrm{SL}_n$ has an eigenvalue λ_g with eigenspace E_g of codimension $m_g \in n/\mathbb{N}$ we call λ_g the *large eigenvalue* and E_g the *large eigenspace* of g.

We leave it to the reader to observe that an element of SL_n can have at most one large eigenvalue.

Now, given distinct initial segments $I, I' \subseteq_e n$ closed under addition it is easy enough to see that we get distinct normal subgroups of $\mathrm{SL}_n^{(I,K')}$. It is similarly easy to see that distinct subgroups of $K^\times$ will likewise give us distinct normal subgroups. But what happens when we factor everything by SZ_n to get to the PSL_n case that we were originally interested in? We adapt the notation from earlier, recalling that $\mathcal{P}$ is defined in Definition 6.2.3 to be the projection map.

Definition 6.2.7. Let $I \subseteq_e n$ be an arbitrary initial segment and let $K' \leq K^\times$. We define

$$\mathrm{PSL}_n^{(I,K')} = \mathcal{P}(\mathrm{SL}_n{}^{(I,K')}).$$

Exactly what happens when we take quotients depends upon the interaction of n with the characteristic of K, as one might imagine. It is possible that every normal subgroup of PSL_n of this form is already of the form $\mathrm{PSL}_n^{(I)}$ for some initial segment $I \subseteq_e n$. More precisely, we have

$$\mathrm{PSL}_n^{(I,K^Z)} = \mathrm{PSL}_n^{(I)} \tag{6.2.2}$$

where $K^Z = \{\alpha \in K^\times \ : \ \alpha I_n \in \mathrm{SZ}_n\}$ and, more generally,

$$\mathrm{PSL}_n^{(I,K')} = \mathrm{PSL}_n^{(I,K'')} \text{ whenever } K'K^Z = K''K^Z. \tag{6.2.3}$$

To see that equation 6.2.2 holds take any $g \in \mathrm{SL}_n^{(I,K')}$ and let $g = \alpha g_1$ where $g_1 \in \mathrm{SL}_n^{(I)}$. We leave it to the reader to write out explicit matrix expressions. But since $\mathcal{P}(\mathrm{SL}_n^{(I)}) = \mathcal{P}(\mathrm{SL}_n^{(I)}\,\mathrm{SZ}_n)$ we get that $\mathcal{P}(g) \in \mathrm{PSL}_n^{(I)}$ and equation 6.2.2 follows.

Now, suppose that $K'K^Z = K''K^Z$ and take any $g \in \mathrm{SL}_n^{(I,K')}$. Let $g = \alpha g_1$, where $g_1 \in \mathrm{SL}_n^{(I)}$, and let $\beta \in K''$ and $\gamma_1, \gamma_2 \in K^Z$ such that $\alpha\gamma_1 = \beta\gamma_2$. Then

$$g = \alpha g_1 = \beta\gamma_2\gamma_1^{-1} = \beta g_2$$

where $g_2 \in \mathrm{SL}_n^{(I)}$. Then $\mathrm{PSL}_n^{(I,K')} \subseteq \mathrm{PSL}_n^{(I,K'')}$ and a similar argument gives the reverse inclusion so that equation 6.2.3 holds also.

Of particular interest is the case where the normal subgroups of PSL_n *are* linearly ordered. Combining this with the assumption that there are no other families of normal subgroups we have the following summary of the situation. The proof is straightforward, and is left to the reader.

Proposition 6.2.8. *The following are equivalent:*

1. *the normal subgroups of* PSL_n *of the form* $\mathrm{PSL}_n^{(I,K')}$ *are linearly ordered;*

2. $K^Z = K^\times$*;*

3. $|K| - 1$ *divides n.*

The question of whether these are the only normal subgroups is partially answered in the next section, though it is opportune to make the following conjecture.

Conjecture 6.2.9. *Let* $n \in \mathfrak{M} \models \mathsf{PA}$ *and let K be an* $\mathfrak{M}$*-finite field. Suppose that H is a proper nontrivial normal subgroup of* PSL_n*. Then there is a subgroup* $K' \leq K^\times$ *and an initial segment I of n such that* $H = \mathrm{PSL}_n^{(I,K')}$.

6.3 Normal subgroup structure of certain nonstandard projective special linear groups

In models elementarily extending the standard model we can show that the list of subgroups of SL_n and PSL_n obtained in the previous section is complete. We proceed to do this in the current section, leaving open the question of how this hypothesis may be weakened.

A good notion for this task is that of a cyclic block in a matrix, taken from the linear algebra literature. Introductory material on cyclic subspaces and vectors may be found in Hoffman and Kunze[19], chapter 7, though our terminology and notation more closely follows that of Lev[28].

For the purposes of this section, n will denote some nonstandard element of M and K will denote some (standard or nonstandard) $\mathfrak{M}$-finite field.

Definition 6.3.1. An $n \times n$ matrix A over a field K is *cyclic* if there is $x \in K^n$ such that the set

$$\{x, Ax, \ldots, A^{n-1}x\}$$

is linearly independent. We call A 1-cyclic if it is similar to a block diagonal matrix

$$\operatorname{diag}(A_1, A_2, \ldots, A_r)$$

where each A_i is cyclic and at most one A_i is 1×1. We say that A is strictly 1-cyclic if none of the A_i are 1×1. The reader should be aware that a cyclic matrix is 1-cyclic and any 2×2 or larger cyclic matrix is strictly 1-cyclic.

The following are well known and again the reader is referred to Hoffman and Kunze for the details.

Lemma 6.3.2. *Let A be an $n \times n$ matrix. Then the following are equivalent:*

1. *A is cyclic;*
2. *the characteristic polynomial of A coincides with the minimum polynomial of A;*
3. *the Jordan normal form of A, taken over the algebraic closure of K, contains one Jordan block per eigenvalue.*

Definition 6.3.3. Let T be a linear transformation on a vector space V of dimension n. For $v \in V$ the *T-cyclic subspace generated by v*, denoted $Z(T;v)$ is the subspace span $\{T^k(v) : k \geq 0\}$. The *T-annihilator* of v is the (unique) monic generator for the ideal of $F[x]$ of polynomials f such that $f(T)[v] = 0$.

Theorem 6.3.4 (Rational decomposition theorem)**.** *Let T be a linear transformation on a vector space V of dimension n. Then there are vectors $v_1, \ldots, v_k \in V$ with respective T-annihilators $p_1, \ldots, p_k$ such that:*

1. *$V = Z(T;v_1) \oplus Z(T;v_2) \oplus \cdots \oplus Z(T;v_k)$;*
2. *for each $1 \leq i \leq k$, p_{i+1} divides p_i.*

Furthermore, k and the p_i are uniquely determined by these requirements.

In matrix terms this theorem states that every invertible matrix A is similar over K to a matrix in block diagonal form

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

where each A_i is cyclic[1]. We restrict our attention to models of PA in which the above theorem holds for groups of nonstandard dimension. Note that if any A_i in the above is 1×1 then similarity coincides with SL_n-conjugacy. By rearranging the blocks we see the following.

[1]Actually, as the reader will observe, it states slightly more than this.

Lemma 6.3.5. *Let $A \in \mathrm{GL}_n(K)$. Then A is similar to a matrix of the form*

$$\begin{pmatrix} \alpha I_{n-l} & \\ & A' \end{pmatrix}$$

where A' is a strictly 1-cyclic $l \times l$ matrix.

Definition 6.3.6. We denote by $\ell(A)$ the greatest $l \in M$ such that A is similar to a matrix of the form above and call this the *cyclic rank* of A.

Most of the work in classifying the normal subgroups of PSL_n is done by the following theorem of Lev. Here, as in Lev's paper[28], $\mathscr{C}^4$ refers to the set of all products of four elements of $\mathscr{C}$.

Theorem 6.3.7 (Lev)**.** *Let $n \geq 2$, $q \geq 4$ and let $\mathscr{C}$ be a 1-cyclic SL_n-conjugacy class of GL_n. Denote*

$$\mathscr{A} = \left\{ A \in \mathrm{GL}_n \smallsetminus Z(\mathrm{GL}_n) \;:\; \det A = (\det \mathscr{C})^4 \right\}.$$

Then $\mathscr{A} \subseteq \mathscr{C}^4$. In particular, if $\mathscr{C} \subseteq \mathrm{SL}_n$ then $\mathrm{SL}_n \smallsetminus \mathrm{SZ}_n \subseteq \mathscr{C}^4$.

What is important here is the *uniformity* with which the elements of SL_n may be expressed as products of conjugates of 1-cyclic elements of SL_n. It is this result which makes 1-cyclic matrices so useful in the current context. The proofs of the above result are long, complicated and spread over multiple papers and this author has not verified all of the claims. The interested reader may try to work out what extension of PA, if any, is needed to prove the above results. We can however check that the statement of the above theorem can be stated in the language of arithmetic, given suitable apparatus for coding and defining the underlying algebraic notions.

We now proceed to show how to determine that there are no normal subgroups other than those of the previous section, given that Lev's result holds within a model $\mathfrak{M}$ of PA

(and the models for which this is the case must include those that elementarily extend the standard model by the discussion of the previous paragraph.)

Lemma 6.3.8. *Suppose that* $|K| \geq 4$ *and* $H \lhd \mathrm{SL}_n$*. Let* $A \in H$ *be a matrix of the form*

$$\begin{pmatrix} \alpha I_{n-l} & \\ & A' \end{pmatrix}$$

where A' *is* 1*-cyclic and* $\alpha \in K^\times$*. Then for every* $B' \in \mathrm{SL}_l$ *we have*

$$\begin{pmatrix} I_{n-l} & \\ & B' \end{pmatrix} \in H.$$

Proof. Let $\mathscr{C}$ denote the conjugacy class of A and take any $B' \in \mathrm{SL}_l$. Let $A_1, A_2, A_3, A_4 \in \mathscr{C}$ such that $B'A_1'A_2'A_3'A_4'$ is nonscalar. Now, by applying Theorem 6.3.7 to the 1-cyclic block A' and using the fact that $\det(B'A_1'A_2'A_3'A_4') = \det(A')^4$, it follows that

$$\begin{pmatrix} \alpha^4 I_{n-l} & \\ & B'A_1'A_2'A_3'A_4' \end{pmatrix} \in \mathscr{C}^4 \subseteq H.$$

But H is a normal subgroup so

$$\begin{aligned} &\begin{pmatrix} \alpha^4 I_{n-l} & \\ & B'A_1'A_2'A_3'A_4' \end{pmatrix} \begin{pmatrix} \alpha I_{n-l} & \\ & A_4' \end{pmatrix}^{-1} \cdots \begin{pmatrix} \alpha I_{n-l} & \\ & A_1' \end{pmatrix}^{-1} \\ &\qquad = \begin{pmatrix} I_{n-l} & \\ & B' \end{pmatrix} \in H \end{aligned}$$

as required. ☐

Lemma 6.3.9. *Suppose that $|K| \geq 4$ and $H \lhd \mathrm{SL}_n$. Let $A \in H$ be a matrix of the form*

$$\begin{pmatrix} \alpha I_{n-l} & \\ & A' \end{pmatrix}$$

where A' is strictly 1-cyclic and $\alpha \in K^\times$. Then the following holds.

1. *If $l > n/\mathbb{N}$ then $H = \mathrm{SL}_n$.*

2. *If $l \in n/\mathbb{N}$ then for all $m \in l \cdot \mathbb{N}$ and all $B' \in \mathrm{SL}_m$ we have that*

$$\begin{pmatrix} I_{n-m} & \\ & B' \end{pmatrix} \in H.$$

Proof. We prove only (1) since (2) follows by similar reasoning. We begin by showing that if $l \geq \lfloor \frac{n}{2} \rfloor$ then $H = \mathrm{SL}_n$. Let $k = n - l$ and observe that by the previous lemma we may assume that $\alpha = 1$. Furthermore, by the same lemma, a rearrangement of rows and columns and the fact that $k \leq l$, it follows that for any $B' \in \mathrm{SL}_k$

$$\begin{pmatrix} B' & & \\ & I_{l-k} & \\ & & I_k \end{pmatrix} = \begin{pmatrix} B' & \\ & I_l \end{pmatrix} \in H.$$

Let B' be cyclic and observe that

$$\begin{pmatrix} B' & \\ & A' \end{pmatrix} = \begin{pmatrix} B' & \\ & I_l \end{pmatrix} \begin{pmatrix} I_k & \\ & A' \end{pmatrix} \in H$$

and that this matrix is 1-cyclic, and strictly 1-cyclic except in case $k = 1$. In this case we are done by the normality of H and Theorem 6.3.7.

Suppose now that $l < \lfloor \frac{n}{2} \rfloor$ but $l > n/\mathbb{N}$. Then

$$\begin{pmatrix} I_{n-2l} & & \\ & A' & \\ & & I_l \end{pmatrix} \begin{pmatrix} I_{n-l} & \\ & A' \end{pmatrix} = \begin{pmatrix} I_{n-2l} & & \\ & A' & \\ & & A' \end{pmatrix} \in H$$

and

$$\begin{pmatrix} A' & \\ & A' \end{pmatrix}$$

is a strictly 1-cyclic matrix since A' is. Repeating this process until $l > \left\lfloor \frac{n}{2} \right\rfloor$ we finish in the case of the previous argument and hence we are done here also. $\square$

Lemma 6.3.10. *Suppose that $|K| \geq 4$. Let $H \lhd \mathrm{SL}_n$ contain a matrix A of the form*

$$\begin{pmatrix} \alpha I_{n-l} & \\ & A' \end{pmatrix}$$

where A' is strictly 1-cyclic and $l \in n/\mathbb{N}$. Let $K' = \langle \alpha \rangle$ and $I = l \cdot \mathbb{N}$. Then $\mathrm{SL}_n^{(I,K')} \leq H$.

Proof. It follows from the second case of the previous lemma and the normality of H that $\mathrm{SL}_n^{(I,1)} \leq H$. Now take any $0 \neq r \in \mathbb{Z}$ and any matrix $B \in \mathrm{SL}_n$ of the form

$$\begin{pmatrix} \alpha^r I_{n-k} & \\ & B' \end{pmatrix}.$$

In a similar fashion to the proof of the previous lemma we may replace A by some product of conjugates of A so as to assume without loss that $k \leq l$. Then BA^{-r} is of the form

$$\begin{pmatrix} I_{n-l} & \\ & \begin{pmatrix} I_{l-k} & \\ & B' \end{pmatrix} A'^{-r} \end{pmatrix}$$

so is contained in H. Thus $B = (BA^{-r})A^r \in H$ and the proof is complete. $\square$

The theorem is then straightforward from the above lemmas and the fact that any matrix that is not strictly 1-cyclic is SL_n-conjugate to a matrix of the form $\mathrm{diag}(\alpha I_{n-l}, A')$ where A' is strictly 1-cyclic. We state the theorem only for the case where $\mathfrak{M}$ elementarily extends the standard model, leaving the more general case for the interested reader to investigate.

Theorem 6.3.11. *Let $\mathfrak{M}$ be a nonstandard model of* PA *with domain M, elementarily extending the standard model. Let $n \in M$ be nonstandard and let K be an $\mathfrak{M}$-finite field with $|K| \geq 4$. Suppose that $H \lhd \mathrm{PSL}_n(K)$ is a proper nontrivial normal subgroup. Then there is $I \subseteq_e n$ closed under addition and $K' \leq K^\times$ such that $H = \mathrm{PSL}_n^{(I,K')}(K)$.*

Proof. Let $H \lhd \mathrm{PSL}_n(K)$ and lift this to a normal subgroup $H\,\mathrm{SZ}_n$ of SL_n. By lemma 6.3.10 we have

$$H\,\mathrm{SZ}_n = \bigcup_{A \in H\,\mathrm{SZ}_n} \mathrm{SL}_n^{(\ell(A)\cdot\mathbb{N}, \alpha_A)}$$

and it suffices now to observe that

$$\bigcup_{A \in H\,\mathrm{SZ}_n} \mathrm{SL}_n^{(\ell(A)\cdot\mathbb{N}, \alpha_A)} = \mathrm{SL}_n^{(I,K')}$$

where $I = \bigcup_{A \in H\,\mathrm{SZ}_n} \ell(A) \cdot \mathbb{N}$ and $K' = \langle \alpha_A \,:\, A \in H\,\mathrm{SZ}_n \rangle$ where $\ell(A)$ is the cyclic rank of A from Definition 6.3.6. $\square$

6.4 Nonstandard projective symplectic groups

Less is known about the cases of the nonstandard analogues of other classical groups. Here we give some comments on the symplectic case.

We begin with a brief discussion of the countably infinite case covered by Gardener in his paper on infinite dimensional classical groups[15]. Gardener's main result is that for a symplectic space V of countably infinite dimension, $\mathrm{Aut}(V)/\mathcal{E}$ is simple, where $\mathcal{E}$ is the subgroup of $\mathrm{Aut}(V)$ consisting of those elements that possess an eigenspace of finite codimension. Gardener uses this to deduce the composition series for $\mathrm{Aut}(V)$, namely

$$1 \lhd \mathcal{F} \lhd \mathcal{E} \lhd \mathrm{Aut}(V)$$

where $\mathcal{F}$ is the subgroup of $\mathrm{Aut}(V)$ of elements possessing a fixed space of finite codimension and $\mathcal{E}$ is the subgroup of elements possessing an eigenspace of finite codimension.

That there are no more normal subgroups (like those in the previous sections where there were a nontrivial restricted choice of eigenvalues) and why this does not depend upon the field as the special linear case does, follows from the following result which is standard in the area of symplectic groups.

Lemma 6.4.1. *Let $g \in \mathrm{Sp}_{2n}(K)$ and suppose that $\lambda \in K$ is an eigenvalue for g of multiplicity m. Then λ^{-1} is also an eigenvalue for g of multiplicity m.*

The composition series result follows from this since if λ is the eigenvalue for g with an eigenspace of finite codimension, then λ^{-1} is an eigenvalue for g with an eigenspace of finite codimension. It then follows that these eigenspaces have nontrivial intersection whence $\lambda v = \lambda^{-1} v$ for some nonzero vector v so that $\lambda = \lambda^{-1}$ and hence $\lambda = 1$ or else $\lambda = -1$.

A similar argument runs through in the nonstandard case, at which point factoring through by the centre of the symplectic group yields the result that the normal subgroups of $\mathrm{PSp}_{2n}(K)$ corresponding to subgroups of $\mathrm{Sp}_{2n}(K)$ of the form $\mathrm{Sp}_{2n}(K)^{(I,K')}$ are necessarily linearly ordered. We thus have the following conjecture (recalling that the center of $\mathrm{Sp}_{2n}(K)$ is $\{1, -1\}$.)

Conjecture 6.4.2. *Let $\mathfrak{M}$ be a nonstandard model of* PA *and let n be a nonstandard element of $\mathfrak{M}$. If $H \lhd \mathrm{PSp}_{2n}(K)$ then there is $I \subseteq_e \mathfrak{M}$ closed under addition such that*

$$H = \mathcal{P}(\mathrm{Sp}_{2n}(K)^{(I,-1)}).$$

The result in the earlier material allowing for us to answer the question in the case of the nonstandard special linear groups in models elementarily extending the standard model was a uniformity result stating effectively that any 'smaller' element of SL_n could be constructed as a product of four conjugates of any given 'larger' element, where 'larger' and 'smaller' were defined formally in terms of the cyclic rank, which is the dimension of the nontrivial part of the expression of the matrix of the group element in block diagonal form (by nontrivial here we mean the part of the block diagonal matrix that differs from a scalar matrix.) Essentially the result, due to Lev for special linear groups, took the form: for all groups G of a family $\mathcal{G}$ tthere is $k \in \mathbb{N}$ and a partial ordering $\preceq$ of each $G \in \mathcal{G}$ such that whenever g and h are elements of G with $g \preceq h$ there is a sequence of up to k conjugates of h whose product is g. Unfortunately there is little in the literature that helps towards the production of such a result for symplectic groups, and it is unlikely that the normal subgroup structure conjecture above can be proved without one.

Chapter 7

Nonstandard measure of nonstandard finite groups

7.1 Measure of S_n and its bounded subgroups

One may verify, as we shall do shortly, by internalising the argument for finite symmetric groups, that the order of S_n in the sense of $\mathfrak{M}$ is $n!$. The order of the bounded subgroups of S_n, in the sense of the measures defined in chapter 3 is the subject of the subsequent proposition. It should be clear to the reader how to adapt these arguments to the alternating groups as well as a few others.

Recall that, as defined in chapter 4, S_n contains precisely the minimal codes for sequences of length n whose elements are pairwise distinct and less than n. Thus we need to count these (minimal codes for) sequences.

Let $\mathsf{MinCode}(x)$ have the intended meaning 'x is a minimal code for the sequence it codes' and let $\mathsf{Distinct}(x)$ have the intended meaning 'the elements coded by x are pairwise distinct.' Define

$$\mathsf{Perm}(x,n) = |x| = n \ \wedge\ \forall y{<}|x|\,([x]_y < n)\ \wedge\ \mathsf{Distinct(x)}\ \wedge\ \mathsf{MinCode(x)}.$$

This can be taken to be the formula referred to in chapter 4: it is true precisely if x is a minimal code for a permutation on the first n elements of $\mathfrak{M}$. The reader should identify $x \in \mathrm{S}_n$ with the more formal $\mathfrak{M} \models \mathsf{Perm}(x,n)$, but the former notation makes things a little clearer.

Proposition 7.1.1. *Let $n \in M$. Then $|\mathrm{S}_n| = n!$.*

Proof. Let $\phi(x)$ be the formula $|\mathrm{S}_x| = x!$ and observe that $\mathfrak{M} \models \phi(0)\ \wedge\ \phi(1)$ trivially. Suppose that $\mathfrak{M} \models \phi(k)$ for some $k \in M$ and take any $x \in M$ such that $\mathfrak{M} \models \mathsf{Perm}(x,k+1)$. Then there is precisely one x' such that $\mathfrak{M} \models \mathsf{Perm}(x',k)$ and (thought of as a sequence) x' is obtained from x by deleting the single instance of k. Let $\psi(x)$ be the Skolem function yielding x' given x as its sole parameter. That is,

$$\begin{aligned}\psi(x) = x' \Leftrightarrow_{\mathrm{def}} \forall y{<}\left|x'\right| (&\\ &\big((\exists z{\leq}y\ [x]_z = k) \wedge \big([x]_{y+1} = [x']_y\big)\big) \vee \\ &\big((\forall z{\leq}y\ [x]_z \neq k) \wedge \big([x]_y = [x']_y\big)\big)\big).\end{aligned}$$

Then there are $k+1$ elements $z \in M$ with $\mathfrak{M} \models \mathsf{Perm}(z,k+1)\ \wedge\ x' = \psi(z)$, one for each of the places where we can insert k. But then since ψ is a function it follows that $|\{x \in M\ :\ \mathfrak{M} \models \mathsf{Perm}(x,k+1)\}| = |\mathrm{S}_{k+1}| = (k+1)\,|\mathrm{S}_k|$ and the result follows by induction. □

Proposition 7.1.2. *Let $n \in \mathfrak{M}$ be nonstandard and let $I \subseteq_e n$ be closed under addition. Then*

$$\mu(\mathrm{S}_n^{[I]}) = \nu(\mathrm{S}_n^{[I]}) = \sup\left\{ \left\lfloor \frac{n!}{(n-k)!} \right\rfloor \;:\; k \in I \right\}$$
$$= \inf\left\{ \left\lfloor \frac{n!}{(n-k)!} \right\rfloor \;:\; k < n, k \notin I \right\}$$

Proof. By a standard inclusion-exclusion argument we know that the number of r-permutations having no fixed points is

$$r! \sum_{k=0}^{r} \frac{(-1)^k}{k!}.$$

It is evident from this that the number of r-permutations with support of cardinality r is

$$\binom{n}{r} r! \sum_{k=0}^{r} \frac{(-1)^k}{k!}$$
$$= \frac{n!}{(n-r)!} \sum_{k=0}^{r} \frac{(-1)^k}{k!}.$$

These results may be readily internalised in PA by the use of suitable inductive arguments, but we do not do so here. Thus there are

$$\sum_{j=0}^{r} \frac{n!}{(n-j)!} \sum_{k=0}^{j} \frac{(-1)^k}{k!}$$

n-permutations having support of cardinality at most r.

Observe that for $j \geq 2$ we have

$$\frac{1}{3} \leq \sum_{k=0}^{j} \frac{(-1)^k}{k!} \leq 1 \tag{7.1.1}$$

and by explicit calculation we have

$$\sum_{k=0}^{0} \frac{(-1)^k}{k!} = 1 \text{ and } \sum_{k=0}^{1} \frac{(-1)^k}{k!} = 0$$

. so we see that, for $r \geq 2$,

$$\begin{aligned}
&\sum_{j=0}^{r} \left\lfloor \frac{n!}{3(n-j)!} \right\rfloor \\
&\leq \sum_{j=0}^{r} \left(\frac{n!}{(n-j)!} \sum_{k=0}^{j} \frac{(-1)^k}{k!} \right) \\
&\leq \sum_{j=0}^{r} \frac{n!}{(n-j)!}.
\end{aligned}$$

The reader should be aware that the cases $r = 0$ and $r = 1$ may be handled by explicit calculation.

Now, take any nonzero $k \in I$. Then $2k \in I$ by closure under addition. Thus there are at least

$$\sum_{j=0}^{2k} \frac{n!}{3(n-j)!}$$

distinct permutations in $\mathrm{S}_n^{[I]}$.

We then have

$$\begin{aligned}
\sum_{j=0}^{2k} \frac{n!}{3(n-j)!} &\geq 3 \sum_{j=0}^{k} \frac{n!}{3(n-j)!} \\
&\geq \frac{n!}{(n-k)!}
\end{aligned}$$

so that there are at least

$$\frac{n!}{(n-k)!}$$

distinct permutations in $\mathrm{S}_n^{[I]}$ (the last inequality being produced by taking a single term of the sum.) Hence

$$\mu(\mathrm{S}_n^{[I]}) \supseteq \sup \left\{ \frac{n!}{(n-k)!} \ : \ k \in I \right\}.$$

We now deal with the outer measure. Take any $k \in n \setminus I$. Then $\left\lfloor \frac{k}{2} \right\rfloor \in n \setminus I$. Observe that

$$\nu(\mathrm{S}_n^{[I]}) \subseteq \sum_{j=0}^{\lfloor k/2 \rfloor} \frac{n!}{(n-j)!}$$

since permutations in $\mathrm{S}_n^{[I]}$ have support of cardinality less than $\lfloor \frac{k}{2} \rfloor$. Also

$$\sum_{j=0}^{\lfloor k/2 \rfloor} \frac{n!}{(n-j)!} \leq \left\lfloor \frac{k}{2} \right\rfloor \frac{n!}{(n-\lfloor \frac{k}{2} \rfloor)!}$$
$$\leq \frac{n!}{(n-k)!}$$

since

$$\left\lfloor \frac{k}{2} \right\rfloor \leq \frac{(n-\lfloor \frac{k}{2} \rfloor)!}{(n-k)!}.$$

Thus we have

$$\nu(\mathrm{S}_n^{[I]}) \subseteq \inf\left\{ \frac{n!}{(n-k)!} \ : \ k \in n \smallsetminus I \right\}$$

which completes the proof since straightforward overspill and underspill arguments show that

$$\sup\left\{ \frac{n!}{(n-k)!} \ : \ k \in I \right\} = \inf\left\{ \frac{n!}{(n-k)!} \ : \ k \in n \smallsetminus I \right\}.$$

□

7.2 Internal indices: Measure of transversals of external subgroups

When a subgroup of S_n, say, is not $\mathfrak{M}$-finite, it is not possible to define the index in the usual way, since this will not evaluate to an element of M. Thus we look to using our measure construction from chapter 3 for this purpose. The definition we take here is essentially that of chapter 3 with the exception that we are restricting the $\mathfrak{M}$-finite subsets to those contained by a transversal and the $\mathfrak{M}$-finite supersets to those containing a transversal.

Definition 7.2.1. We define the *measures of the index* of a subgroup H within a group G by the following.

$$\mu(G:H) = \sup\left\{|x| \;:\; x \in M, [\![x]\!] \subseteq G, \forall g,h{\in}[\![x]\!]\left(g^{-1}h \in H \to g = h\right)\right\}$$
$$\nu(G:H) = \inf\left\{|x| \;:\; x \in M, [\![x]\!] \subseteq G, \forall g{\in}G\,\exists h{\in}[\![x]\!]\left(g^{-1}h \in H\right)\right\}$$

We leave open the question as to whether this is equivalent to measuring a transversal. If there exists a measurable transversal then it follows easily that these two measures above agree, though the more general questions such as whether measurability in terms of the above definition implies the existence of a measurable transversal are left open. The reader should also see that these measures of index agree with the expected values on any $\mathfrak{M}$-finite quotient. For example $\mu(\mathrm{S}_n : \mathrm{A}_n) = \nu(\mathrm{S}_n : \mathrm{A}_n) = 2$. (Note that an $\mathfrak{M}$-finite quotient always gives rise to an $\mathfrak{M}$-finite and hence measurable transversal.)

In what follows we often need to talk about an $\mathfrak{M}$-finite set contained within a possibly external transversal and an $\mathfrak{M}$-finite superset of such a transversal. We define two pieces of terminology for these.

Definition 7.2.2. We call a subset of a transversal a *subtransversal* and a supserset of a transversal a *supertransversal*, prepending the adjective $\mathfrak{M}$-finite as applicable.

This allows us to measure the indices of one bounded subgroup within another. Again, the arguments applied here to the symmetric group case are easily adapted to the alternating group case.

Proposition 7.2.3. *The index of any bounded subgroup of* S_n *within* S_n *is measurable and*

$$\mu(\mathrm{S}_n : \mathrm{S}_n^{[I]}) = \nu(\mathrm{S}_n : \mathrm{S}_n^{[I]}) = \sup\{(n-k)! \,:\, k < n, k \notin I\}$$
$$= \inf\{(n-k)! \,:\, k \in I\}$$

for any initial segment I *of* $\mathfrak{M}$ *closed under addition.*

Proof. Standard underspill and overspill arguments show that

$$\sup\{(n-k)! \,:\, k \in {}_{<}n \smallsetminus I\} = \inf\{(n-k)! \,:\, k \in I\}$$

so we need only check that

$$\mu(\mathrm{S}_n : \mathrm{S}_n^{[I]}) \supseteq \sup\{(n-k)! \,:\, k \in {}_{<}n \smallsetminus I\}$$

and

$$\nu(\mathrm{S}_n : \mathrm{S}_n^{[I]}) \subseteq \sup\{(n-k)! \,:\, k \in {}_{<}n \smallsetminus I\}.$$

We begin with μ. Take any $k \in {}_{<}n \smallsetminus I$. We need to show that there exists an $\mathfrak{M}$-finite subtransversal of $\mathrm{S}_n^{[I]}$ in S_n of cardinality at least $(n-k)!$. To do this we begin by showing that any $\mathfrak{M}$-finite subset, say X, of cardinality less than $(n-k)!$ can be extended with an element x such that for all $y \in X$ we have $\deg(x^{-1}y) \geq \left\lfloor \frac{k}{2} \right\rfloor$: this is then used inductively to obtain the required $\mathfrak{M}$-finite subtransversal.

Let $\phi(l,k,n)$ be the following formula capturing the required meaning from the above paragraph:

$$l < (n-k)! \to \forall c\,((|c| \leq l \,\wedge\, [\![c]\!] \subseteq \mathrm{S}_n)$$
$$\exists x \in \mathrm{S}_n\, \forall y \in [\![c]\!] \left(\deg(x^{-1}y) \geq \left\lfloor \tfrac{k}{2} \right\rfloor\right)).$$

We wish to prove that $(\mathfrak{M},k,n) \models \forall l\, \phi(l,k,n)$. To do this fix n and k and suppose that $l < (n-k)!$. Take any $\mathfrak{M}$-finite subset $[\![c]\!]$ of S_n of cardinality l and observe that for $g \in [\![c]\!]$ there are at most

$$\sum_{j=0}^{\lfloor k/2 \rfloor} \frac{n!}{(n-j)!} < \frac{n!}{(n-k)!}$$

elements h of S_n with $\deg(h^{-1}g) \leq \left\lfloor \frac{k}{2} \right\rfloor$, this being an application of the right hand inequality of 7.1.1. Thus there are at most

$$|c|\, \frac{n!}{(n-k)!}$$

elements h of S_n with $\deg(h^{-1}g) \leq \left\lfloor \frac{k}{2} \right\rfloor$ for some $g \in [\![c]\!]$. Since $[\![c]\!] \leq l < (n-k)!$ it follows that

$$|c|\, \frac{n!}{(n-k)!} < n!$$

so these elements do not exhaust S_n. Thus there is $x \in \mathrm{S}_n$ satisfying the existential subformula of ϕ.

Now, starting with the empty set and applying ϕ at the inductive step we see that we can construct a subtransversal of $\mathrm{S}_n^{[I]}$ in S_n of cardinality $(n-k)!$ as is required to prove the containment involving μ.

For the ν containment take any $k \in I$. We need to show that that there is no transversal containing an $\mathfrak{M}$-finite subset of cardinality $(n-k)!$. (This being sufficient to show that $\nu(\mathrm{S}_n : \mathrm{S}_n^{[I]}) < (n-k)!$.) We can achieve this by showing that given any subset X of S_n of cardinality $(n-k)!$ there are x and y in X such that $\deg(x^{-1}y) \in I$. Let $l \notin I$ (for otherwise we are done). Then, by the left hand inequality of 7.1.1, for each $x \in X$ there are at least

$$\sum_{r=0}^{\left\lfloor \frac{l-1}{2} \right\rfloor} \frac{n!}{3(n-r)!}$$

elements y with $\deg(x^{-1}y) < l/2$. (Intuitively these are the elements closest to x with respect to a permutation metric.) Then there are at least

$$\begin{aligned}
&(n-k)!\sum_{r=0}^{\lfloor\frac{l-1}{2}\rfloor}\frac{n!}{3(n-r)!}\\
\geq&(n-k)!\left(\frac{n!}{3(n-k)!}+\frac{n!}{3(n-k-1)!}+\frac{n!}{3(n-k-2)!}\right)\\
>&(n-k)!\frac{n!}{(n-k)!}=n!
\end{aligned}$$

elements in X, using the observation that we necessarily have $l > 2(k+2)$ in this case since $k \in I < l$. This is the required contradiction and we conclude that no transversal of $\mathrm{S}_n^{[I]}$ within S_n can contain a subset of cardinality $(n-k)!$, which suffices to prove that

$$\nu(\mathrm{S}_n : \mathrm{S}_n^{[I]}) \subseteq \sup\{(n-k)! \ : \ k \in {}_{<}n \smallsetminus I\}$$

and this completes the proof. $\square$

We conclude this section with a conjecture of a generalisation of the above result and a more general question.

Conjecture 7.2.4. *The index of any bounded subgroup of* S_n *within another such bounded subgroup is measurable and*

$$\begin{aligned}
\mu(\mathrm{S}_n^{[J]} : \mathrm{S}_n^{[I]}) = \nu(\mathrm{S}_n^{[J]} : \mathrm{S}_n^{[I]}) &= \sup\{(l-k)! \ : \ k < l, l \in J, k \notin I\}\\
&= \inf\{(n-k)! \ : \ k \in I, l > J, l < n\}
\end{aligned}$$

for any initial segments $I \subseteq J$ *of* $\mathfrak{M}$ *closed under addition.*

Question 7.2.5. Let $H < K \leq \mathrm{S}_n$ be measurable. Then is it the case that

$$\begin{aligned}
\mu(K:H) = \nu(K:H) &= \sup\{\lfloor x/y\rfloor : \ |H| < y < x < |K|\}\\
&= \inf\{\lceil x/y\rceil : y < |H| \text{ and } |K| < x\}\,?
\end{aligned}$$

7.3 Non measurable groups

The purpose of this section is a brief discussion of non-measurable groups. We begin with a construction of a non-measurable group. This construction has much in common with those of chapter 3.

Proposition 7.3.1. *Let $n \in M$ be nonstandard. There exists a non-measurable subgroup G of S_n.*

Proof. Let $I \subseteq_e n$ be closed under addition. Let $(d_n)_{n\in\mathbb{N}}$ enumerate the codes of $\mathfrak{M}$-finite sets with $\mathbb{N} < |d_n| < I$. Let $G_0 = 1$, the group consisting of just the identity. We proceed to inductively construct a group generated by 2-cycles with disjoint support. At stage n we assume that G_n has been constructed. Now, since $|d_n| < I$, there is a 2-cycle $(k, k+1)$ in S_n with $k+1 \in I$ (which is assured by the previous stages of the induction) and $k > x$ for any x in the support of G_n. Let $G_{n+1} = \langle G_n, \{(k, k+1)\}\rangle$ and let $G = \bigcup_{n\in\mathbb{N}} G_n$. Then $\mu(G) = \mathbb{N}$ since every $\mathfrak{M}$-finite subset of G is necessarily finite by the above construction and $\nu(G) > I$, again by the construction above. □

Clearly the group constructed above has non-measurable support. This immediately opens the following question, which time did not permit the author to address.

Question 7.3.2. Is there a subgroup G of S_n, for some nonstandard n, such that G is non-measurable but $\mathrm{supp}(G)$ is measurable? Is it further possible to ensure that G has $\mathfrak{M}$-finite support, for example Ω_n?

The author's suspicions are that the answer to both questions is yes, though there is no evidence of substance in either direction. A related question concerns transversals.

Question 7.3.3. Let H be a measurable subgroup of a measurable group G. Then is it necessary that any transversal of H in G be measurable?

Here this author suspects that there exist non-measurable transversals, especially in the case where the index of H in G is nonstandard, for one has much choice of coset representatives and there should be sufficient for a diagonalisation construction to go through.

Chapter 8

Conclusion

The purpose of this thesis has been to look at the intuition of size given by initial segments of a model of Peano arithmetic both in its own right and in the context of nonstandard finite groups and at the normal subgroup structure of some nonstandard groups, giving some illustration of what happens when the notions of measure and finite groups are brought together in nonstandard models. As such the work has mainly been of an exploratory nature and there are many questions and loose ends remaining. Many approaches to problems raised in this thesis drew blanks and did not result in material worthy of inclusion. Thus this thesis is mainly a compendium of what did work and ideas of where future progress may be made. Now we bring proceedings to a close.

8.1 A brief recapitulation of the thesis

Having dealt with preliminary material in chapters 1 and 2, we began our investigation in earnest with a look at initial segments as a measure of size in their own right, considering measures of non-$\mathfrak{M}$-finite sets via initial segments constructed as limits of sizes of included or including $\mathfrak{M}$-finite sets.

The constructions are looked upon as analogues to the Lebesgue construction on the real line and a few similarities emerged. For bounded sets these measure constructions appear to work reasonably well, the only known counterexamples where the two measures μ and ν disagree being the pathological diagonalisation constructions of section 3.2. The non-Lebesgue measurable sets require the axiom of choice to construct and it appears likely that something similar can be said of the non-$\mathfrak{M}$-measurable sets constructed, though such questions go beyond the scope of this thesis: choice is near universally assumed in model theory and questions relating to its absence would require certain foundational investigations before analysis of the implications for the measure construction could be attempted.

For unbounded sets, in particular the cofinal ω-sequences, we looked briefly at the idea of temporarily end-extending the model to allow bounds to be adjoined whilst the set in question is gauged by sets coded in the end-extension. This allowed for the outer measure construction to be refined so as to agree with the inner measure in some cases, especially extensional sets and those close to being extensional. Questions of how extensional sets and unbounded extended-measurable sets are related, whilst interesting, appeared to be a difficult proposition, especially in light of the nature of open questions regarding whether a model contains extensional or elementary extensional ω-sequences at all.

The measures chapter concluded with a brief look at a possible future direction for the measures work, namely measuring and approximating non-$\mathfrak{M}$-finite functions on bounded subsets of models of arithmetic, this being related to the measures discussed in an analogous fashion to the relation between Lebesgue measure and integration (hence the proposed integral notation.)

We then began our investigations into nonstandard finite groups by looking at nonstandard permutations on $\mathfrak{M}$-finite sets, explaining what they are and proceeding to give an interesting result that a necessary and sufficient condition for $\mathbb{N}$ to be strong in $\mathfrak{M}$ is that $g \in \mathrm{S}_n^{\mathrm{Sym}_n}$ only if τ_g is coded in the sense of the standard system of $\mathfrak{M}$. This set the scene for our results concerning the normal subgroup structure of nonstandard symmetric and alternating groups. We gave a classification of the normal subgroup structure of all nonstandard alternating and symmetric groups, echoing the results of Baer[4], Schreier and Ulam[36] from the early 20th century regarding infinite symmetric groups. Our proof had a number of similarities to the proof of the result in the infinite case, again relying on the fact that every permutation is a product of two involutions.

We concluded the chapter on the normal subgroup structure of nonstandard symmetric and alternating groups with a look at a result due to Macpherson and Neumann regarding supplements of normal subgroups of symmetric groups, with a view to producing similar results in the nonstandard case. Time did not permit a proper investigation beyond examining the possibility of a reasonably direct adaptation of the result in question. This direct adaptation proved, as one would suspect, not to be possible.

The normal subgroup structure results inspired a search for similar results concerning normal subgroups of other nonstandard groups. Our findings were presented in chapter 6. It was quickly evident that the normal subgroup structure of the nonstandard analogues of other finite simple groups would be important and necessary to look at first. The cyclic groups of nonstandard prime order presented little difficulty, so the work then proceeded to look at the nonstandard nonabelian finite simple groups. A cursory look at the theory of Lie groups as expounded by Carter[10] gave little to aid a uniform approach to this problem (as is possible to the simplicity of finite simple groups of Lie type.) Thus we began with an attempt to classify the normal subgroup structure of special linear groups. As such this was a difficult task and, whilst much effort was expended in the search for concrete results, little of substance could be found, though we did manage to apply a result of Lev in the case that model $\mathfrak{M}$ elementary extends the standard model $\mathbb{N}$. The generalisation of this result to arbitrary models and the similar result for symplectic groups were left as conjectures.

We concluded the material of the thesis with a brief look at an application of the measures of chapter 3 to properties of groups such as measuring indices of the normal subgroups of S_n.

8.2 Future Research

With regards to the measures work, there is much that can be done in this new area. The analogy to Lebesgue integration is one suggestion and there are a plethora of foundational questions that may be asked such as the following.

Question 8.2.1. Is AC necessary to construct a non-$\mathfrak{M}$-measurable subset of M?

This question is motivated by the well known situation regarding Lebesgue measure on the real line. On another front, one is obliged to ask the following.

Question 8.2.2. What properties are necessary for a subset X of M to be $\mathfrak{M}$-measurable? What properties form a sufficient condition for $\mathfrak{M}$-measurability?

There are also analogous questions for the notation of measure utilising end extensions.

The work on nonstandard symmetric and alternating groups, and also the preliminary work on nonstandard permuations, presents what could be considered the best starting point for future research. There is much in the literature regarding infinite permutation groups and the natural next step is to investigate the possibility of analogous results in the nonstandard case, as we have attempted with the results classifying supplements of bounded subgroups of infinite symmetric groups. This area will also give motivation to questions regarding the differences between finite, nonstandard and infinite combinatorics.

Nonstandard cyclic groups present objects worthy of investigation, in particular one would like to elucidate the structure of nonstandard cyclic groups whose order is a nonstandard power of a standard prime.

In addition, the question of how the external normal subgroup structure of an arbitrary (composite) $\mathfrak{M}$-finite group is related to that of the $\mathfrak{M}$-finite composition factors is possibly worthy of further investigation.

However, the difficulties experienced with the nonstandard classical groups can be taken to indicate that the possible gains in terms of achievable results may well be outweighed by the difficulty of obtaining and proving them. Whilst the normal subgroup structure of nonstandard classical groups presented an interesting question, it appears that there is little to justify further effort in this area: the results that were obtained in the course of this thesis relied on results concerning finite groups and added little to what is already known. Of course the questions themselves are left open and indeed a different slant upon the research may provide inroads, but this author can see little at present to merit a sustained assault upon these problems.

On a final, related note, there is the question regarding how these groups are related to the pseudofinite groups studied by Felgner[13] and Wilson[41], which we phrase as follows.

Question 8.2.3. Given a pseudofinite group G, does there exist a model $\mathfrak{M} \models \mathsf{PA}$ and an $\mathfrak{M}$-finite group G' such that $G \cong G'$?

As such, this question is intended to help place the current work with respect to the current related literature, as needs to be done with regards the notion of nonstandard finite groups considered as infinite groups in their own right.

Bibliography

[1] J. D. Allsup. Initial segments of nonstandard models of peano arithmetic. *M.Phil(Qual) thesis*, September 2001.

[2] J. D. Allsup and R. W. Kaye. Normal subgroups of nonstandard symmetric and alternating groups. To appear.

[3] J. Ax. The elementary theory of finite fields. *Ann. of Math. (2)*, 88:239–271, 1968.

[4] R. Baer. Die Kompositionsreihe der Gruppe alle eieindeutigen Abbildungen einer unendlichen Menge auf sich. *Studia Math*, 5:15–17, 1935.

[5] J. K. Barwise and J. S. Schlipf. An introduction to recursively saturated and resplendant models. *J. Symbolic Logic*, 41(2):531–536, 1976.

[6] M. Bhattacharjee, H. D. Macpherson, R. G. Möller, and P. M. Neumann. *Notes on Infinite Permutation Groups*, volume 12 of *Texts and Readings in Mathematics*. Hindustan Book Agency, New Delhi, 1997. Also available as Lecture Notes in Math., Vol. 1698., Springer-Verlag, 1998.

[7] S. Bigelow. Supplements of bounded permutation groups. *J. Symbolic Logic*, 63(1):89–102, 1998.

[8] R. Bird. *Introduction to Functional Programming using Haskell, second edition*. Prentice Hall Series in Computer Science. Prentice Hall, 1998.

[9] P. J. Cameron. *Oligomorphic permutation groups*, volume 152 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1990.

[10] R. W. Carter. *Simple groups of Lie type*. John Wiley & Sons, London-New York-Sydney, 1972. Pure and Applied Mathematics, Vol. 28.

[11] C. C. Chang and H. J. Keisler. *Model Theory*, volume 73 of *Studies in logic and the foundations of mathematics*. North Holland, Amsterdam, 1973.

[12] K. Ciesielski. *Set Theory for the Working Mathematician*. Number 39 in London Mathematical Society Student Texts. Cambridge University Press, 1997.

[13] U. Felgner. Pseudo-endliche gruppen. *Proceedings of the 8th Easter Conference on Model Theory*, pages 82–96, 1990.

[14] D. H. Fremlin. *Measure Theory*. Torres Fremlin, 25 Ireton Road, Colchester, CO3 3AT, England, 2000.

[15] T. S. Gardener. Infinite dimensional classical groups. *J. London Math. Soc*, 51(2):219–229, 1995.

[16] A. J. Hahn and O. T. O'Meara. *The classical groups and K-theory*, volume 291 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1989. With a foreword by J. Dieudonné.

[17] W. Hodges. *Building Models by Games*. Number 2 in London Mathematical Society Student Texts. Cambridge University Press, 1985.

[18] W. Hodges. *Model theory*. Cambridge University Press, 1993.

[19] K. Hoffman and R. Kunze. *Linear Algebra*. Prentice-Hall, Englewood Cliffs, N. J., 1961.

[20] A. E. Hurd and P. A. Loeb. *An Introduction to Nonstandard Real Analysis*. Academic Press, New York, 1985.

[21] R. W. Kaye. *Models of Peano Arithmetic*, volume 15 of *Oxford Logic Guides*. Oxford University Press, 1989.

[22] R. W. Kaye. A galois correspondende for countable recursively saturated models of peano arithmetic. In R. W. Kaye and H. D. Macpherson, editors, *Automorphisms of First-Order Structures*. Oxford University Press, 1994.

[23] R. W. Kaye, R. Kossak, and H. Kotlarski. Automorphisms of recursively saturated models of arithmetic. *Ann. Pure Appl. Logic*, 55(1):67–99, 1991.

[24] R. W. Kaye and H. D. Macpherson, editors. *Automorphisms of First Order Structures*. Oxford University Press, 1994.

[25] L. A. S. Kirby and J. B. Paris. Initial segments of models of Peano's axioms. In *Set theory and hierarchy theory, V (Proc. Third Conf., Bierutowice, 1976)*, pages 211–226. Lecture Notes in Math., Vol. 619. Springer-Verlag, Berlin, 1977.

[26] R. Kossak and J. B. Paris. Subsets of models of arithmetic. *Arch. Math. Logic*, 32(1):65–73, 1992.

[27] D. Lascar. Les automorphisms d'un ensemble fortement minimal. *J. Symbolic Logic*, 57:238–251, 1992.

[28] A. Lev. Powers of 1-cyclic conjugacy classes in the groups $\mathrm{GL}_n(F)$ and $\mathrm{SL}_n(F)$. *Linear Algebra Appl.*, 257:243–281, 1997.

[29] T. Lindstrøm. An invitation to nonstandard analysis. In N. Cutland, editor, *Nonstandard Analysis and its Applications*. Cambridge University Press, 1988.

[30] H. D. Macpherson and P. M. Neumann. Subgroups of infinite symmetric groups. *J. London Math. Soc.*, 42(1):64–84, 1990.

[31] O. T. O'Meara. *Lectures on linear groups*. American Mathematical Society, Providence, R.I., 1974. Expository Lectures from the CBMS Regional Conference held at Arizona State University, Tempe, Ariz., March 26–30, 1973, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 22.

[32] O. T. O'Meara. *Symplectic groups*, volume 16 of *Mathematical Surveys*. American Mathematical Society, Providence, R.I., 1978.

[33] J. B. Paris and L. Harrington. A mathematical incompleteness in Peano Arithmetic. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 1133–42. Springer-Verlag, Berlin, 1977.

[34] A. Robinson. Non-standard analysis. *Nederl. Akad. Wetensch. Proc. Ser. A 64 = Indag. Math.*, 23:432–440, 1961.

[35] J. J. Rotman. *An introduction to the theory of groups*. Springer-Verlag, New York, 4 edition, 1995.

[36] J. Schreier and S. Ulam. Über die Permutationsgruppe der natürlichen Zahlenfolge. *Studia Math.*, 4:134–141, 1933.

[37] C. A. Smoryński. Recursively saturated nonstandard models of arithmetic. *J. Symbolic Logic*, 46(2):259–286, 1981.

[38] C. A. Smoryński. Lectures on nonstandard models of arithmetic. In *Logic colloquium '82 (Florence, 1982)*, pages 1–70. North-Holland, Amsterdam, 1984.

[39] R. Solomon. On finite simple groups and their classification. *Notices Amer. Math. Soc.*, 42(2):231–239, 1995.

[40] I. Stewart. *Galois Theory, second edition*. Chapman and Hall/CRC, 1989.

[41] J. S. Wilson. On simple pseudofinite groups. *J. London Math. Soc.*, 51(2):471–490, 1995.