ENLARGEMENTS AND THE NON-STANDARD PERSPECTIVE

MARK SCHACHNER

Abstract. The ultraproduct construction is of utmost importance to model theory. In this paper we consider a particular application of the construction– namely, the theory of enlargements– as it pertains to Abraham Robinson's nonstandard analysis and Ramsey theory. We will show the numerous simplifications both of definitions and of proofs which this perspective affords, while assuming only minimal background in mathematical logic and model theory.

Contents

1. Introduction

Model theory is a relatively young branch of mathematics, with the first works explicitly in the subject appearing within the last 80 years, and the first works making implicit use of its techniques appearing in the early 1900s. However, thanks to the power of its methods and a vast array of applications, model theory has developed rapidly since its inception. The goal for this paper will be to exposit one such application, namely Robinson's nonstandard analysis, while assuming only minimal experience in mathematical logic. We will accomplish this via a concise but ground-up introduction of the ultraproduct construction, followed by a general study of enlargements and nonstandard arguments, before finally specializing to the hyperreals and then to Ramsey theory. We hope that, due to our development of nonstandard theory in generality, the transition into the hyperreal field will seem a natural application of the ideas so far explored.

The ultraproduct construction is interesting in part due to its nonconstructive nature. Much of our exposition here will rely on the existence of a specific object– namely, a *nonprincipal ultrafilter* on an infinite set– the existence of which cannot be proved in Zermelo-Fraenkel set theory without the axiom of choice (although we note that the full strength of choice is not required). However, despite the somewhat immodest nature of its construction, the ultraproduct will turn out to be fundamental to our study of nonstandard analysis, largely due to Łós's theorem, which broadly states that ultraproducts have many properties similar to those of their factors.

The reasons for our focus on Robinson's nonstandard analysis are threefold. First, we recognize that much of model theory can seem highly abstract at a first approach, and so through this application we hope to provide a concrete anchor to other fields of mathematics. Second, as noted in our main source [1], much of nonstandard analysis yields novel, insightful proofs of familiar facts, as well as a host of new concepts with significant intuitive backing. Lastly, our study of nonstandard theory in general will allow a concluding diversion into Ramsey theory, demonstrating yet more applications of the principles we exposit in the earlier sections and yielding an unexpected result as we return to pure mathematical logic.

2. Ultrafilters

In the interest of brevity, we omit the standard definitions in mathematical logic and model theory; for a detailed exposition, the reader is encouraged to consult our source [2]. As exposited from these initial definitions alone, it is somewhat unclear how one can construct new models from old ones. The theorems of Löwenheim-Skolem-Tarski give the existence of models of different cardinalities of a single theory, but for our purposes we will need a way to construct a new model from multiple component models, in a manner analagous to the direct product in algebra. The construction we develop in this section, namely the *ultraproduct*, will serve this purpose.

How, in constructing such a composite model, do we determine the theory it satisfies? A statement might have completely different interpretations across our composite model's different component models, and might state something obviously true in one model but obviously false in another. Intuitively, we would like a statement to be true for the composite model if it is true for a"large" subset of the set of component models. However, we will be considering products of infinitely many models. But which subsets of an infinite set can be considered "large"?

Perhaps an appropriate axiomatization of a "large" set would be

- the whole set itself is large;
- any set containing a large subset is itself large; and
- any two large sets must have large intersection (to prevent "crowding" of large sets).

Also, to ensure the composite model is sufficiently complete and consistent, we should require that, if some statement holds for some subset of the component models, then either it or its negation holds in the composite model, but not both. This can be summarized set-theoretically by

- For any set A, either A or its complement is large, but not both.

This is now enough to make a definition. Our formalization of "largeness" is called a *filter* (or *ultrafilter* if it satisfies the fourth condition), and is defined thus:

Definition 2.1 (Filters & Ultrafilters)**.** Let I be a set. A *filter* is a set $\mathcal{F} \subseteq 2^I$, satisfying:

(i) $I \in \mathcal{F}$;
(ii) If $A \in \mathcal{F}$ and $A \subseteq B \subseteq I$, then $B \in \mathcal{F}$; and
(iii) If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

If $\mathcal{F}$ also satisfies

(iv) For all $A \in I$, either $A \in \mathcal{F}$ or $I \setminus A \in \mathcal{F}$, but not both,

then we call $\mathcal{F}$ an *ultrafilter*.

A quick induction gives that property (iii) holds for arbitrary finite intersections as well. There is much to say about filters, but in the interest of brevity we exposit only the following two important properties:

Definition 2.2 (Properties of Filters)**.** Let I be a set, and let $\mathcal{F}$ be a filter over I.

(i) $\mathcal{F}$ is called *improper* if $\varnothing \in \mathcal{F}$; otherwise it is called *proper*.
(ii) $\mathcal{F}$ is called *principal* if there exists $A \subseteq I$ such that

$$\mathcal{F} = \{B \subseteq I : A \subseteq B\}.$$

Otherwise it is called *nonprincipal*.

Additionally, we will need the following fact:

Lemma 2.3. *Let I be a set, and let $\mathcal{E} \subseteq 2^I$. Then the sets*

$$\left\{A \subseteq I : A \supseteq \bigcap Y_i \text{ for some finite } \{Y_i\} \subseteq \mathcal{E}\right\}$$

and

$$\bigcap \{\mathcal{F} : \mathcal{F} \text{ is a filter over } I \text{ containing } \mathcal{E}\}$$

both define the same filter over I.

We will call the resulting filter the *filter generated by $\mathcal{E}$*. We remark that in the first expression we allow for the set $\{Y_i\}$ to be empty, adopting the convention that an empty intersection of sets is equal to I.

Proof. Let $\mathcal{G}$ denote the first expression, and let $\mathcal{H}$ denote the second. We will first show that $\mathcal{G}$ is a filter over I containing $\mathcal{E}$:

- By our remark above, $\mathcal{G}$ contains I, so it has property (i).
- If $A \in \mathcal{G}$, then A contains a finite intersection $\bigcap Y_i$ of elements of $\mathcal{E}$. Thus any superset of A also contains this finite intersection, and so lies in $\mathcal{G}$ as well. Therefore $\mathcal{G}$ has property (ii).
- If $A, B \in \mathcal{G}$, then A contains a finite intersection $\bigcap Y_i$ of elements of E, and B also contains such a finite intersection $\bigcap Z_i$. Thus $A \cap B$ contains the finite intersection $(\bigcap Y_i) \cap (\bigcap Z_i)$, and so lies in $\mathcal{G}$. Therefore $\mathcal{G}$ has property (iii).
- Certainly any element $E \in \mathcal{E}$ is a superset of the finite intersection $\bigcap\{E\}$, so $E \in \mathcal{G}$; thus $\mathcal{G}$ contains $\mathcal{E}$.

Next, we show that $\mathcal{G} \subseteq \mathcal{H}$. Let $A \in \mathcal{G}$, so $A \supseteq \bigcap Y_i$ for some finite $\{Y_i\} \subseteq \mathcal{E}$, and let $\mathcal{F}$ be a filter over I which contains $\mathcal{E}$. Since $\mathcal{F}$ contains $\mathcal{E}$, by property (iii) the intersection $\bigcap Y_i$ lies in $\mathcal{F}$. Then, since A is a superset of this intersection, by property (ii) we have $A \in \mathcal{F}$. Since our choice of $\mathcal{F}$ was arbitrary, this implies that A lies in every filter $\mathcal{F}$ over I which contains $\mathcal{E}$; i.e., $A \in \mathcal{H}$. Thus $\mathcal{G} \subseteq \mathcal{H}$. But since we showed $\mathcal{G}$ is a filter over I containing $\mathcal{E}$, we also have the reverse inclusion; hence $\mathcal{G}$ and $\mathcal{H}$ define the same filter over I. $\square$

Now, we move to two results, crucial to our study of filters. We recall from topology that a set S has the *finite intersection property* if the intersection of any finite, nonempty collection of elements of S is nonempty.

Theorem 2.4 (Ultrafilter Lemma)**.** *Let I be a set, and suppose $\mathcal{E} \subseteq 2^I$ has the finite intersection property. Then $\mathcal{E}$ is contained in some ultrafilter over I.*

Proof. Write $\mathcal{F}$ for the filter generated by $\mathcal{E}$. We note $\mathcal{F}$ is proper; if not then by Lemma 2.3 there would exist some finite subset of $\mathcal{E}$ with empty intersection, which is impossible since $\mathcal{E}$ has the finite intersection property. Now, let $\mathcal{P}$ be the set of proper filters over I containing $\mathcal{F}$, partially ordered by inclusion. We have $\mathcal{P}$ is nonempty since it contains $\mathcal{F}$. Consider a chain $\{\mathcal{F}_\alpha\}$ in $\mathcal{P}$, and write $\mathcal{G} = \bigcup \mathcal{F}_\alpha$. We claim $\mathcal{G}$ is a proper filter over I containing $\mathcal{F}$:

- We have $I \in \mathcal{G}$ since $I \in \mathcal{F}_\alpha$ for any particular α; thus $\mathcal{G}$ has property (i).
- If $A \in \mathcal{G}$, then $A \in \mathcal{F}_\alpha$ for some α, so if $A \subseteq B \subseteq I$ then $B \in \mathcal{F}_\alpha \subseteq \mathcal{G}$; thus $\mathcal{G}$ has property (ii).
- If $A, B \in \mathcal{G}$, then $A \in \mathcal{F}_\alpha, B \in \mathcal{F}_\beta$ for some α, β. But since $\{\mathcal{F}_\alpha\}$ is a chain, we must have either $\mathcal{F}_\alpha \subseteq \mathcal{F}_\beta$ or $\mathcal{F}_\beta \subseteq \mathcal{F}_\alpha$; without loss of generality suppose the former. Then we have $B \in \mathcal{F}_\beta \subseteq \mathcal{F}_\alpha$, so $A, B \in \mathcal{F}_\alpha$. By property (iii), then, $A \cap B \in \mathcal{F}_\alpha \subseteq \mathcal{G}$, so $A \cap B \in \mathcal{G}$. Thus $\mathcal{G}$ has property (iii).
- We have $\mathcal{F} \subseteq \mathcal{G}$ since $\mathcal{F}$ is contained in any $\mathcal{F}_\alpha$. We also have $\varnothing \notin \mathcal{G}$ since $\varnothing \notin \mathcal{F}_\alpha$ for all α. Thus $\mathcal{G}$ is proper and contains $\mathcal{F}$.

Thus we have that any chain in $\mathcal{P}$ has an upper bound in $\mathcal{P}$. By Zorn's Lemma, then, there exists a maximal proper filter containing $\mathcal{F}$; call it $\mathcal{H}$.

We will now show that $\mathcal{H}$ is actually an ultrafilter. Let $A \subseteq I$. We have two cases: $A \in \mathcal{H}, A \notin \mathcal{H}$. In the first case, we cannot have $I \setminus A \in \mathcal{H}$, since this would imply $\varnothing = A \cap (I \setminus A) \in \mathcal{H}$– an impossibility since we showed $\mathcal{H}$ was proper. So suppose $A \notin \mathcal{H}$. Put $\mathcal{E} = \mathcal{H} \cup \{A\}$, and let $\mathcal{H}'$ denote the filter generated by $\mathcal{E}$. Then $\mathcal{H}'$ is a filter which properly contains $\mathcal{H}$, since it contains A while $\mathcal{H}$ does not; by maximality, then, we must have that $\mathcal{H}'$ is improper, i.e., $\varnothing \in \mathcal{H}'$. However, by Lemma 2.3, this implies that $\varnothing$ contains some finite intersection $\{Y_i\}$ of elements of $\mathcal{E}$. These elements cannot all come from $\mathcal{H}$; if they did then their intersection $\varnothing$ would also lie in $\mathcal{H}$, contradicting the properness of $\mathcal{H}$. Therefore $Y_i = A$ for at least one i. Let $Y = \bigcap\{Y_i : Y_i \neq A\}$; then $Y \in \mathcal{H}$ and $Y \cap A$ is empty. This implies $Y \subseteq I \setminus A$, so by property (ii) we must have $I \setminus A \in \mathcal{H}$. In both cases, exactly one of $A, I \setminus A$ lies in $\mathcal{H}$; hence $\mathcal{H}$ is in fact an ultrafilter.

$\square$

Corollary 2.5. *Let $\mathcal{F}$ be an ultrafilter over a set I. Then:*

(i) $\mathcal{F}$ contains a singleton if and only if it contains a finite set.
(ii) $\mathcal{F}$ is principal if and only if it contains a finite set.

Moreover, if I is infinite, then

(iii) there exists a nonprincipal ultrafilter over I, and
(iv) any nonprincipal ultrafilter over I contains all cofinite subsets of I.

Proof.

(i) The forward direction is immediate; to show the backward direction, let $a_1, \ldots, a_n \in I$ and suppose $\{a_i\} \notin \mathcal{F}$ for each $1 \leq i \leq n$. Then $I \setminus \{a_i\} \in \mathcal{F}$ for each $1 \leq i \leq n$, so the intersection

$$I \setminus \{a_1, \ldots, a_n\} = \bigcap_{1 \leq i \leq n} (I \setminus \{a_i\})$$

lies in $\mathcal{F}$. Thus $\{a_1, \ldots, a_n\} \notin \mathcal{F}$. By contrapositive, if $\mathcal{F}$ contains $\{a_1, \ldots, a_n\}$, then it contains some singleton $\{a_i\}$.

(ii) By (i), we need only show the result for singletons. For the forward direction, suppose that $\mathcal{F}$ is principal, so there exists $A \subseteq I$ such that

$$\mathcal{F} = \{B \subseteq I : A \subseteq B\}.$$

Now, suppose for contradiction that A has two distinct elements a, a'. Then $\{a\}, \{a'\}$ are both proper subsets of A, so neither is large; since $\mathcal{F}$ is an ultrafilter this implies that $I \setminus \{a\}, I \setminus \{b\}$ are both large. Thus their intersection $I \setminus \{a, b\}$ is large. But this set does not contain A, a contradiction. Thus we must have $|A| = 1$, so $\mathcal{F}$ contains a singleton.

For the backward direction, note that if $\mathcal{F}$ contains a singleton $\{a\}$ then it contains all supersets of $\{a\}$; moreover, since any proper subset of a singleton is empty, and an ultrafilter is proper, $\mathcal{F}$ cannot contain a proper subset of $\{a\}$. This implies that if $\mathcal{F}$ contains a singleton, then it consists exactly of all supersets of that singleton, and is thus principal.

(iii) Note that if I is infinite then the collection of cofinite subsets of I has the finite intersection property; thus by Theorem 2.4 it is contained in some ultrafilter $\mathcal{F}$. This ultrafilter cannot contain any finite subsets of I since it contains all the cofinite subsets; thus by contrapositive of (ii) $\mathcal{F}$ is nonprincipal.

(iv) Follows directly from the contrapositive of (ii).

□

We conclude this section with a few notational remarks. First, recall that we conceptualized filters as a method of determining which subsets of a set are "large". Indeed, hereafter we will use the phrase "A is large" to mean A lies in a given filter, so long as the particular filter is clear from context. Secondly, we adopt the following notation from [1]: if $\mathcal{F}$ is a filter on a set I and R is a unary relation on I, then we write

$$[\![R(i)]\!] := \{i \in I : R(i)\}.$$

Again, we will only use this notation when the filter and underlying set are both clear from context.

3. The Ultraproduct

In this section, we define and show basic results regarding the ultraproduct. As we saw in the previous section, an ultraproduct will be defined as a model composed of other models, with the truth or falsity of a sentence φ determined by whether the subset of models satisfying φ is large or not. Throughout this section, we fix a set I, an ultrafilter $\mathcal{F}$ on I, and a collection of models $(\mathfrak{A}_i)$ of first-order logic indexed by I.

Definition 3.1. Two elements $(a_i), (b_i) \in \prod \mathfrak{A}_i$ are called *$\mathcal{F}$-equivalent* if the set $[\![a_i = b_i]\!]$ is large.

Recall $[\![a_i = b_i]\!]$ denotes the set $\{i \in I : a_i = b_i\}$.

Lemma 3.2. *$\mathcal{F}$-equivalence is an equivalence relation.*

Proof.
For any $(a_i) \in \prod \mathfrak{A}_i$ we have $[\![a_i = a_i]\!] = I$, which is large; thus the relation is reflexive. For any $(a_i), (b_i) \in \prod \mathfrak{A}_i$, the two sets $[\![a_i = b_i]\!], [\![b_i = a_i]\!]$ are equal, so if one is large so too must be the other; thus the relation is symmetric. Lastly, suppose $(a_i), (b_i), (c_i) \in \prod \mathfrak{A}_i$ such that $[\![a_i = b_i]\!], [\![b_i = c_i]\!]$ are both large. Then $[\![a_i = c_i]\!] \supseteq [\![a_i = b_i]\!] \cap [\![b_i = c_i]\!]$, so $[\![a_i = c_i]\!]$ is large. Thus the relation is transitive. □

Motivated by Lemma 3.2, we write $[a_i]$ for the equivalence class of (a_i) under this relation. With this in hand, we are ready to define the ultraproduct.

Definition 3.3 (The Ultraproduct)**.**
Let I be a set, let $\mathcal{F}$ be an ultrafilter on I, and, for each $i \in I$, let $\mathfrak{A}_i$ be a model of first-order logic. The *ultraproduct of the* $\mathfrak{A}_i$ *modulo* $\mathcal{F}$ is the model $\prod_{\mathcal{F}} \mathfrak{A}_i$ of first-order logic with universe $\{[a_i] : (a_i) \in \prod\{\mathfrak{A}_i\}\}$, and the following interpretation mapping of the function, relation, and constant symbols of first-order logic:

(i) If f is an n-ary function symbol which is interpreted by the map f_i for each $\mathfrak{A}_i$, then the interpretation of f in $\prod_{\mathcal{F}} \mathfrak{A}_i$ is given by
$$([a_i^1], [a_i^2], \ldots, [a_i^n]) \mapsto [f_i(a_i^1, a_i^2, \ldots, a_i^n)].$$
(ii) If R is an n-ary relation symbol which is interpreted by the relation R_i for each $\mathfrak{A}_i$, then the interpretation of R in $\prod_{\mathcal{F}} \mathfrak{A}_i$ is given by
$$R([a_i^1], [a_i^2], \ldots, [a_i^n]) \quad \text{iff} \quad [\![R_i(a_i^1, a_i^2, \ldots, a_i^n)]\!] \text{ is large.}$$
(iii) If C is a constant symbol which is interpreted by the constant c_i for each $\mathfrak{A}_i$, then the interpretation of C in $\prod_{\mathcal{F}} \mathfrak{A}_i$ is given by $[c_i]$.

If all the $\mathfrak{A}_i$ are in fact the same model $\mathfrak{A}$, we call $\prod_{\mathcal{F}} \mathfrak{A}_i$ an *ultrapower* of $\mathfrak{A}$, and denote it $\prod_{\mathcal{F}} \mathfrak{A}$.

We note that in order to make well-defined arguments with ultraproducts, we fix representatives for each $[a_i]$.

The primary reason we care about the ultraproduct construction is due to the following theorem, due to Łós but also called the *fundamental theorem of ultraproducts*:

Theorem 3.4 (Łós)**.** *Let $\mathfrak{A}$ be an ultraproduct of a collection $\{\mathfrak{A}_i\}$ of models of first-order logic; let φ be any sentence. Then $\mathfrak{A} \models \varphi$ if and only if $[\![\mathfrak{A}_i \models \varphi]\!]$ is large.*

The proof of Theorem 3.4 draws heavily on the inductive structure of first-order logic. In the interest of brevity, we will not reproduce it here, directing the reader to [2] for a thorough treatment. For us it will suffice to say that the theorem is easy to show for terms in first-order logic, and then a simple but somewhat tedious induction gives the theorem for formulas and finally sentences.

We conclude this section with a brief demonstration of the power of the ultraproduct construction– a quick proof of the compactness theorem.

Theorem 3.5 (Compactness)**.** *Let T be a theory such that any finite subset of T is satisfiable. Then T is satisfiable.*

Proof. Let I be the set of finite subsets of T; our hypothesis is that each $\Sigma \in I$ has some model $\mathfrak{A}_\Sigma$. Let $\mathcal{E} = \{[\![\varphi \in \Sigma]\!] : \varphi \in T\} \subseteq 2^I$, and let $\mathcal{F}$ be the filter generated by $\mathcal{E}$. Now, we have that any finite intersection

$$[\![\varphi_1 \in \Sigma]\!] \cap [\![\varphi_2 \in \Sigma]\!] \cap \cdots \cap [\![\varphi_n \in \Sigma]\!]$$

is nonempty since it contains the set $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$. Thus, by Theorem 2.4, $\mathcal{E}$ lies in some ultrafilter $\mathcal{G}$.

Now, fix $\varphi \in T$. If φ lies in some $\Sigma \in I$, then $\mathfrak{A}_\Sigma \models \varphi$; thus $[\![\varphi \in \Sigma]\!] \subseteq [\![\mathfrak{A}_\Sigma \models \varphi]\!]$. Since $[\![\varphi \in \Sigma]\!] \in \mathcal{G}$, we have $[\![\mathfrak{A}_\Sigma \models \varphi]\!] \in \mathcal{G}$; by Theorem 3.4, then, $\prod_{\mathcal{G}} \mathfrak{A}_\Sigma \models \varphi$. Hence $\prod_{\mathcal{G}} \mathfrak{A}_\Sigma$ models all $\varphi \in T$.

□

4. Enlargements

In this section, we will begin our development of nonstandard theory in generality, by showing that the ultraproduct allows us to construct an *enlargement* of any set, with certain elements called *nonstandard entities*. These concepts are defined as follows:

Definition 4.1. Let A be a set. The *superstructure* over A is the set $\mathbb{U}(A)$ obtained by taking the power set countably many times, i.e., the set

$$A \cup 2^A \cup 2^{(2^A)} \cup \cdots .$$

Definition 4.2. [Enlargements]

(i) Let $\mathfrak{A} = (A, I_1), \mathfrak{B} = (B, I_2)$ be models of first-order logic. A *transfer map* is a function $(*) : \mathbb{U}(A) \to \mathbb{U}(B)$, along with the following recursive extension to the terms, formulas and sentences of $\mathfrak{A}$:
  - If t is a term of $\mathfrak{A}$, then ${}^*t$ is the term of $\mathfrak{B}$ obtained by replacing all elements $X \in \mathbb{U}(A)$ with ${}^*X$.
  - If φ is a formula or sentence of $\mathfrak{A}$, then ${}^*\varphi$ is obtained by replacing all terms t in φ with ${}^*t$.

(ii) We say $\mathfrak{B}$ is a *enlargement* of $\mathfrak{A}$ if there exists a transfer map as in (i) which satisfies:
  - The *transfer principle*: Let φ be a sentence in $\mathfrak{A}$. Then φ is true iff ${}^*\varphi$ is true.
  - *Countable saturation*: If $\{A_i : i \in \mathbb{N}\}$ is a collection of nonempty subsets of A with the finite intersection property, then $\bigcap_{i\in\mathbb{N}} {}^*A_i$ is nonempty.

We remark that, since the Cartesian product $A \times A$ is encoded in set theory as an element of $2^{2^{(2^A)}}$, the functions and relations in A lie in the domain of the transfer map– e.g., a function $f : A \to A$ has an enlargement ${}^*f : {}^*A \to {}^*A$ in the above definition.

The transfer principle implies a number of useful properties of the map $(*)$. For example, if $A_1, A_2, a \in \mathbb{U}(A)$, then we have $a \in A_1 \cup A_2$ iff $a \in A_1$ or $a \in A_2$. Applying transfer, this implies that ${}^*a \in {}^*(A_1 \cup A_2)$ iff ${}^*a \in {}^*A_1$ or ${}^*a \in {}^*A_2$. This implies that

$$ {}^*(A_1 \cup A_2) = {}^*(A_1) \cup {}^*(A_2). $$

Identical arguments can be used to show that, for any $A_1, A_2 \in \mathbb{U}(A)$,

- $2^{{}^*A_1} = {}^*(2^{A_1})$;
- ${}^*(A_1 \cap A_2) = {}^*A_1 \cap {}^*A_2$;
- ${}^*(A_1 \setminus A_2) = {}^*A_1 \setminus {}^*A_2$; and
- ${}^*\varnothing = \varnothing$.

We observe also that if S is finite, then ${}^*S = \{{}^*s : s \in S\}$. This is because if $S = \{s_1, \ldots, s_n\}$, then membership in S can be expressed by the first-order sentence

$$ (x = s_1) \vee (x = s_2) \vee \cdots \vee (x = s_n). $$

(Note that for infinite S this argument breaks down at this point, since the above sentence would be infinite and thus not a sentence of first-order logic.) Applying transfer to the above sentence gives that $x \in {}^*S$ iff $x = {}^*s_i$ for some $1 \leq i \leq n$.

Beyond these properties, enlargements are interesting because they can contain additional *nonstandard entities* which make certain novel modes of argument possible. Some examples of these nonstandard entities are:

- Let $\mathfrak{A} = (A, \leq)$ be any totally ordered set; then any enlargement $\mathfrak{B}$ must contain some element which bounds A above. To show this, for each $a \in A$ let $[a, \infty) = \{x \in a : a \leq x\}$. Then for any finite collection $\{a_1, \ldots, a_n\}$, the intersection $\bigcap_{i \leq n} [a_i, \infty)$ is the set $\left[\max_{i \leq n} a_i, \infty\right)$, which contains $\max_{i \leq n} a_i$, so is nonempty; thus the set $\{[a, \infty) : a \in A\}$ has the finite intersection property. Since $\mathfrak{B}$ is an enlargement, then, there exists

  $$ b \in \bigcap \{{}^*[a, \infty) : a \in A\}. $$

  By the transfer principle, the set ${}^*[a, \infty)$ is the set $\{\beta \in B : a \leq \beta\}$; therefore b is larger than every element of A.
- Let $\mathfrak{A}$ be a metric space, and let a be a limit point of A. Then any enlargement $\mathfrak{B}$ of $\mathfrak{A}$ contains elements "infinitely close to a" in the following sense. For each $n \in \mathbb{N} \setminus \{0\}$, let $U_n = B_{1/n}(a)$. Then the collection $\{U_n \setminus \{a\} : n \in \mathbb{N}\}$ can be quickly shown to have the finite intersection property. By saturation, then, there exists some $b \in \bigcap_{n \in \mathbb{N}} {}^*(U_n \setminus \{a\})$. By transfer, we have that

  $$ {}^*(U_n \setminus \{a\}) = \{b \in B : {}^*d(a, b) < 1/n\} \setminus \{{}^*a\}, $$

  where ${}^*d$ is the enlargement of the metric to B. This seems to imply that the "distance" between ${}^*a$ and b is less than any positive real number, but nonzero.
- In general, if A is any infinite set, then the set $\{A \setminus \{a\} : a \in A\}$ will have the finite intersection property, and so the intersection $\bigcap {}^*(A \setminus \{a\}) = \bigcap ({}^*A \setminus \{{}^*a\})$ will be nonempty. But no $a \in A$ can lie within this intersection.

In order to demonstrate the existence of enlargements, we need the following result, an immediate corollary of Łós's theorem:

Corollary 4.3. *Let $\mathfrak{B}$ be an ultrapower of $\mathfrak{A}$, indexed by some set I. Then there exists an elementary embedding $(*)$ of $\mathfrak{A}$ into $\mathfrak{B}$, given by ${}^*a = [a]$, the equivalence class of the constant sequence $(a, a, a, \ldots)$.*

Proof. Let $\varphi(x_0, \ldots, x_n)$ be some formula of first-order logic, and let $a_0, \ldots, a_n$ lie in the universe of $\mathfrak{A}$. If $\varphi(a_0, \ldots, a_n)$ holds, then the set $[\![\varphi({}^*(a_0)_i, \ldots, {}^*(a_n)_i)]\!]$ is all of I and thus large. If instead $\varphi(a_0, \ldots, a_n)$ does not hold, then the set $[\![\varphi({}^*(a_0)_i, \ldots, {}^*(a_n)_i)]\!]$ is empty and thus not large. By Theorem 3.4, this implies that $\mathfrak{B} \models \varphi({}^*a_0, \ldots, {}^*a_n)$ iff $\mathfrak{A} \models \varphi(a_0, \ldots, a_n)$. Thus $(*)$ is an elementary embedding. $\square$

Theorem 4.4. *Let $\mathfrak{A}$ be a model over a countable language; let $\mathcal{F}$ be a nonprincipal ultrafilter over $\mathbb{N}$. Then the ultrapower $\prod_{\mathcal{F}} \mathfrak{A}$ is an enlargement of $\mathfrak{A}$.*

Proof. For any $a \in A$, let ${}^*a$ again be the equivalence class of the constant sequence $(a, a, a, \ldots)$. Corollary 4.3 then gives that this is an elementary embedding, so that the transfer principle is satisfied. To show countable saturation, let $\{A^k : k \in \mathbb{N}\}$ be a collection of nonempty subsets of A with the finite intersection property. Then for each $n \in \mathbb{N}$ there exists $x_n \in \bigcap_{i=1}^{n} A_n$; let $x = [x_n]$. Now, we claim $x \in \bigcap {}^*A_i$. To show this, we fix $i \in \mathbb{N}$. Then by construction $x_n \in A_i$ for all $i \geq n$. Thus the set $[\![x_n \in A_i]\!]$ is cofinite, so by Corollary 2.5 it is large. Therefore $x \in {}^*A_i$. Hence $\bigcap {}^*A_i$ contains x and is thus nonempty. $\square$

Theorem 4.4 has a relatively simple statement and proof, but it will allow us to make some highly nontrivial constructions and facilitate the remainder of the paper.

5. The Hyperreals

For this section and the next, we fix a nonprincipal ultrafilter $\mathcal{F}$ over $\mathbb{N}$, which exists by Corollary 2.5. Now that we have developed a sufficient amount of non-standard theory, the definition of the hyperreal numbers is quite brief:

Definition 5.1. The *hyperreal numbers* are the enlargement ${}^*\mathbb{R} = \prod_{\mathcal{F}} \mathbb{R}$, where $\mathbb{R}$ is the usual model of the real numbers over the language of first-order logic.

A hyperreal number is thus an equivalence class of sequences of real numbers, usually identified by one of its representatives. The hyperreals inherit all the functions and relations from the real numbers, and thanks to the transfer principle satisfy all the same first-order properties. However, as mentioned in Section 4, what makes ${}^*\mathbb{R}$ interesting is the nonstandard entities it introduces. For instance, $\mathbb{R}$ is totally ordered, so ${}^*\mathbb{R}$ must contain some element bounding $\mathbb{R}$ above– we will call such elements *unlimited*, and all other hyperreals *limited*. Moreover, since 0 is a limit point of $\mathbb{R}$, there exist nonzero elements of ${}^*\mathbb{R}$ which are nonetheless lesser than any positive real number; we will call such elements *infinitesimal*.

Our first introduction to reasoning using the hyperreals involves the following definitions:

Definition 5.2. Let $x, y \in {}^*\mathbb{R}$.

(i) We say x, y are *infinitely close* if $|x - y|$ is infinitesimal. We write $x \simeq y$. This is an equivalence relation; we call the equivalence class of x the *halo* of x and write $\mathrm{hal}(x)$.

(ii) If x is infinitely close to some element of A, we write $x \simeq A$. If every element of A is infinitely close to some element of A', we write $A \simeq A'$. We write $\mathrm{hal}(A)$ for the set $\bigcup_{x \in A} \mathrm{hal}(x)$.

An important notion when dealing with interactions between $\mathbb{R}$ and ${}^*\mathbb{R}$ is the *shadow* of a hyperreal number. This is defined as the unique real number infinitely close to the given hyperreal. We conclude this section by showing that the shadow of any limited hyperreal is well-defined.

Theorem 5.3. *Let $x \in {}^*\mathbb{R}$ be limited. Then there exists a unique $r \in \mathbb{R}$ such that $x \simeq r$.*

Proof. Let $r = \sup\{\alpha \in \mathbb{R} : \alpha < x\}$. Then there exists no real number strictly greater than r and strictly less than x. This implies that $x - r$ cannot be real; if it were equal to some $\varepsilon \in \mathbb{R}$, then $r + \varepsilon/2$ would be a real number strictly greater than r and strictly less than x. Therefore $x - r$ is infinitesimal. The same argument run in reverse shows that if r is some hyperreal infinitely close to x, then $r = \sup\{\alpha \in \mathbb{R} : \alpha < x\}$. This proves both existence and uniqueness. □

6. A Tour of Non-Standard Analysis

In this section, we present a selection of theorems of real analysis, stated and proved with nonstandard methods. The nonstandard perspective will make some proofs almost immediate, while others will be slightly shortened but much more intuitive. We start by translating a few basic concepts into their analogs in ${}^*\mathbb{R}$:

Theorem 6.1. *Let $\{s_n : n \in \mathbb{N}\}$ be a sequence in $\mathbb{R}$; let $\{s_n : n \in {}^*\mathbb{N}\}$ be its enlargement. Then:*

(i) (s_n) is bounded if and only if s_n is limited for all $n \in {}^*\mathbb{N}$.
(ii) (s_n) converges to $L \in \mathbb{R}$ if and only if $s_n \simeq L$ for all unlimited n.
(iii) (s_n) is Cauchy if and only if $s_m \simeq s_n$ for all unlimited m, n.
(iv) L is a limit point of (s_n) if and only if there exists unlimited n such that $s_n \simeq L$.

Proof.

(i) We have (s_n) is bounded iff there exists some $N \in \mathbb{N}$ such that $-N \leq s_n \leq N$ for all $n \in \mathbb{N}$. By transfer, this is true iff $-N \leq s_n \leq N$ for all $n \in {}^*\mathbb{N}$ as well, that is, iff all s_n are limited.

(ii) We have $(s_n) \to L$ iff for all $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that if $n \geq N_\varepsilon$ then $|s_n - L| < \varepsilon$. Suppose $(s_n) \to L$, and let $N \in {}^*\mathbb{N}$ be unlimited; then $N > N_\varepsilon$ for all $\varepsilon > 0$, so $|s_N - L| < \varepsilon$ for all $\varepsilon > 0$, i.e., $s_N \simeq L$. Conversely, suppose $s_N \simeq L$ for all unlimited N. Then for all $\varepsilon > 0$, there certainly exists a hypernatural N with the first-order property that if n exceeds N then $|s_n - L|$ does not exceed ε (take any unlimited hypernatural, for instance). By transfer, then, there exists a natural number with this property, i.e., $(s_n) \to L$.

(iii) We have (s_n) is Cauchy iff for all $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that if $n, m \geq N_\varepsilon$ then $|s_n - s_m| < \varepsilon$. Suppose (s_n) is Cauchy, and let $N, M \in {}^*\mathbb{N}$ be unlimited; then $N, M > N_\varepsilon$ for all $\varepsilon > 0$, so $|s_N - s_M| < \varepsilon$ for all $\varepsilon > 0$, i.e., $s_N \simeq s_M$. Conversely, suppose $s_N \simeq s_M$ for all unlimited N, M; then there exists a hypernatural with the first-order property that if n, m exceed N then $|s_n - s_m|$ does not exceed ε (take any unlimited hypernatural, for instance).

By transfer, then, there exists a natural number with this property, i.e., (s_n) is Cauchy.

(iv) Suppose L is a limit point of (s_n). Then for all $k \in \mathbb{N}$ there exists $n_k \in \mathbb{N}$ such that $0 < |s_{n_k} - L| < 1/k$. In particular this implies that if $k' > k$ then $0 < |s_{n_{k'}} - L| < 1/k' < 1/k$. Now, let K be unlimited, and let $N = n_K$, i.e., the K-th element of the enlargement of (s_n). Then $N > k$ for all $k \in \mathbb{N}$, so $0 < |s_N - L| < 1/k$ for all $k \in \mathbb{N}$, i.e., $s_N \simeq L$. Conversely, suppose there exists $N \in {}^*\mathbb{N}$ such that $s_N \simeq L$; then for all $\varepsilon > 0$ there exists $N \in {}^*\mathbb{N}$ such that $|s_N - L| < \varepsilon$. By transfer, then, there also exists such an $N \in \mathbb{N}$ for each $\varepsilon > 0$.

□

Having translated these definitions into their analogs in ${}^*\mathbb{R}$, we move to our first proof in real analysis using nonstandard techniques:

Theorem 6.2 (The Cauchy convergence criterion)**.** *Let (s_n) be a sequence in $\mathbb{R}$. Then (s_n) converges if and only if (s_n) is Cauchy.*

Proof. First, suppose (s_n) converges to some L. Then $s_N \simeq L$ for all unlimited N. For any unlimited N, M, then, $s_N \simeq L \simeq s_M$, so by transitivity $s_N \simeq s_M$; thus (s_n) is Cauchy. Conversely, suppose (s_n) is Cauchy, and let $L = s_N$, where N is some unlimited hypernatural. Then for any unlimited M, $s_M \simeq s_N \simeq L$, so $s_M \simeq L$; thus (s_n) converges. □

In this new context, the Cauchy convergence criterion becomes nearly trivial, essentially taking the form "if all s_N are infinitely close to each other, then all s_N are infinitely close to some L"– and we simply take an arbitrary s_N for our choice of L! We next treat an important result concerning convergence, the Bolzano-Weierstrass theorem:

Theorem 6.3 (Bolzano-Weierstrass)**.** *Suppose (s_n) is a bounded sequence in $\mathbb{R}$. Then (s_n) has a limit point L.*

Proof. Take any unlimited hypernatural N; since (s_n) is bounded, we have s_N is limited. Thus, by Theorem 5.3, s_N has a shadow L. This L satisfies $L \simeq s_N$, and is thus a limit point of (s_n). □

Again, our nonstandard treatment makes the result almost immediate. Note that in this case, we have not used the full strength of our characterization of limit points in ${}^*\mathbb{R}$. Putting this and Theorem 5.3 to use, we obtain not just the existence of a limit point, but rather a full characterization of the limit points of (s_n)– namely, they are exactly the shadows of the s_N for all unlimited N.

Before we move to further results, we will need a translation of continuity:

Theorem 6.4. *Let $f : \mathbb{R} \to \mathbb{R}$. Then f is continuous at $c \in \mathbb{R}$ iff $x \simeq c$ implies ${}^*f(x) \simeq {}^*f(c)$.*

Proof. The continuity of f at c is expressed by the first-order property that for all $\varepsilon > 0$ there exists some $\delta_\varepsilon > 0$ such that if $|x - c| < \delta_\varepsilon$ then $|f(x) - f(c)| < \varepsilon$. Suppose this is true, and suppose $x \simeq c$. Then $|x - c| < \delta_\varepsilon$ for all $\varepsilon > 0$, so $|{}^*f(x) - {}^*f(c)| < \varepsilon$ for all $\varepsilon > 0$. Thus $x \simeq c$ implies ${}^*f(x) \simeq {}^*f(c)$. Conversely, suppose ${}^*f(x) \simeq {}^*f(c)$ whenever $x \simeq c$. Fix $\varepsilon > 0$. Then there exists a positive hyperreal δ_ε with the first-order property that if $|x-c| < \delta_\varepsilon$ then $|{}^*f(x)-{}^*f(c)| < \varepsilon$

(take any positive infinitesimal, for instance). By transfer, then, there exists a positive real number δ_ε with this property, so that f is continuous at c. $\square$

Using this translation, the proofs of the following familiar theorems resemble intuition much more clearly:

Theorem 6.5 (Intermediate value theorem)**.** *Let $f : [0,1] \to [0,1]$ be continuous with $f(0) = 0, f(1) = 1$. Then for any $y \in [0,1]$ there exists $x \in [0,1]$ such that $f(x) = y$.*

Proof. Let N be some unlimited hypernatural. We note that for any $n \in \mathbb{N}$ there exists $k \in \{0, 1, \ldots, n-1\}$ such that y lies between $f\left(\frac{k}{n}\right)$ and $f\left(\frac{k+1}{n}\right)$, so by transfer such a $K \in {}^*\mathbb{N}$ exists for N as well. Now, we note that $\frac{K}{N} \simeq \frac{K+1}{N}$, so by continuity $f\left(\frac{K}{N}\right) \simeq f\left(\frac{K+1}{N}\right)$. Thus any hyperreal lying between the two has the same shadow x. This implies that $f(x)$ has the same shadow as y; since both are real we thus have $f(x) = y$. $\square$

Theorem 6.6 (Contraction mapping principle)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz map with constant $c < 1$. Then f has a fixed point.*

Proof. Fix $a \in \mathbb{R}$, let $b = |f(a) - a|$, and inductively define a sequence $\{a_n : n \in \mathbb{N}\}$ by $a_0 = a$, $a_{n+1} = f(a_n)$. A quick induction shows that $|a_n - a_{n-1}| \leq bc^{n-1}$. Thus for each $n \in \mathbb{N}$ we have

$$|a_n - a_0| \leq \sum_{i=1}^{n} |a_i - a_{i-1}| = \sum_{i=1}^{n} bc^{n-1},$$

which is bounded above by $\frac{b}{1-c}$. Let N be an unlimited hypernatural; by boundedness $|a_N - a_0|$ is limited, so a_N is as well. By transfer, we have $|a_{N+1} - a_N| \leq bc^N$; since $c < 1$ and N is unlimited, c^N is infinitesimal, so this implies $a_{N+1} \simeq a_N$.

Now, let x be the shadow of a_N. Then $x \simeq a_N$, so by continuity we have $f(x) \simeq f(a_N) = a_{N+1} \simeq a_N \simeq x$. Since $f(x)$ and x are both real, we hence have $f(x) = x$. $\square$

We conclude our tour through nonstandard analysis by considering the hyperreal translations of open and closed sets, as well as those of boundedness and compactness.

Theorem 6.7 (Basic Topology)**.** *Let $A \subseteq \mathbb{R}$. Then:*

(i) A is open iff $hal(A) \subseteq {}^*A$.
(ii) A is closed iff $\mathbb{R} \cap hal({}^*A) \subseteq A$.
(iii) A is bounded iff ${}^*A$ contains no unlimited elements.
(iv) A is compact iff ${}^*A \simeq A$.

Proof.

(i) Let A be open, and fix $x \in \text{hal}(A)$; then $x \simeq a$ for some $a \in A$. By openness, there exists some real $\varepsilon > 0$ such that $B_\varepsilon(a) \subseteq A$. Now, $|x - a| < \varepsilon$, so $x \in {}^*B_\varepsilon(a) \subseteq {}^*A$. Therefore $\text{hal}(A) \subseteq {}^*A$. Conversely, suppose $\text{hal}(A) \subseteq {}^*A$. Then, for any $a \in A$, there exists $N \in {}^*\mathbb{N}$ such that $|x - a| < 1/N$ implies $x \in {}^*A$; take for instance any unlimited hypernatural. Thus by transfer there exists $N \in \mathbb{N}$ such that $|x - a| < 1/N$ implies $x \in A$, so that A is open.

(ii) Suppose A is closed, and suppose $x \in \mathbb{R} \cap \mathrm{hal}({}^*A)$. Then for all positive $\varepsilon \in \mathbb{R}$ there exists some $a \in {}^*A$ such that $|x - a| < \varepsilon$, so by transfer the same is true of A. Thus, for fixed $\varepsilon > 0$, there exists $a \in A \cap B_\varepsilon(x)$. Thus $x \in \overline{A} = A$. Conversely, suppose $\mathbb{R} \cap \mathrm{hal}({}^*A) \subseteq A$, and suppose $x \in \overline{A}$. Then for all $n \in \mathbb{N}$ there exists $a_n \in A$ such that $|a_n - x| < 1/n$. By transfer, then, if N is any unlimited hypernatural then there exists $a_N \in {}^*A$ such that $|a_N - x| < 1/N$, i.e., $a_N \simeq x$. Thus $x \in \mathbb{R} \cap \mathrm{hal}({}^*A)$, so by hypothesis $x \in A$. Thus A is closed.

(iii) We have A is unbounded iff there exists an unbounded sequence $\{a_n : n \in \mathbb{N}\}$ in A. Let $N \in {}^*\mathbb{N}$ be unlimited; by our translation of bounded sequences, $a_N \in {}^*A$ is then unlimited. Thus A is bounded iff ${}^*A$ contains no unlimited element.

(iv) We have ${}^*A \not\simeq A$ iff there exists $a \in {}^*A$ such that $a \not\simeq A$, i.e., $|a - x|$ is not infinitesimal for all $x \in A$. This in turn is true iff for each $x \in A$ there exists a ball B_x of radius ε_x centered at x not containing a. Now, we note that the set $\mathcal{B} = \{B_x(x) : x \in A\}$ is an open cover of A; let $\{B_1, \dots, B_n\}$ be a finite subset. If this is a subcover, then by transfer we have ${}^*A \subseteq {}^*B_1 \cup \cdots \cup {}^*B_n$. Thus a lies in some ${}^*B_i$, so $|a - x| < \varepsilon_i$, an impossibility. Thus $\mathcal{B}$ has no finite subcover, which is true iff A is not compact.

□

The last characterization, that of compactness, is often called *Robinson's criterion*. It is immensely useful in the simplification of certain theorems, including an almost one-line proof of the Tychonoff theorem in topology. Another example is that of the Heine-Borel theorem, which is usually proved using a careful, somewhat tedious dissection argument, but which can be proven elementarily via nonstandard arguments as follows:

Theorem 6.8 (Heine-Borel theorem)**.** *Let $A \subseteq \mathbb{R}$. Then A is compact if and only if A is closed and bounded.*

Proof. Suppose A is compact, so ${}^*A \simeq A$, and fix $x \in \mathbb{R} \cap \mathrm{hal}({}^*A)$. Then there exist $a \in {}^*A, a' \in A$ such that $x \simeq a \simeq a'$. By transitivity, then, $x \simeq a'$; since x, a' are both real, this implies $x = a' \in A$. Thus $\mathbb{R} \cap \mathrm{hal}({}^*A) \subseteq A$, so A is closed. Moreover, since all elements of ${}^*A$ are infinitely close to $A \subseteq \mathbb{R}$, ${}^*A$ contains no unlimited elements; thus A is bounded.
Conversely, suppose A is closed and bounded, and fix $a \in {}^*A$. Since A is bounded, we have a is limited; let x be its shadow. Then $x \in \mathbb{R}$ and $x \simeq {}^*A$, so since A is closed we have $x \in A$. Thus $a \simeq A$, so $A \simeq {}^*A$. Hence A is compact. □

7. Ramsey's Theorem and Variation

In this conclusory section, we exposit one more interesting application of the nonstandard approach, by proving Ramsey's theorem, an important combinatorial result. The theorem is stated using the following definitions and notation:

Definition 7.1. Let A be a set; for $k \in \mathbb{N}$, we denote by $\binom{A}{k}$ the set of subsets of A of cardinality k.

(i) An (n,k)-*coloring* of A is a cover of $\binom{A}{k}$ by n pairwise disjoint sets, i.e., a collection $\mathcal{C} = \{C_i : 1 \leq i \leq n\}$ such that $\binom{A}{k} \subseteq \bigcup_{i=1}^n C_i$ and $C_i \cap C_j = \varnothing$ for all distinct $i, j \in \{1, \dots, n\}$. If n and k are left unspecified we simply call $\mathcal{C}$ a *coloring* of A.

(ii) If $\mathcal{C}$ is an (n,k)-coloring of A, we say that $B \subseteq A$ is *monochromatic* if $\binom{B}{k} \subseteq C_i$ for some $1 \leq i \leq n$.

With this in hand, Ramsey's theorem is stated thus:

Theorem 7.2 (Ramsey's theorem).
Any coloring of an infinite set admits an infinite monochromatic subset.

Our proof of Theorem 7.2, adapted from [1], will go like this: let $R(n,k)$ be the sentence which states "any (n,k)-coloring of an infinite set admits an infinite monochromatic subset." Then, we show the following three steps:

(i) For all $n, k \in \mathbb{N}$, $R(1,k)$ and $R(n,1)$ hold.
(ii) For all $k \in \mathbb{N}$, $R(2,k)$ implies $R(2,k+1)$.
(iii) For all $n \in \mathbb{N}, n \geq 2$, $R(2,k) \wedge R(n,k)$ implies $R(n+1,k)$.

These facts along with two quick inductions will imply $R(n,k)$ for all $n,k \in \mathbb{N}$. The dependence of our proof on nonstandard theory largely comes from the following lemma:

Lemma 7.3. *Let A be an infinite set, along with some enlargement ${}^*A$. Fix $k \in \mathbb{N}$, let $C \subseteq \binom{A}{k}$, and let N be an unlimited hypernatural. Then there exists a sequence $\{s_n : n \in \mathbb{N}\}$ in A with the following property: if $n_1, \dots, n_k \in \mathbb{N}$ with $n_1 < \cdots < n_k$, then*

$$\{s_{n_1}, \dots, s_{n_{k+1}}\} \in C \quad \textit{iff} \quad \{s_{n_1}, \dots, s_{n_k}, s_N\} \in {}^*C.$$

Proof. Since A is infinite, there exists an injection $f : \mathbb{N} \subseteq A$. Fix $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{N}$, and for $n \geq k$ suppose inductively that $\alpha_1, \dots, \alpha_n \in \mathbb{N}$ are such that for all $n_1 < \cdots < n_{k+1} \leq n$ we have

$$\{f(\alpha_{n_1}), \dots, f(\alpha_{n_{k+1}})\} \in C_1 \quad \text{iff} \quad \{f(\alpha_{n_1}), \dots, f(\alpha_{n_k}), {}^*f(N)\} \in {}^*C_1.$$

Now, let $\varphi_n(x)$ be the following formula:

$$\begin{aligned}\varphi_n(x) = \bigwedge \big\{ &\{f(\alpha_{n_1}), \dots, f(\alpha_{n_k}), f(x)\} \in C_1 : \\ &\quad n_1 < \cdots < n_k \leq n, \{f(\alpha_{n_1}), \dots, f(\alpha_{n_k}), f(N)\} \in {}^*C_1 \big\} \\ \wedge \bigwedge \big\{ &\{f(\alpha_{n_1}), \dots, f(\alpha_{n_k}), f(x)\} \notin C_1 : \\ &\quad n_1 < \cdots < n_k \leq n, \{f(\alpha_{n_1}), \dots, f(\alpha_{n_k}), f(N)\} \notin {}^*C_1 \big\}.\end{aligned}$$

Essentially, φ_n states that, so long as we consider subsequences of $\mathbb{N}$ bounded above by n, the behavior of x with respect to $\mathcal{C}$ matches that of N with respect to ${}^*\mathcal{C}$. We note that φ_n is in fact a formula of first-order logic; since there are only finitely many subsequences of $\{1, \dots, n\}$, there are only finitely many sentences which φ conjoins.

Now, it is certainly true that $\exists N({}^*\varphi_n(N))$ holds; i.e., there exists an element of ${}^*\mathbb{N}$ which exceeds s_n and satisfies ${}^*\varphi_n$. By transfer, then, there exists an element of $\mathbb{N}$ which exceeds α_n and satisfies φ_n. Denote this element by α_{n+1}. By induction this gives an infinite sequence $\{\alpha_n : n \in \mathbb{N}\}$; the image of this sequence under f satisfies the lemma. ∎

Now, we are ready to move to the proof:

Proof of Ramsey's theorem. As stated previously, we proceed in three steps. Throughout this proof, k and n are natural numbers, and A is some infinite set. We also fix an enlargement ${}^*A$ of A, which exists by Theorem 4.4.

Step (i). Suppose $\mathcal{C}$ is a $(1,k)$-coloring of A. This is to say that $\binom{A}{k} \subseteq C$ for some C. But then A itself is a monochromatic subset; thus $R(1,k)$ holds for all k. Next, suppose $\mathcal{C}$ is an $(n,1)$-coloring of A. This means that

$$\binom{A}{1} \subseteq C_1 \cup \cdots \cup C_n$$

for some $C_1, \cdots, C_n$. Now, observe that, since the correspondence $\{a\} \leftrightarrow a$ is bijective, we have

$$|A| = \left|\binom{A}{1}\right| \leq \sum_{i=1}^{n} |C_i|;$$

since A is infinite, the sum on the right side is as well. Since no finite sum of finite cardinals is infinite, this implies at least one C_i is infinite. Taking $B = \{a \in A : \{a\} \in C_i\}$ thus gives $R(n,1)$.

Step (ii). Suppose we have $R(2,k)$; we will show $R(2,k+1)$. Thus we begin with a coloring $\mathcal{C} = \{C_1, C_2\}$ of $\binom{A}{k+1}$. Fix an unlimited hypernatural N; let $S = \{s_n : n \in \mathbb{N}\}$ be the sequence in A given by Lemma 7.3. Now, let $\mathcal{C}' = \{C_1', C_2'\}$ be the $(2,k)$-coloring of S given by

$$\{s_{n_1}, \ldots, s_{n_k}\} \in C_1' \quad \text{if} \quad \{s_{n_1}, \ldots, s_{n_k}, s_N\} \in {}^*C_1,$$

and $\{s_{n_1}, \ldots, s_{n_k}\} \in C_2'$ otherwise. Since we assumed $R(2,k)$, under this coloring S has some infinite monochromatic subset B. Suppose $\binom{B}{k} \subseteq C_1'$; then for any $\{s_{n_1}, \ldots, s_{n_{k+1}}\} \in \binom{B}{k+1}$ we have

$$\begin{aligned}
&\{s_{n_1}, \ldots, s_{n_{k+1}}\} \in \binom{B}{k+1} \\
&\Rightarrow \{s_{n_1}, \ldots, s_{n_k}\} \in \binom{B}{k} \\
&\Rightarrow \{s_{n_1}, \ldots, s_{n_k}\} \in C_1' \\
&\Rightarrow \{s_{n_1}, \ldots, s_{n_k}, s_N\} \in {}^*C_1 \\
&\Rightarrow \{s_{n_1}, \ldots, s_{n_k}, s_{n_{k+1}}\} \in C_1.
\end{aligned}$$

Thus $\binom{B}{k+1} \subseteq C_1$. A similar chain of implications shows that if $\binom{B}{k} \subseteq C_2'$ then $\binom{B}{k+1} \subseteq C_2$. Thus B is an infinite monochromatic subset under the coloring $\mathcal{C}$ as well, so that $R(2,k+1)$ holds.

Step (iii). Suppose $R(2,k)$ and $R(n,k)$ both hold. We will show $R(n+1,k)$ holds; let $\mathcal{C} = \{C_1, \ldots, C_{n+1}\}$ be an $(n+1,k)$-coloring of A. Then let $C_1' = \bigcup_{i=1}^n C_i, C_2' = C_{n+1}$. Then $\mathcal{C}' = \{C_1', C_2'\}$ is a $(2,k)$-coloring of A, so by $R(2,k)$ there exists an infinite monochromatic subset B. Now, we have two cases: $\binom{B}{k} \subseteq C_1', \binom{B}{k} \subseteq C_2'$. In the first case, expanding definitions we get $\binom{B}{k} \subseteq \bigcup_{i=1}^n C_i$; thus by $R(n,k)$ there exists an infinite $B' \subseteq A$ such that $\binom{B'}{k} \subseteq C_i$ for some $1 \leq i \leq n$. In the second case, we already have that $\binom{B}{k} \subseteq C_{n+1}$. Thus in both cases $\mathcal{C}$ admits an infinite monochromatic subset, so that $R(n+1,k)$ holds.

Putting the above steps together, we have the following inductions:

- From (i) we get $R(2,1)$; this along with (ii) implies $R(2,k)$ for all k.
- From (i) we get $R(1,k)$ for all k; this along with the above and (iii) implies $R(n,k)$ for all $n,k \in \mathbb{N}$.

This completes the proof. $\square$

The above nonstandard methods can in fact be used to prove the following variation of Ramsey's theorem, studied by Paris and Harrington in [3]:

Theorem 7.4 (Strengthened finite Ramsey theorem)**.**
Let $k,n,m \in \mathbb{N}$. Then there exists $r \in \mathbb{N}$ such that for any (n,k)-coloring $\mathcal{C} = \{C_1, C_2, \ldots, C_n\}$ of $\{1,\ldots,r\}$ there exists a monochromatic subset B satisfying $|B| \geq m$ and $|B| \geq \min(B)$.

The nonstandard proof of Theorem 7.4 is somewhat shorter than that of Theorem 7.2, but requires nonstandard theory which we will not develop here. A quick and thorough proof using the theory of hyperfinite sets can be found in [1]. We mention Theorem 7.4 here only for the following surprising fact: despite its appearances as a relatively simple and natural statement in combinatorics, this theorem is unprovable from the usual axioms of Peano arithmetic. This provides one last look at the surprising power of nonstandard techniques.

8. Acknowledgments

I am deeply grateful to my mentor, Sarah Reitzes, for providing me with guidance, both as I was introduced to model theory and logic and as I specialized to this topic. I thank her, Isabella Scott, and Gabriela Pinto for their efforts to support the logic cohort during this online REU, as well as Professor May and the numerous other faculty who worked diligently to adapt their lectures and activities to these times. I am also specifically grateful to Prof. Malliaris for her expertise and advice which helped me immensely with my vision for this paper.

References

[1] Goldblatt, R. (2012). Lectures on the hyperreals: an introduction to nonstandard analysis (Vol. 188). Springer Science & Business Media.
[2] Chang, C. C., & Keisler, H. J. (1990). Studies in Logic and the Foundations of Mathematics. Model Theory, Third Edition, 73.
[3] Paris, J.; Harrington, L. (1977). "A Mathematical Incompleteness in Peano Arithmetic". In Barwise, J. (ed.). Handbook of Mathematical Logic. Amsterdam, Netherlands: North-Holland.
[4] Chang, C. C. (1973). H. Jerome Keisler. Model theory. Actes du Congrès International des Mathèmaticiens 1970, Gauthier-Villars, Paris 1971, Vol. 1, pp. 141?150. The Journal of Symbolic Logic, 38(4), 648-648.