The Internal Set Theory Approach to Nonstandard Analysis

Gregory Beauregard

May 4, 2016

Abstract

Although epsilon-delta analysis has enjoyed great success in giving rigor to modern mathematics, nonstandard analysis can be used to rigorously reformulate modern arguments with infinite and infinitesimal arguments. The internal set theory approach to nonstandard analysis is an axiomatic approach that gives us a fresh point of view to look at mathematics. Together, nonstandard analysis with the internal set theory approach allows us to give novel proofs for classically difficult results.

Acknowledgments

First, I would like to express my gratitude toward my thesis advisor, Dr. Paul Fili, for suggesting the topic of nonstandard analysis and supporting me throughout the development of this thesis. Without him, this thesis would not have been possible.

Further, I would like to thank Dr. David Wright for agreeing to be the second reader for this thesis and provide feedback on my work.

I would also like to express gratitude toward all of the mathematics faculty who have contributed to my understanding of classical analysis since without them this thesis could not have begun. In particular I would like to thank Dr. Jiří Lebl for his support in Advanced Calculus.

Additionally, I thank my friends who listened to discussions of my work and provided me with invaluable advice and suggestions.

Finally, I would like to thank my parents and family for always supporting me and encouraging me to pursue my interests.

Contents

1 Introduction

1.1 History

The development of standard epsilon-delta analysis [1] is often regarded as a step forward for mathematics. By doing away with the older infinitesimal approach to analysis, analysis was given the rigor the infinitesimal approach of Euler et al. lacked. However, the infinitesimal approach was remarkably given rigor by Abraham Robinson in 1966.[6] This approach came to be known as nonstandard analysis, or NSA.

In the decades that followed NSA's introduction, a variety of different NSA approaches were developed.[5] In this thesis, I explore and elucidate the internal set theory (IST) approach first introduced by Edward Nelson [4] by applying it to a variety of analysis topics.

1.2 Motivations

As classical analysis has been very successful in giving rigor to modern mathematics, it is valid to ask why we should eschew classical analysis for NSA. This question is a little misguided; instead of thinking about nonstandard analysis as a replacement for classical analysis, one can think of it as an extension of our analysis tool set. Still, we should want motivations for including this tool.

One of the primary motivations for nonstandard analysis is that it significantly simplifies a lot of classically difficult proofs. Even better, it often simplifies the proofs by allowing us to give rigor to intuitive notions of why some theorems are true. It's not an accident that nonstandard analysis helps give rigor to analysis proof intuition. Notions of infinite numbers and infinitesimals were used before the development of modern analysis to give "proofs" for all of calculus. Using nonstandard analysis we can give many of these arguments rigor.

Finally, one of the most important reasons to learn the IST flavor of nonstandard analysis in particular is it represents a radically different viewpoint for thinking about problems. By introducing IST to our toolbox we are free to think about any mathematics problem from a new point of view that may inspire new ideas.

1.3 Preface

In this thesis I give a general introduction to IST suitable for a reader not familiar with any flavor of NSA; however, I will generally assume an existing understanding of basic analysis topics. It is worth note that I often borrowed ideas from and was heavily influenced by the discussion in Robert's *Nonstandard Analysis.*[5] Still, this thesis represents my own approach to discussing and introducing IST.

My discussion in this thesis develops IST and analysis topics while working entirely on $\mathbb{R}$, although this was not strictly necessary. Instead of restricting myself to $\mathbb{R}$, I have generally made a concerted effort to give definitions and proofs that are valid in any metric space or topological space where this would make sense. In doing so, I hope to make it obvious how the results may be extended.

2 Internal Set Theory

2.1 Introduction to IST

The internal set theory, or IST, approach to NSA sets itself apart as an axiomatic approach to NSA. Instead of relying on special constructions to create “infinite” and “infinitesimal” numbers, IST modifies set theory by adding three new axioms and the unary predicate *standard* that reveal new numbers in sets we are already familiar with. The resulting set theory is known as ZFC+IST, after the Zermelo–Fraenkel set theory axioms with the axiom of choice.

In particular, IST adds the Idealization, Standardization, and Transfer axioms. When modifying set theory with new axioms it is natural to be hesitant or question its validity. In some sense, it is quite a radical departure from the usual methods of proving theorems, but the axiomatic approach provides a useful point of view that hides the machinery of other NSA approaches. However, we must be certain when changing set theory that we can rely on existing results; indeed, it has been shown [4] we can. To this end we say ZFC+IST is a conservative extension of ZFC.

Theorem 1. *If a mathematical statement is true without IST, it remains true with IST.*

This result is important conceptually when understanding IST. It will remain true in spite of the new numbers IST reveals. Therefore, we can be assured the axioms added by IST correctly deal with the new concepts in a way that avoids contradictions. This is done in part by the introduction of new terminology.

One of the most important concepts in IST is the notion of standardness. IST introduces the unary predicate *standard* as a new term to describe sets. Just as every set can be described as either finite or infinite, every set is either standard or nonstandard. Notably, each mathematical object we deal with is represented by a set. Hence, functions and numbers are all either standard or nonstandard. Figuring out if a given set is standard or nonstandard is in general a difficult question, just as determining if a given set is finite or infinite can be difficult. However, the IST axioms will provide us a guideline for making this determination.

Here we used the terms finite and infinite to describe the size of sets. As this can be confusing when conflated with the introduction of numbers one might be excused to also deem infinite, we will never refer to the numbers we introduce with such terms. Our terminology choices are in line with the spirit of Theorem 1. For example, one can state without IST, classically, that all elements of $\mathbb{N}$ are finite. It would be confusing to refer to a revealed element of $\mathbb{N}$ as infinite while this remains true.

Another new phrase we will often use is *classical*. Since we are modifying set theory itself for our method, we must be careful to accurately use the terms that describe what we mean.

Definition 1 (Classical)**.** A mathematical statement or formula is classical if it is stated without the use of IST.

One should exercise caution when calling something classical to be certain that it is completely free of IST. This means the statement cannot use IST-exclusive terminology. An example of a common classical formula is the ε definition of a convergent sequence $\{x_n\} \to L$ given by

$$\exists L \in \mathbb{R}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon.$$

One can be certain that if a set is uniquely defined classically, that set is standard. For example, $\mathbb{N}$, $\mathbb{R}$, and 3 are standard sets. We will prove this later, but it is useful to state now.

Finally, note that we must be careful when defining sets using IST terminology. The axiom of specification in ZF that allows us to form subsets of a set satisfying a formula is valid only for classical formulas. For example, we can show with IST there does not exist a set containing only the standard elements of $\mathbb{N}$; forming this set with the axiom of specification would be a contradiction. The takeaway is that nonclassical properties are not necessarily set-forming. In practice, this won't limit us as much as it would seem since the Standardization axiom makes a good substitute.

2.2 The IST Axioms

A solid understanding of the IST axioms is vital for working with IST. Although there are only three axioms, they reveal a rewarding NSA point of view. We will list each axiom here before going through them in detail.

The following notation will be useful when discussing and using the axioms of IST.

$\forall^{\mathsf{s}} x \ \rightarrow$	for all standard x	$\exists^{\mathsf{s}} x \ \rightarrow$	exists a standard x
$\forall^{\mathsf{f}} x \ \rightarrow$	for all finite x	$\exists^{\mathsf{f}} x \ \rightarrow$	exists a finite x
$\forall^{\mathsf{sf}} x \ \rightarrow$	for all standard, finite x	$\exists^{\mathsf{sf}} x \ \rightarrow$	exists a standard, finite x

Axiom 1 (Idealization)**.** *The following are equivalent for a classical formula $R(x, y)$:*

(i) For every standard and finite F there exists an x such that $R(x, y)$ is true for all $y \in F$.

(ii) There exists an x such that $R(x, y)$ is true for all standard y.

Axiom 2 (Standardization)**.** *If E is a standard set and P is a property of E, then there exists a unique standard $A \subseteq E$ denoted by ${}^{\mathsf{s}}\{x \in E : P(x)\}$ with the property that the standard elements of A are the standard elements of E satisfying P.*

Axiom 3 (Transfer)**.** *If F is a classical formula with fixed standard parameters $a_1, \ldots, a_n$, then $F(x, a_1, \ldots, a_n)$ is true for all x if and only if it is true for all standard x.*

2.2.1 Idealization Axiom

Axiom (Idealization)**.** *The following are equivalent for a classical formula* $R(x, y)$*:*

(i) $\forall^{\mathsf{sf}} F, \exists x, \forall y \in F, R(x, y)$

(ii) $\exists x, \forall^{\mathsf{s}} y, R(x, y)$

The Idealization axiom explicitly provides for the existence of nonstandard elements of familiar sets. To see this, we will use Idealization to show there are nonstandard elements of $\mathbb{N}$. As we mentioned before, uniquely classically definable sets are all standard, so we can assume for now when interpreting our results that the classically constructed elements of $\mathbb{N}$ (such as 4 and 7) are standard; as we will see, this is a consequence of the Transfer axiom.

Consider the classical formula $R(x, y)$ given by

$$y \in \mathbb{N} \implies (x \in \mathbb{N} \wedge x > y).$$

This meets the condition of Idealization if

$$\forall^{\mathsf{sf}} F, \exists x, \forall y \in F, y \in \mathbb{N} \implies (x \in \mathbb{N} \wedge x > y).$$

In other words, it meets the condition of Idealization if for every standard and finite set we can find a $x \in \mathbb{N}$ larger than every member of that set in $\mathbb{N}$. If $F \cap \mathbb{N} = \varnothing$, $R(x, y)$ is vacuously true; otherwise, since every finite subset of $\mathbb{N}$ has a maximum, we can choose $x = \max(F \cap \mathbb{N}) + 1$ to meet the first Idealization condition.

Since we have met the first condition of Idealization, we know the second condition must be true. Therefore,

$$\exists x, \forall^{\mathsf{s}} y, y \in \mathbb{N} \implies (x \in \mathbb{N} \wedge x > y).$$

That is, there exists an $x \in \mathbb{N}$ strictly greater than every standard $y \in \mathbb{N}$. It follows this x is necessarily nonstandard.

One might naively call this nonstandard number infinite; after all, it meets the driving concept behind what we think an infinite number should be. However, we are reserving the term infinite for what it is already used for in sets; e.g., $\mathbb{N}$ is infinite. In fact, the nonstandard number we revealed is finite; after all, every natural number is finite (see Theorem 1). Instead, we will call this number *unlimited*.

Definition 2 (Unlimited)**.** $x \in \mathbb{R}$ is unlimited $\iff \forall^{\mathsf{s}} y > 0, |x| > y$.

Definition 3 (Limited)**.** $x \in \mathbb{R}$ is limited $\iff$ x is not unlimited.

As our definition of unlimited suggests, we can easily use Idealization with the appropriate formulas on $\mathbb{R}$ to reveal unlimited positive and negative real numbers similarly to how we did for $\mathbb{N}$.

Another type of number often used in NSA is *infinitesimals*. Positive infinitesimals can be revealed in $\mathbb{R}$ with Idealization using $E = \{x \in \mathbb{R} : x > 0\}$ and $R(x, y)$ given by $y \in E \implies (x \in E \wedge x < y)$. Negative infinitesimals are revealed similarly.

Definition 4 (Infinitesimal)**.** $x \in \mathbb{R}$ is infinitesimal $\iff \forall^{\mathsf{s}} \varepsilon > 0, |x| < \varepsilon$.

Note that with this definition, 0 is an infinitesimal. It is useful to define a notation for the concept of two numbers differing by not more than an infinitesimal.

Definition 5 (Infinitely Close)**.** The following are equivalent:

(i) x is infinitely close to y

(ii) $x \simeq y$

(iii) $\forall^{\mathsf{s}} \varepsilon > 0, |x - y| < \varepsilon$

We have defined a number as limited if it is bounded above by a standard number. Although we will often use that definition, the following characterization of this property on $\mathbb{R}$ more readily generalizes.

Definition 6 (Near Standard)**.** $x \in \mathbb{R}$ is near standard $\iff$

$$\exists^{\mathsf{s}} y \in \mathbb{R}, x \simeq y.$$

As we will often be working with unlimited and limited numbers, the following notation will be useful.

$\forall^{\ell} x$	$\rightarrow$	for all limited x	$\exists^{\ell} x$	$\rightarrow$	exists a limited x
$\forall^{\mathsf{u}} x$	$\rightarrow$	for all unlimited x	$\exists^{\mathsf{u}} x$	$\rightarrow$	exists an unlimited x
$\forall^{\mathsf{ns}} x$	$\rightarrow$	for all near standard x	$\exists^{\mathsf{ns}} x$	$\rightarrow$	exists a near standard x

We have shown there exist nonstandard elements in $\mathbb{N}$ and $\mathbb{R}$, so it is natural to wonder when sets contain nonstandard elements.

Theorem 2. *Every infinite set contains nonstandard elements.*

Proof. Consider Idealization with the infinite set E and formula $R(x, y)$ given by $x \in E \wedge x \neq y$. The Idealization condition is met if

$$\forall^{\mathsf{sf}} F, \exists x, \forall y \in F, x \in E \wedge x \neq y.$$

To show such an x exists, note that if F is finite and E is infinite, $E \setminus F$ is nonempty so we can pick any element of this set for our x. Therefore, Idealization provides us an $x \in E$ such that $x \neq y$ for all standard y. In other words, E contains a nonstandard element. To show E contains more than one nonstandard element, note that for a given nonstandard x, $E \setminus \{x\}$ is still infinite. □

Theorem 2's contrapositive provides the following suggestive corollary we will extend in section 2.2.3.

Corollary 3. *If every element of a set is standard, that set is finite.*

As a final result of the Idealization axiom, we will show there exists a finite set containing every standard set. This may at first seem quite surprising in light of Russell's paradox and the corresponding nonexistence of a universal set, but this set is not itself a universal set and a careful treatment of the axiom of specification prevents a contradiction.

Theorem 4. *There is a finite set containing every standard set.*

Proof. Consider Idealization with the formula $R(x, y)$ given by x is finite and $y \in x$. We require that

$$\forall^{\mathsf{sf}} F, \exists x, \forall y \in F, x \text{ is finite} \wedge y \in x.$$

We may simply choose $x = F$ in which case Idealization applies and furnishes

$$\exists x, \forall^{\mathsf{s}} y, x \text{ is finite} \wedge y \in x.$$

In other words, there exists a finite set x that contains every standard set. □

Although Theorem 4 is itself suggestive, the following corollary shows how it will be used in practice.

Corollary 5. *Every set has a finite subset containing all its standard elements.*

Proof. Denote a finite set containing all standard sets from Theorem 4 as x. A finite subset of E containing all its standard elements is given by $x \cap E$. □

As an example, Corollary 5 states there is a finite set that contains every standard element of $\mathbb{R}$. Although counterintuitive, Theorem 4 and Corollary 5 suggest that in some sense we can think of the standard elements as being the accessible elements of a set.

2.2.2 Standardization Axiom

Axiom (Standardization)**.** *If E is standard and P a property, then*

$$\exists^{\mathsf{s}} A \subseteq E, \forall^{\mathsf{s}} x, x \in A \iff x \in E \wedge P(x).$$

It's important to note the property P does not need to be classical. Note we dropped uniqueness of A from this statement of Standardization; we will show later uniqueness is a consequence of the Transfer axiom. We use ${}^{\mathsf{s}}\{x \in E : P(x)\}$ to denote the unique set that Standardization provides.

Since $\mathbb{N}$ is standard (we've merely stated this; it's a consequence of the Transfer axiom), we can use it with the Standardization axiom. Let $E = \mathbb{N}$ and $P(x)$ be x is standard. Standardization then provides us a unique standard subset of $\mathbb{N}$ whose standard elements are the standard elements of $\mathbb{N}$. Since $\mathbb{N}$ is standard and the standard A we constructed is unique, we can conclude A is in fact $\mathbb{N}$ itself.

2.2.3 Transfer Axiom

Axiom (Transfer)**.** *If $F(x, a_1, \dots, a_n)$ is a classical formula with fixed standard parameters $a_1, \dots, a_n$, then*

$$\forall x, F(x, a_1, \dots, a_n) \iff \forall^{\mathsf{s}} x, F(x, a_1, \dots, a_n).$$

Note there are n parameters here. We should take care not to write down a formula with an unlimited number of parameters, but we might run out of pens before this becomes an issue.

We will use the following consequence of Transfer often enough that it's worth stating formally.

Corollary 6 (Dual Form of Transfer)**.** *If $F(x, a_1, \dots, a_n)$ is a classical formula with fixed standard parameters $a_1, \dots, a_n$, then*

$$\exists x, F(x, a_1, \dots, a_n) \iff \exists^{\mathsf{s}} x, F(x, a_1, \dots, a_n).$$

Proof. Take the contrapositive of the Transfer axiom. □

If there is a unique x satisfying such a formula, it is a consequence of Corollary 6 that this x is standard. Therefore, every set uniquely defined by a classical formula is standard.

Theorem 7. *If a set is uniquely defined by a classical formula with standard parameters, it is standard.*

Proof. If our set is defined uniquely by a classical formula with standard parameters, it meets the conditions of Corollary 6 with the set itself as x. Since our set exists and is unique, Corollary 6 tells us it's standard. □

From Theorem 7, we conclude that $\mathbb{N}$ is standard since it is uniquely defined by a classical formula. Further, numbers like 5 and 28 are standard by the same reasoning. In referencing Theorem 7, e.g. to prove $x + y$ is standard if x and y are standard, we will often simply cite Dual Transfer instead.

Theorem 8. 0 *is the only standard infinitesimal.*

Proof. Consider the notion of an infinitesimal x in Definition 4 given by

$$\forall^{\mathsf{s}}\varepsilon > 0, |x| < \varepsilon.$$

If we fix x to a standard number, this is a classical formula of ε with fixed standard parameters. In this case, Transfer applies and states

$$\forall\varepsilon > 0, |x| < \varepsilon.$$

This is a classical statement, and the only such x is 0. □

Theorem 9. *If A and B are standard, $A = B \iff A$ and B have the same standard elements.*

Proof. Given A, B standard, consider the classical formula $x \in A \implies x \in B$. Then, Transfer applies to the statement

$$\forall x, x \in A \implies x \in B.$$

This is exactly the statement $A \subseteq B$. Transfer tells us this is equivalent to

$$\forall^{\mathsf{s}} x, x \in A \implies x \in B.$$

Since $A = B$ exactly when $A \subseteq B \wedge B \subseteq A$, we conclude $A = B$ when

$$\forall^{\mathsf{s}} x, x \in A \iff x \in B.$$

□

Corollary 10. *The standard set provided by Standardization is unique.*

Proof. Standardization provides a standard set A and defines its standard elements. By Theorem 9, a standard set is uniquely defined by its standard elements. Therefore, A is unique. $\square$

This brings us to a point we brushed aside earlier. Idealization concerns statements about standard and finite sets. In using it with sets like $\mathbb{N}$, we assumed when interpreting our theorems that standard and finite sets contained only standard elements and that those standard elements were the usual numbers we are familiar with. In Theorem 7 we showed that the uniquely classically definable sets of e.g. π and 9 were standard. This proves the standard elements are the usual familiar elements as claimed, but we haven't yet justified that every standard and finite set contains only these standard elements. Corollary 3 is certainly related when it states that if every element of a set is standard, that set is finite; however, we can do better.

Theorem 11. *The following are equivalent:*

(i) E *is standard and finite*

(ii) $\forall x \in E$*,* x *is standard*

Proof. We showed in the proof of Theorem 2 by Idealization that the statement

$$\exists x \in E,\ x \text{ is nonstandard}$$

is equivalent to

$$\forall^{\mathsf{sf}} F, \exists x, \forall y \in F, x \in E \wedge x \neq y.$$

This itself is equivalent to

$$\forall^{\mathsf{sf}} F, \exists x \in E, x \notin F.$$

Finally, this is equivalent to

$$\forall^{\mathsf{sf}} F, E \not\subseteq F.$$

We can take the contrapositive of the first and last item in the equivalence to show the following lemma we will use to complete the theorem.

Lemma 12. *$\forall x \in E$, x is standard $\iff \exists^{\mathsf{sf}} F, E \subseteq F$.*

Suppose E is standard and finite, then $\Longleftarrow$ of Lemma 12 with $F = E$ shows $\forall x \in E$, x is standard. This proves that (i) $\implies$ (ii).

Suppose $\forall x \in E$, x is standard. Then, Lemma 12 states $\exists^{\mathsf{sf}} F, E \subseteq F$. F standard $\implies$ $\mathscr{P}(F)$ standard by Dual Transfer and F finite $\implies$ $\mathscr{P}(F)$ finite classically. Since $\mathscr{P}(F)$ is standard and finite, we know by (i) $\implies$ (ii) that every element of $\mathscr{P}(F)$ is standard. Since E is an element of $\mathscr{P}(F)$, E is standard. Since $E \subseteq F$ and F is finite, E is finite as well. Therefore, we have (ii) $\implies$ (i). □

We have stated in Theorem 1 that anything true without IST is still true with IST; however, we must occasionally take care in its application. Consider the statement $P(n)$ given by n is standard. Suppose n is standard, then $n+1$ is standard by Dual Transfer. Therefore, $\forall n \in \mathbb{N}, P(n) \implies P(n+1)$ since the statement is true vacuously when n is not standard. We know $P(1)$ since 1 is standard, so a naive application of the induction principle would lead us to conclude $\forall n \in \mathbb{N}, n$ is standard. Since we know this to be false, we should examine what went wrong.

The induction principle follows from the fact that $\mathbb{N}$ is well-ordered. That is, every nonempty subset of $\mathbb{N}$ has a least element. This is still true, provided the subset of $\mathbb{N}$ exists. As we have mentioned before, nonclassical formulas cannot be used with the axiom of specification. As a result, we cannot use nonclassical properties with induction and anticipate our findings will be well-defined or true in general. Therefore, we must restrict our use of the induction principle to classical properties from this point forward.

Although we must constrain ourselves when applying induction, we can formulate a sort of nonstandard induction valid for any property, classical or not.

Theorem 13 (Nonstandard Induction)**.** *If $P(n)$ is a property such that $P(1)$ holds and $\forall^{\mathsf{s}} n \in \mathbb{N}, P(n) \implies P(n+1)$, then $\forall^{\mathsf{s}} n \in \mathbb{N}, P(n)$.*

Proof. Suppose $P(1)$ holds and $\forall^{\mathsf{s}} n \in \mathbb{N}, P(n) \implies P(n+1)$. Consider the standard set $A = {}^{\mathsf{s}}\{n \in \mathbb{N} : P(n)\}$. Our assumptions ensure that

$$\forall^{\mathsf{s}} n \in \mathbb{N}, n \in A \implies n+1 \in A.$$

Since A is standard, Transfer applies and furnishes

$$\forall n \in \mathbb{N}, n \in A \implies n+1 \in A.$$

Since our assumptions ensure $1 \in A$ and this is a classical formula, induction applies and states $\forall n \in \mathbb{N}, n \in A$. That is, $A = \mathbb{N}$. Standardization ensures the standard elements of A are exactly the standard elements satisfying $P(n)$, so we may conclude that $\forall^{\mathsf{s}} n \in \mathbb{N}, P(n)$. $\square$

2.3 Axiom of Choice

It is often interesting to note whether or not mathematics proofs invoke the axiom of choice. In this paper we have used ZFC+IST and mentioned that ZFC+IST is a conservative extension of ZFC. However, ZF+IST is not a conservative extension of ZF. In fact, there are many things we can prove with ZF+IST that are unprovable with ZF alone as they would normally require the axiom of choice.

One of the most notable things provable in ZF+IST that is unprovable in ZF is the boolean prime ideal theorem, or BPI. Taken as an axiom, ZF+BPI is strictly weaker than ZFC while still being strong enough to prove many important results usually said to rely on the axiom of choice. Although many theorems can be proven equivalent to BPI, a notable example is the ultrafilter lemma [2]; we will prove the ultrafilter lemma using ZF+IST.

Definition 7 (Filter)**.** $\mathcal{F} \subseteq \mathscr{P}(X)$ is a filter on the set X if it meets the following conditions:

(i) $X \in \mathcal{F}$.

(ii) $\varnothing \notin \mathcal{F}$.

(iii) $\forall (A,B) \in \mathcal{F} \times \mathscr{P}(X), A \subseteq B \implies B \in \mathcal{F}$.

(iv) $\forall A, B \in \mathcal{F}, A \cap B \in \mathcal{F}$.

Definition 8 (Ultrafilter)**.** A filter $\mathcal{U}$ on X is an ultrafilter if it meets the following additional condition:

(v) $\forall A \subseteq X, A \in \mathcal{U} \vee X \setminus A \in \mathcal{U}$.

Lemma 14 (Ultrafilter Lemma)**.** *Every filter on a set X is a subset of an ultrafilter on X.*

Theorem 15. $\mathsf{ZF}+\mathsf{IST}$ *implies the ultrafilter lemma.*

Proof. Assume $\mathsf{ZF}+\mathsf{IST}$. By Transfer, the ultrafilter lemma holds for every filter on a set X if and only if it holds for every standard filter on a standard set X. Therefore, let $\mathcal{F}$ be a standard filter on a standard set X.

Consider the classical formula $R(x,y)$ such that the Idealization condition is given by

$$\forall^{\mathsf{sf}} F, \exists x, \forall y \in F, y \in \mathcal{F} \implies (x \in X \wedge x \in y).$$

To meet this condition we need to show that for any standard and finite $F \subseteq \mathcal{F}$ there is an $x \in X$ contained in every element of F. Condition (iv) of Definition 7 implies that since F is finite, $\bigcap_{y \in F} y \in \mathcal{F}$. Since $\varnothing \notin \mathcal{F}$, $\bigcap_{y \in F} y \neq \varnothing$. Therefore, there is an $x \in X$ in every $y \in F$. Idealization applies and states

$$\exists x, \forall^{\mathsf{s}} y, y \in \mathcal{F} \implies (x \in X \wedge x \in y).$$

That is, there is an $x \in X$ satisfying $x \in y$ for every standard $y \in \mathcal{F}$. Taking this x, use Standardization to define

$$\mathcal{U} = {}^{\mathsf{s}}\{A \in \mathscr{P}(X) : x \in A\}.$$

We prove $\mathcal{U}$ is an ultrafilter satisfying $\mathcal{F} \subseteq \mathcal{U}$. Consider the statement $\forall^{\mathsf{s}} A \in \mathcal{F}, A \in \mathcal{U}$. By Idealization, every standard $A \in \mathcal{F}$ satisfies $x \in A$.

Since this A is standard, it is also in $\mathcal{U}$ by the Standardization condition. Therefore, our statement holds. Since the statement is a classical formula of A with fixed standard parameters, Transfer applies and tells us $\forall A \in \mathcal{F}, A \in \mathcal{U}$. That is, $\mathcal{F} \subseteq \mathcal{U}$. Next we prove $\mathcal{U}$ is an ultrafilter by checking it satisfies each condition of Definition 7 and Definition 8.

(i): Since X is standard and $x \in X$, $X \in \mathcal{U}$ by our Standardization condition.

(ii): Since $\varnothing$ is standard, $\varnothing \in \mathcal{U}$ if and only if it satisfies $x \in \varnothing$ by our Standardization condition. Since this is false, $\varnothing \notin \mathcal{U}$.

(iii): Consider the statement

$$\forall (A, B) \in \mathcal{U} \times \mathscr{P}(X), A \subseteq B \implies B \in \mathcal{U}.$$

This is a classical formula of (A, B) with fixed standard parameters, so Transfer applies and tells us this is equivalent to

$$\forall^{\mathsf{s}} (A, B) \in \mathcal{U} \times \mathscr{P}(X), A \subseteq B \implies B \in \mathcal{U}.$$

Given a standard A, B with $A \in \mathcal{U}$ and $A \subseteq B$, we know that $x \in A$ by the Standardization condition that ensures the standard A in $\mathcal{U}$ are exactly the standard elements of $\mathscr{P}(X)$ satisfying $x \in A$. Since $x \in A$, $A \subseteq B$ ensures $x \in B$. Since B is standard, $B \in \mathcal{U}$ by the Standardization condition.

(iv): Consider the statement

$$\forall A, B \in \mathcal{U}, A \cap B \in \mathcal{U}.$$

This is a classical formula of A, B with fixed standard parameters, so Transfer applies and tells us this is equivalent to

$$\forall^{\mathsf{s}} A, B \in \mathcal{U}, A \cap B \in \mathcal{U}.$$

Suppose $A, B \in \mathcal{U}$ are standard. Since A and B are standard, $x \in A$ and $x \in B$ by the Standardization condition. Therefore, $x \in A \cap B$. Since A and B are standard, $A \cap B$ is standard by Dual Transfer. Therefore, $A \cap B \in \mathcal{U}$ by the Standardization condition.

(v): Consider the statement

$$\forall A \subseteq X, A \in \mathcal{U} \vee X \setminus A \in \mathcal{U}.$$

This is a classical formula of A with fixed standard parameters, so Transfer applies and tells us this is equivalent to

$$\forall^{\mathsf{s}} A \subseteq X, A \in \mathcal{U} \vee X \setminus A \in \mathcal{U}.$$

Suppose $A \subseteq X$ is standard. If $x \in A$, then $A \in \mathcal{U}$ by the Standardization condition and we are done. If $x \notin A$, then $x \in X \setminus A$. We know $X \setminus A$ is standard by Dual Transfer, so $X \setminus A \in \mathcal{U}$ by the Standardization condition.

Therefore, every standard filter on a standard set X is a subset of an ultrafilter on X. By Transfer, this is equivalent to the ultrafilter lemma. □

By Theorem 15, we can conclude ZF+IST implies BPI. However, we have left open the question of whether or not ZF+IST can be used to prove something ZF+BPI cannot. Actually, it can't; it has been shown ZF+IST is a conservative extension of ZF+BPI.[2] That is, the things provable with ZF+IST are exactly the things provable with ZF+BPI.

3 Nonstandard Analysis

3.1 Introduction to NSA

The techniques and results of Internal Set Theory allow us to develop analysis with new definitions resulting in a discussion of the topic we will call nonstandard analysis, or NSA. In some cases, the standard analysis proofs and techniques have direct NSA analogues. However, in other cases, NSA proofs can provide a fresh approach to a classical problem.

3.1.1 Sequences

The usual definition for a convergent sequence $\{x_n\} \to L$ is

$$\exists L \in \mathbb{R}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon.$$

We can give an equivalent definition of a convergent sequence using IST. This definition makes rigorous the intuitive notion of a sequence being infinitely close to its limit at unlimited n.

Theorem 16. *The following are equivalent for a standard* $\{x_n\}$*:*

(i) $\lim_{n\to\infty} x_n = L$

(ii) $\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{u}} n \in \mathbb{N}, x_n \simeq L$

Proof. We will first show (i) $\implies$ (ii). (i) means

$$\exists L \in \mathbb{R}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon.$$

Since our sequence $\{x_n\}$ is standard and converges to a unique number by a classical formula, its limit L must be standard by Dual Transfer. Further, if the formula is true $\forall \varepsilon$, we can choose only standard ε to show

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{s}} \varepsilon > 0, \exists N \in \mathbb{N}, [\forall n \geq N, |x_n - L| < \varepsilon].$$

As all parameters of the bracketed formula of N are now standard, Dual Transfer tells us $\exists N$ as soon as $\exists^{\mathsf{s}} N$. Therefore, we may say

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{s}} \varepsilon > 0, \exists^{\mathsf{s}} N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon.$$

If N is standard, then $n \geq N$ is satisfied for all unlimited n. In this case, we may write

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{s}} \varepsilon > 0, \forall^{\mathsf{u}} n \in \mathbb{N}, |x_n - L| < \varepsilon.$$

Using Definition 5, we may rewrite this as

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{u}} n \in \mathbb{N}, x_n \simeq L.$$

This is exactly (ii).

Next, we will show (ii) $\implies$ (i). Assuming (ii) we have

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{u}} n \in \mathbb{N}, x_n \simeq L.$$

Then, we know the following holds

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{s}} \varepsilon > 0, [\exists N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon].$$

To see this is true, note we can pick any unlimited $N \in \mathbb{N}$ and the assumption that $x_n \simeq L$ in this case guarantees $|x_n - L| < \varepsilon$ since ε is standard. The bracketed expression is a classical formula of ε with standard parameters, so Transfer tells us this formula is true $\forall^{\mathsf{s}} \varepsilon > 0$ as soon as it is true $\forall \varepsilon > 0$. Therefore, we have

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon.$$

Finally, there exists a standard L implies there exists an L, so we may conclude

$$\exists L \in \mathbb{R}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - L| < \varepsilon.$$

This is exactly (i). □

Therefore, Theorem 16 formalizes the intuitive notion of convergent sequences being infinitely close to their limit at every unlimited term. The process of finding the limit of a known convergent sequence is therefore equivalent to finding the standard L the sequence is infinitely close to at a given unlimited term. This is where the notion of a standard part of a number comes in useful.

3.1.2 The Standard Part

A useful concept in NSA is the notion of a standard part of a number. The *standard part* of a limited number is the unique standard number it is infinitely close to. The existence and uniqueness of this number is an important cornerstone of NSA that many of its proofs rely on.

Definition 9. If $x \in \mathbb{R}$ is limited, the standard part of x is denoted x^* or $\mathrm{st}(x)$ and defined as

$$\mathrm{st}(x) = \sup {}^{\mathsf{s}}\{y \in \mathbb{R} : y \leq x\}.$$

Theorem 17. *If $x \in \mathbb{R}$ is limited, the standard part exists and is standard.*

Proof. If $I = {}^{\mathsf{s}}\{y \in \mathbb{R} : y \leq x\}$ is bounded above, then the supremum exists. Further, since we know ${}^{\mathsf{s}}\{y \in \mathbb{R} : y \leq x\}$ is standard and unique (by Standardization) and the supremum is unique, then we know if the supremum exists it must be standard by Dual Transfer. Therefore, we just need to show I is bounded above.

Since x is limited, $\exists^{\mathsf{s}} M \in \mathbb{R}, |x| \leq M$. If $t \in I$ is standard, then we know

$$t \leq x \leq |x| \leq M.$$

Therefore, we know $\forall^{\mathsf{s}} t \in I, t \leq M$. Since I and M are standard, Transfer applies and we may conclude $\forall t \in I, t \leq M$. Hence, I is bounded above ensuring $\sup(I) = x^*$ exists and is standard. ☐

Theorem 18. *If the standard part of x exists, then $x \simeq x^*$.*

Proof. Let $I = {}^{\mathsf{s}}\{y \in \mathbb{R} : y \leq x\}$ so that $x^* = \sup(I)$. Suppose for the sake of contradiction there is a standard $\varepsilon > 0$ such that $x - x^* \geq \varepsilon$. Then,

$$x^* < x^* + \varepsilon \leq x.$$

Since x^* and ε are standard, $x^* + \varepsilon$ is standard by Dual Transfer. Further, since $x^* + \varepsilon \leq x$, $x^* + \varepsilon \in I$. However, this is contradiction because if $x^* < x^* + \varepsilon$, then $x^* + \varepsilon \in I$ contradicts x^* being an upper bound of I.

Similarly, suppose for the sake of contradiction there is a standard $\varepsilon > 0$ such that $x^* - x \geq \varepsilon$. Then,

$$x \leq x^* - \varepsilon < x^*.$$

Hence, $\forall^{\mathsf{s}} t \in I, t \leq x \leq x^* - \varepsilon \implies \forall^{\mathsf{s}} t \in I, t \leq x^* - \varepsilon$. Transfer applies and furnishes $\forall t \in I, t \leq x^* - \varepsilon$. Therefore $x^* - \varepsilon$ was an upper bound for I, contradicting x^* being the least upper bound.

Our assumptions must have been false so we conclude $\forall^{\mathsf{s}} \varepsilon > 0, |x - x^*| < \varepsilon$. That is, $x \simeq x^*$. □

Lemma 19. *The property of being infinitely close is transitive. That is,*

$$x \simeq y \wedge x \simeq z \implies y \simeq z.$$

Proof. Suppose $x \simeq y$ and $x \simeq z$. Hence, $\forall^{\mathsf{s}} \varepsilon > 0, |x - y| < \varepsilon$ and $\forall^{\mathsf{s}} \varepsilon > 0, |x - z| < \varepsilon$. For a given standard $\varepsilon > 0$, $\varepsilon/2$ is standard by Dual Transfer. Therefore by the triangle inequality,

$$|y - z| = |y - x + x - z| \leq |y - x| + |x - z| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since we could have picked any standard $\varepsilon > 0$, we conclude $\forall^{\mathsf{s}} \varepsilon > 0, |y - z| < \varepsilon$. That is, $y \simeq z$. □

Lemma 20. *If x and y are standard, then $x \simeq y \implies x = y$.*

Proof. Suppose x and y are standard and $x \simeq y$. This implies $x - y$ is infinitesimal. Since $x - y$ is standard by Dual Transfer, we can conclude $x - y = 0$ since Theorem 8 states 0 is the only standard infinitesimal. Therefore, $x = y$. □

Theorem 21. *If the standard part exists, it is unique.*

Proof. Suppose there are two numbers y and z such that $\mathrm{st}(x) = y$ and $\mathrm{st}(x) = z$. By Theorem 18, $x \simeq y$ and $x \simeq z$. Therefore, by Lemma 19, $y \simeq z$. Since y and z are standard by Theorem 17, Lemma 20 states $y = z$. □

Lemma 22. *If x is limited and $x \simeq y$, then y is limited.*

Proof. If x is limited then $\exists^{\mathsf{s}} M > 0, |x| \leq M$. Further, if $x \simeq y$, then $|y - x| < 1$. Therefore,

$$|y| = |y - x + x| \leq |y - x| + |x| < 1 + M.$$

Since $M + 1$ is standard by Dual Transfer, we conclude y is limited. □

Lemma 23. *If $x, y, a, b \in \mathbb{R}$ and $x \simeq a \wedge y \simeq b$, then:*

(i) $x + y \simeq a + b$

(ii) *If x and y are limited, $xy \simeq ab$.*

Proof. First we prove (i). Given $x, y, a, b \in \mathbb{R}$, suppose $x \simeq a$ and $y \simeq b$. Given a standard $\varepsilon > 0$ so that $\varepsilon/2$ is standard, we have by the triangle inequality that

$$|(x + y) - (a + b)| = |x - a + y - b| \leq |x - a| + |y - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence,

$$\forall^{\mathsf{s}} \varepsilon > 0, |(x + y) - (a + b)| < \varepsilon.$$

Therefore we conclude $x + y \simeq a + b$.

Next, we prove (ii).

$$\begin{aligned} |xy - ab| &= |xy - (a + x - x)b| \\ &= |x(y - b) + (x - a)b| \\ &\leq |x(y - b)| + |(x - a)b| \\ &= |x||y - b| + |x - a||b| \end{aligned}$$

Since x is limited, it is bounded above by a standard $A > 0$. Since y is limited and $y \simeq b$, b is limited by Lemma 22. Since b is limited, it is bounded above by a standard $B > 0$. Given a standard $\varepsilon > 0$, we know from Dual Transfer that $\varepsilon/(A + B)$ is standard and is therefore strictly greater than $|x - a|$ and $|y - b|$ since $x \simeq a$ and $y \simeq b$. Hence,

$$\begin{aligned}|xy - ab| &\leq |x||y - b| + |x - a||b| \\ &< A\frac{\varepsilon}{A + B} + \frac{\varepsilon}{A + B}B \\ &= \varepsilon\end{aligned}$$

Therefore, we have

$$\forall^{\mathsf{s}}\varepsilon > 0, |xy - ab| < \varepsilon.$$

In other words, $xy \simeq ab$. □

Theorem 24. *If $x, y \in \mathbb{R}$ are limited:*

(i) $x \geq 0 \implies x^* \geq 0$

(ii) $(x + y)^* = x^* + y^*$

(iii) $(x \cdot y)^* = x^* \cdot y^*$

(iv) $x \geq y \implies x^* \geq y^*$

Proof. We begin with (i). We will instead prove the contrapositive, $x^* < 0 \implies x < 0$. Suppose $x^* < 0$. By Theorem 18, $\forall^{\mathsf{s}}\varepsilon > 0, |x - x^*| < \varepsilon$. Since x^* is standard by Theorem 17, $-x^* > 0$ is standard by Dual Transfer. Therefore,

$$x - x^* \leq |x - x^*| < -x^*.$$

Finally, $x - x^* < -x^* \implies x < 0$.

Next, we prove (ii). We know that $x + y \simeq (x + y)^*$ by Theorem 18. Likewise, $x \simeq x^*$ and $y \simeq y^*$. Lemma 23 states $x + y \simeq x^* + y^*$. Therefore we know from Lemma 19 that $(x+y)^* \simeq x^* + y^*$. Since $(x+y)^*$ is standard by Theorem 17 and $x^* + y^*$ is standard by Dual Transfer, $(x+y)^* = x^* + y^*$ by Lemma 20.

Now we prove (iii). We know that $x \cdot y \simeq (x \cdot y)^*$ by Theorem 18. Likewise, $x \simeq x^*$ and $y \simeq y^*$. Since x and y are limited, Lemma 23 states $x \cdot y \simeq x^* \cdot y^*$. Therefore we know from Lemma 19 that $(x \cdot y)^* \simeq x^* \cdot y^*$. Since $(x \cdot y)^*$ is standard by Theorem 17 and $x^* \cdot y^*$ is standard by Dual Transfer, $(x \cdot y)^* = x^* \cdot y^*$ by Lemma 20.

Finally, we prove (iv). Suppose $x \geq y$. Then, $x - y \geq 0$. Since $x - y$ is limited (Theorem 25), by (i) we have $(x - y)^* \geq 0$. By (ii), $x^* + (-y)^* \geq 0$. By (iii), $x^* - y^* \geq 0$. Therefore, $x^* \geq y^*$. □

As an immediate application of the standard part, we may observe that instead of the nonstandard definition we gave for for a convergent sequence, $\exists^{\mathsf{s}} L \in \mathbb{R}, \forall^{\mathsf{u}} n \in \mathbb{N}, x_n \simeq L$, we could have just as easily defined $\exists L \in \mathbb{R}, \forall^{\mathsf{u}} n \in \mathbb{N}, \mathrm{st}(x_n) = L$ and had the proof proceed much the same.

3.1.3 Real Numbers

As we will often work with real numbers, it is useful to formalize rules for our most common nonstandard manipulations.

Theorem 25. *If $\varepsilon, \delta \in \mathbb{R}$ are infinitesimal, $\ell, k \in \mathbb{R}$ are limited, $u, v \in \mathbb{R}$ are unlimited, and $x \in \mathbb{R}$, then the following hold:*

(i) $\mathrm{st}(\varepsilon) = 0$.

(ii) $\varepsilon \neq 0 \implies 1/\varepsilon$ *is unlimited.*

(iii) $1/u$ *is infinitesimal.*

(iv) $x \not\simeq 0 \implies 1/x$ *is limited.*

(v) $\varepsilon + \delta$ *and* $\varepsilon \cdot \delta$ *are infinitesimal.*

(vi) $\ell + k$ *and* $\ell \cdot k$ *are limited.*

(vii) $u \cdot v$ *is unlimited and* $\mathrm{sgn}(u) = \mathrm{sgn}(v) \implies u + v$ *is unlimited.*

(viii) $u + \ell$ *is unlimited.*

(ix) $x \not\simeq 0 \implies x \cdot u$ *is unlimited.*

(x) $\ell \cdot \varepsilon$ *is infinitesimal.*

Proof. (i): The definition of infinitesimal implies $\varepsilon \simeq 0$. Since $\mathrm{st}(\varepsilon) \simeq \varepsilon$ by Theorem 18, Lemma 19 ensures $\mathrm{st}(\varepsilon) \simeq 0$. Since the standard part is standard by Theorem 17 and 0 is standard, Lemma 20 states $\mathrm{st}(\varepsilon) = 0$.

(ii): Suppose $\varepsilon \neq 0$. Since ε is infinitesimal, we know $\forall^{\mathsf{s}} y > 0, |\varepsilon| < y$. For a given $y > 0$, this implies $|1/\varepsilon| > 1/y$. Given a standard $a > 0$, pick $y = 1/a > 0$ so that $|1/\varepsilon| > a$. We conclude that $\forall^{\mathsf{s}} a > 0, |1/\varepsilon| > a$. This is exactly the statement $1/\varepsilon$ is unlimited.

(iii): Since u is unlimited, we know $\forall^{\mathsf{s}} y > 0, |u| > y$. For a given $y > 0$, this implies $|1/u| < 1/y$. Given a standard $a > 0$, pick $y = 1/a > 0$ so that $|1/u| < a$. We conclude that $\forall^{\mathsf{s}} a > 0, |1/u| < a$. This is exactly the statement $1/u$ is infinitesimal.

(iv): Suppose $x \not\simeq 0$. This is exactly the statement that x is not infinitesimal. Therefore, $\exists^{\mathsf{s}} y > 0, |x| \geq y$. Given a y, this implies $|1/x| \leq 1/y$. Since y is standard, $1/y > 0$ is standard by Dual Transfer. Let $M = 1/y$ so that $\exists^{\mathsf{s}} M > 0, |1/x| \leq M$. This is exactly the statement $1/x$ is limited.

(v): Since ε and δ are infinitesimal, $\varepsilon \simeq 0$ and $\delta \simeq 0$. Since ε and δ are limited, Lemma 23 states $\varepsilon + \delta \simeq 0 + 0$ and $\varepsilon \cdot \delta \simeq 0 \cdot 0$. These are exactly the statements that $\varepsilon + \delta$ and $\varepsilon \cdot \delta$ are infinitesimal.

(vi): Since ℓ and k are limited, $\exists^{\mathsf{s}} a > 0, |\ell| \leq a$ and $\exists^{\mathsf{s}} y > 0, |k| \leq y$. Therefore, $|\ell + k| \leq |\ell| + |k| \leq a + y$ and $|\ell \cdot k| = |\ell| \cdot |k| \leq a \cdot y$. Since a and y are standard, $M = a + y$ and $N = a \cdot y$ are standard by Dual Transfer. Therefore, $\exists^{\mathsf{s}} M > 0, |\ell + k| \leq M$ and $\exists^{\mathsf{s}} N > 0, |\ell \cdot k| \leq N$. These are exactly the statements that $\ell + k$ and $\ell \cdot k$ are limited.

(vii): Since u and v are unlimited, $\forall^{\mathsf{s}} a > 0, |u| > a$ and $\forall^{\mathsf{s}} y > 0, |v| > y$. Therefore given an a, y, $|u \cdot v| = |u| \cdot |v| > a \cdot y$ and, since $\mathrm{sgn}(u) = \mathrm{sgn}(v)$, $|u + v| = |u| + |v| > a + y$. Given standard $M > 0$, choose a and y such that $a = y = \sqrt{M}$ so that $a \cdot y = M$ is standard by Dual Transfer. Therefore, we may conclude $\forall^{\mathsf{s}} M > 0, |u \cdot v| > M$. That is, $u \cdot v$ is unlimited. Given a standard $N > 0$, choose a and y such that $a = y = M/2$ so that

$a + y = M$ is standard by Dual Transfer. Therefore, we may conclude $\forall^{\mathsf{s}} N > 0, |u + v| > N$. That is, $u + v$ is unlimited.

(viii): ℓ limited means $\exists^{\mathsf{s}} M > 0, |\ell| \leq M$ and u unlimited means $\forall^{\mathsf{s}} y > 0, |u| > y$. We have $|u| = |u + \ell - \ell| \leq |u + \ell| + |\ell| \leq |u + \ell| + M$. Hence, given a standard $y > 0$, $|u + \ell| \geq |u| - M > y - M$. Given a standard $a > 0$, choose $y = a + M$ which is standard by Dual Transfer. Then, $|u + \ell| > a$ so that we may conclude $\forall^{\mathsf{s}} a > 0, |u + \ell| > a$. This is exactly the statement that $u + \ell$ is unlimited.

(ix): Suppose $x \not\simeq 0$. This is exactly the statement that x is not infinitesimal. Therefore, $\exists^{\mathsf{s}} M > 0, |x| \geq M$. Since u is unlimited, $\forall^{\mathsf{s}} y > 0, |u| > y$. For a given $y > 0$, we have $|x \cdot u| = |x| \cdot |u| > M \cdot y$. Given a standard $a > 0$, choose $y = a/M$ which is standard by Dual Transfer. Then, we have $|x \cdot u| > a$ so that we may conclude $\forall^{\mathsf{s}} a > 0, |x \cdot u| > a$. This is exactly the statement that $x \cdot u$ is unlimited.

(x): ℓ limited means $\exists^{\mathsf{s}} M > 0, |\ell| \leq M$ and ε infinitesimal means $\forall^{\mathsf{s}} y > 0, |\varepsilon| < y$. For a given $y > 0$, we have $|\ell \cdot \varepsilon| = |\ell| \cdot |\varepsilon| < M \cdot y$. Given a standard $x > 0$, choose $y = x/M$ which is standard by Dual Transfer. Then, we have $|\ell \cdot \varepsilon| < x$ so that we may conclude $\forall^{\mathsf{s}} x > 0, |\ell \cdot \varepsilon| < x$. This is exactly the statement that $\ell \cdot \varepsilon$ is infinitesimal. □

Using these rules and our previously defined NSA tools, we can already begin to give NSA proofs that formalize our intuition.

Example 1. The sequence defined by $x_n = 1/n$ converges to 0.

Proof. For any unlimited $\nu \in \mathbb{N}$, $1/\nu$ is infinitesimal so that $\mathrm{st}(1/\nu) = 0$ by Theorem 25. Therefore, $\{x_n\}$ converges to 0. □

3.2 Core Analysis Topics

3.2.1 Continuity

The modern definition of function continuity is one that taken many forms and a number of different mathematicians to develop and has culminated in the following definition for function continuity.

Definition 10 (Continuity)**.** If $S \subseteq \mathbb{R}$ and $f\colon S \to \mathbb{R}$ is a function, then we say f is continuous at $c \in S$ if

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in S, |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon.$$

Definition 11 (Continuous Function)**.** If $S \subseteq \mathbb{R}$ and $f\colon S \to \mathbb{R}$ is a function, then we say f is a continuous function if f is continuous at all $c \in S$.

Definition 10 is not the definition Cauchy gave in his celebrated *Cours d'Analyse de l'École Royale Polytechnique.*[1] Instead, he considered a function to be continuous at a point when an infinitesimal change in the domain around that point produced an infinitesimal change in the range. In fact, we can give this definition rigor using NSA.

Definition 12 (S-Continuity)**.** If $S \subseteq \mathbb{R}$ and $f\colon S \to \mathbb{R}$ is a function, then we say f is S-continuous at $c \in S$ if

$$\forall x \in S, x \simeq c \implies f(x) \simeq f(c).$$

S-continuity is not quite equivalent to continuity, but it encapsulates the notion of continuity we care about in the following sense.

Theorem 26. *If $S \subseteq \mathbb{R}$ is standard and $f\colon S \to \mathbb{R}$ is a standard function, then the following are equivalent for a standard $c \in S$:*

(i) f is continuous at c.

(ii) f is S-continuous at c.

Proof. First, we prove (i) $\implies$ (ii). Suppose f is continuous at c so that

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in S, |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon.$$

The formula holds for all ε so we can take in particular a standard ε to find

$$\forall^{\mathsf{s}} \varepsilon > 0, \exists \delta > 0, [\forall x \in S, |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon].$$

The bracketed expression is a classical formula of δ with standard parameters, so we may apply Dual Transfer and find

$$\forall^{\mathsf{s}}\varepsilon > 0, \exists^{\mathsf{s}}\delta > 0, \forall x \in S, |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon.$$

In the case that we had $x \simeq c$, the δ requirement is automatically satisfied so that the expression may be rewritten as

$$\forall^{\mathsf{s}}\varepsilon > 0, \forall x \in S, x \simeq c \implies |f(x) - f(c)| < \varepsilon.$$

This may be written with the definition of infinitely close so that we may conclude

$$\forall x \in S, x \simeq c \implies f(x) \simeq f(c).$$

This shows (i) $\implies$ (ii).

Next, we prove that (ii) $\implies$ (i). Suppose $\forall x \in S, x \simeq c \implies f(x) \simeq f(c)$. Then, we know that

$$\forall^{\mathsf{s}}\varepsilon > 0, [\exists\delta > 0, \forall x \in S, |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon].$$

This is true because we can take an infinitesimal δ to have the expression hold automatically. The bracketed expression is a classical formula of ε with standard parameters, so we may apply Transfer to show

$$\forall\varepsilon > 0, \exists\delta > 0, \forall x \in S, |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon.$$

This is exactly the definition of continuity, so (ii) $\implies$ (i). □

Although we defined continuous functions using the classical definition of continuity, we can give an equally good definition with S-continuity.

Theorem 27. *If $S \subseteq \mathbb{R}$ is standard and $f\colon S \to \mathbb{R}$ is a standard function, then the following are equivalent:*

(i) f is a continuous function.

(ii) $\forall^{\mathsf{s}}x \in S$, f is S-continuous at x.

Proof. Consider the sets $\{x \in S : f \text{ is continuous at } x\}$ and ${}^{\mathsf{s}}\{x \in S : f \text{ is S-continuous at } x\}$. Because f is S-continuous at standard x exactly when it is continuous by Theorem 26, by Standardization these sets must have the same standard elements. Further, since the sets are standard they are equivalent by Theorem 9 as they have the same standard elements. In the case that f is a continuous function, the sets are exactly S. In this case, the standard elements of ${}^{\mathsf{s}}\{x \in S : f \text{ is S-continuous at } x\}$ are equivalent to the standard elements of the first set exactly when $\forall^{\mathsf{s}} x \in S$, f is S-continuous at x. □

One of the nice properties of S-continuity is its definition is very easy to manipulate; this allows simple proofs of many properties.

Theorem 28. *Suppose $A, B \subseteq \mathbb{R}$ such that $f\colon A \to \mathbb{R}$ and $g\colon B \to \mathbb{R}$ are functions and $S \subseteq A \cap B$. If f and g are S-continuous at $c \in S$, then the function $h\colon S \to \mathbb{R}$ is S-continuous at c for each of the following definitions:*

(i) $h(x) = f(x) + g(x)$.

(ii) $h(x) = f(x) - g(x)$.

(iii) *If $f(c)$ and $g(c)$ are limited,* $h(x) = f(x)g(x)$.

(iv) *If $f(c)$ is limited, $g(c) \not\simeq 0$, and $\forall x \in S, g(x) \neq 0$,* $h(x) = f(x)/g(x)$.

Proof. (i): Suppose we have an $x \in S$ such that $x \simeq c$. Since f and g are both S-continuous at c, $f(x) \simeq f(c)$ and $g(x) \simeq g(c)$. By Lemma 23, $f(x) + g(x) \simeq f(c) + g(c)$. Hence, h is S-continuous at c.

(ii): Suppose we have an $x \in S$ such that $x \simeq c$. Since f and g are both S-continuous at c, $f(x) \simeq f(c)$ and $g(x) \simeq g(c) \implies -g(x) \simeq -g(c)$. By Lemma 23, $f(x) - g(x) \simeq f(c) - g(c)$. Hence, h is S-continuous at c.

(iii): Suppose we have an $x \in S$ such that $x \simeq c$. Since f and g are both S-continuous at c, $f(x) \simeq f(c)$ and $g(x) \simeq g(c)$. Since $f(c)$ and $g(c)$ are limited, by Lemma 23, $f(x)g(x) \simeq f(c)g(c)$. Hence, h is S-continuous at c.

(iv): Suppose we have an $x \in S$ such that $x \simeq c$. Since f and g are both S-continuous at c, $f(x) \simeq f(c)$ and $g(x) \simeq g(c)$. Since $g(c) \not\simeq 0 \implies$

$1/g(c)$ is limited (by Theorem 25), we may use Lemma 23 to rewrite this as $g(x)/g(c) \simeq 1$. We know that

$$|g(c)| = |g(c) + g(x) - g(x)| \leq |g(x) - g(c)| + |g(x)|.$$

Further, since $g(c) \not\simeq 0$, $\exists^{\mathsf{s}} y > 0, |g(c)| \geq y$. Take this y. Additionally, $g(x) \simeq g(c) \implies \forall^{\mathsf{s}} b > 0, |g(x) - g(c)| < b$. In particular, we may choose $b = y/2$. Then,

$$\begin{aligned} |g(x)| &\geq |g(c)| - |g(x) - g(c)| \\ &\geq y - |g(x) - g(c)| \\ &> y - y/2 \\ &= y/2 \end{aligned}$$

Since $a = y/2$ is standard by Dual Transfer, $\exists^{\mathsf{s}} a > 0, |g(x)| \geq a$. That is, $g(x) \not\simeq 0$. In this case $1/g(x)$ is limited (by Theorem 25) and we may use Lemma 23 to rewrite $g(x)/g(c) \simeq 1$ as $1/g(x) \simeq 1/g(c)$. Since $f(c)$ and $1/g(c)$ are limited, we can now use Lemma 23 to find that $f(x)/g(x) \simeq f(c)/g(c)$. Hence, h is S-continuous at c. □

Another important concept is the notion of uniform continuity. Classically, we define uniform continuity as follows.

Definition 13 (Uniform Continuity)**.** If $S \subseteq \mathbb{R}$ and $f \colon S \to \mathbb{R}$ is a function, then we say f is uniformly continuous if

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in S, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Just like with continuity, we can define an NSA counterpart to uniform continuity.

Definition 14 (S-Uniform Continuity)**.** If $S \subseteq \mathbb{R}$ and $f \colon S \to \mathbb{R}$ is a function, then we say f is S-uniformly continuous if

$$\forall x \in S,\ f \text{ is S-continuous at } x.$$

Likewise, we can show S-uniform continuity characterizes uniform continuity in the following sense.

Theorem 29. *If $S \subseteq \mathbb{R}$ is standard and $f\colon S \to \mathbb{R}$ is a standard function, then the following are equivalent:*

(i) f is uniformly continuous.

(ii) f is S-uniformly continuous.

Proof. First, we show (i) $\implies$ (ii). Suppose (i) so that

$$\forall \varepsilon > 0, \exists \delta > 0, [\forall x, y \in S, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon].$$

The formula is true for all ε as it is true for all standard ε. In this case the bracketed formula is a classical formula of δ with standard parameters so Dual Transfer applies and shows that

$$\forall^{\mathsf{s}} \varepsilon > 0, \exists^{\mathsf{s}} \delta > 0, \forall x, y \in S, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

In the case that $x \simeq y$, the δ condition is automatically satisfied and we may write

$$\forall^{\mathsf{s}} \varepsilon > 0, \forall x, y \in S, x \simeq y \implies |f(x) - f(y)| < \varepsilon.$$

Using the definition of infinitely close, this may be rewritten

$$\forall x, y \in S, x \simeq y \implies f(x) \simeq f(y).$$

Hence, (i) $\implies$ (ii).

Next we show that (ii) $\implies$ (i). Suppose (ii). In this case we know that

$$\forall^{\mathsf{s}} \varepsilon > 0, [\exists \delta > 0, \forall x, y \in S, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon].$$

This is true because we can take an infinitesimal δ and the expression will hold automatically by our assumption. The bracketed expression is a classical formula of ε with standard parameters, so Transfer applies and furnishes

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in S, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Therefore, (ii) $\implies$ (i). ☐

One of the advantages of the NSA approach to continuity is it can be easier to understand than the classical definition.

Example 2. The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is a continuous function, but not a uniformly continuous function.

Proof. Since f is standard, to show it is a continuous function we need to show it is S-continuous at every standard x. We want to consider the y such that $y \simeq x$. Since this means $x - y$ is infinitesimal, every such y can be represented as $x + \varepsilon$ where ε is an infinitesimal. Therefore, given x and $x + \varepsilon$, we need to show that $f(x) \simeq f(x + \varepsilon)$.

$$\begin{aligned} f(x) &= x^2 \\ f(x+\varepsilon) &= (x+\varepsilon)^2 \\ &= x^2 + 2x\varepsilon + \varepsilon^2 \end{aligned}$$

If $\operatorname{st}\big(f(x)\big) = \operatorname{st}\big(f(x+\varepsilon)\big)$, we can conclude by Lemma 19 that $f(x) \simeq f(x+\varepsilon)$. Applying Theorem 25 for real numbers and Theorem 24 for the standard part, we find for $f(x)$

$$\operatorname{st}\big(f(x)\big) = \operatorname{st}(x^2) = x^2.$$

For $f(x + \varepsilon)$, we have

$$\begin{aligned} \operatorname{st}\big(f(x+\varepsilon)\big) &= \operatorname{st}(x^2 + 2x\varepsilon + \varepsilon^2) \\ &= \operatorname{st}(x^2) + \operatorname{st}(2x\varepsilon) + \operatorname{st}(\varepsilon^2) \\ &= x^2. \end{aligned}$$

Therefore, $f(x) \simeq f(x+\varepsilon)$. Hence, f is a continuous function.

Since f is standard, we can show it's not S-uniformly continuous to show it's not uniformly continuous. To do this we need to show $\exists x, y \in \mathbb{R}, x \simeq y \not\Longrightarrow f(x) \simeq f(y)$. We can pick $x = \nu$ and $y = \nu + 1/\nu$ for some positive unlimited ν. We know that $x \simeq y$, but $f(\nu) = \nu^2$ and $f(\nu{+}1/\nu) = \nu^2{+}2{+}1/\nu^2$. Hence, $|f(x){-}f(y)| = 2{+}1/\nu^2 \implies f(x) \not\simeq f(y)$. Therefore, f is not uniformly continuous. □

Example 3. The function $f\colon (0,1) \to \mathbb{R}$ defined by $f(x) = 1/x$ is not uniformly continuous.

Proof. Since f is standard, we can show it's not S-uniformly continuous to show it's not uniformly continuous. We need to show $\exists x, y \in (0,1), x \simeq y \not\Longrightarrow f(x) \simeq f(y)$. We can pick $x = 1/\nu$ and $y = 1/(\nu+1)$ for some positive unlimited ν. As these are both infinitesimal by Theorem 25, $x \simeq y$. $f(1/\nu) = \nu$ and $f(1/(\nu+1)) = \nu + 1$. However, $|f(x) - f(y)| = 1 \implies f(x) \not\simeq f(y)$. Therefore, f is not uniformly continuous. □

These examples illustrate bringing rigorous intuition to distinguishing between continuous and uniformly continuous functions. S-continuous functions are required to meet the S-continuity condition infinitely close to every standard element in its domain; on the other hand, S-uniformly continuous functions are required to meet the S-continuity condition at every element in its domain. This suggests that to distinguish continuity and uniform continuity we should look at the elements of the domain not infinitely close to a standard element in the domain. Indeed, in Example 2 we found a contradiction in unlimited elements of $\mathbb{R}$ and in Example 3 we found a contradiction in positive infinitesimal elements of $(0,1)$.

3.2.2 Differentiability

It is worth noting how quickly our tools can be extended. Just as with continuity, there is an NSA analog to the notion of a derivative.

Definition 15 (S-Differentiable)**.** Let $I \subseteq \mathbb{R}$ be an interval and $f\colon I \to \mathbb{R}$ be a function with $c \in I$. We say f is S-differentiable at c if we have

$$\exists^{\mathsf{s}} L \in \mathbb{R}, \forall x \in I, x \simeq c \implies \frac{f(x) - f(c)}{x - c} \simeq L.$$

Further, S-differentiability shares characteristics with differentiability in sense that S-continuity shares characteristics with continuity.

Proposition 30. *Let I be a standard interval and $f\colon I \to \mathbb{R}$ be a standard function with a standard $c \in I$. Then, the following are equivalent:*

(i) f *is differentiable at* c.

(ii) f *is S-differentiable at* c.

Although we will not dive into the topic of differentiability, we can immediately see the parallels with continuity that allow the topic to be attacked with NSA.

3.3 Selected Theorems and Topics

3.3.1 Topology

Basic topological properties can be restated with NSA similarly to how continuity and differentiability can be stated. These versions are often useful when proving things in the NSA framework.

Definition 16 (Open)**.** $E \subseteq \mathbb{R}$ is open $\iff$

$$\forall y \in E, \exists \varepsilon > 0, \forall x \in \mathbb{R}, |x - y| < \varepsilon \implies x \in E.$$

Definition 17 (S-Open)**.** $E \subseteq \mathbb{R}$ is S-open $\iff$

$$\forall^{\mathsf{ns}} x \in \mathbb{R}, x^* \in E \implies x \in E.$$

Lemma 31. *E is S-open* $\iff$

$$\forall^{\mathsf{s}} y \in E, \forall x \in \mathbb{R}, x \simeq y \implies x \in E$$

Proof. Suppose E is S-open so that $\forall^{\mathsf{ns}} x \in \mathbb{R}, x^* \in E \implies x \in E$. The near standard $x \in \mathbb{R}$ with $x^* \in E$ are exactly the $x \in \mathbb{R}$ with a standard $y \simeq x$ satisfying $y \in E$ so that the requirement may be rewritten $\forall^{\mathsf{s}} y \in E, \forall x \in \mathbb{R}, x \simeq y \implies x \in E$. □

Theorem 32. *Given a standard* $E \subseteq \mathbb{R}$*, the following are equivalent:*

(i) E *is open.*

(ii) E *is S-open.*

Proof. We will use Lemma 31's variation of S-open. We first show (i) $\implies$ (ii). Suppose E is open, then we know

$$\forall y \in E, \exists \varepsilon > 0, [\forall x \in \mathbb{R}, |x - y| < \varepsilon \implies x \in E].$$

The formula is true for all y so it is true for all standard y. In this case, as E is standard, the bracketed formula is a classical formula of ε with standard parameters. Therefore, Dual Transfer applies and furnishes

$$\forall^{\mathsf{s}} y \in E, \exists^{\mathsf{s}} \varepsilon > 0, \forall x \in \mathbb{R}, |x - y| < \varepsilon \implies x \in E.$$

In the case that $x \simeq y$, the ε condition is automatically satisfied and we have

$$\forall^{\mathsf{s}} y \in E, \forall x \in \mathbb{R}, x \simeq y \implies x \in E.$$

Hence, (i) $\implies$ (ii).

Now we show that (ii) $\implies$ (i). Suppose (ii). In this case we know

$$\forall^{\mathsf{s}} y \in E, [\exists \varepsilon > 0, \forall x \in \mathbb{R}, |x - y| < \varepsilon \implies x \in E].$$

This is true because we can pick any infinitesimal ε and our assumption ensures the expression will hold. The bracketed expression is now a classical formula of y with standard parameters, so Transfer applies and provides

$$\forall y \in E, \exists \varepsilon > 0, \forall x \in \mathbb{R}, |x - y| < \varepsilon \implies x \in E.$$

Hence, (ii) $\implies$ (i). □

Definition 18 (Closed)**.** $E \subseteq \mathbb{R}$ is closed $\iff$ E^{c} is open.

Definition 19 (S-Closed)**.** $E \subseteq \mathbb{R}$ is S-closed $\iff$

$$\forall^{\mathsf{ns}} x \in \mathbb{R}, x \in E \implies x^* \in E.$$

Theorem 33. *Given a standard $E \subseteq \mathbb{R}$, the following are equivalent:*

(i) E is closed.

(ii) *E is S-closed.*

Proof. E is closed exactly when E^{c} is open. Since E is standard, E^{c} is standard by Dual Transfer. By Theorem 32, E^{c} is open exactly when it is S-open. In this case,

$$\forall^{\mathsf{ns}} x \in \mathbb{R}, x^* \notin E \implies x \notin E.$$

Taking the contrapositive, we find

$$\forall^{\mathsf{ns}} x \in \mathbb{R}, x \in E \implies x^* \in E.$$

This is exactly the definition of S-closed. □

Definition 20 (Compact)**.** E is compact $\iff$ every open cover of E has a finite subcover.

Definition 21 (S-Compact)**.** E is S-compact $\iff$

$$\forall x \in E, x \in E \implies x^* \in E.$$

Theorem 34. *Given a standard $E \subseteq \mathbb{R}$, the following are equivalent:*

(i) *E is compact.*

(ii) *E is S-compact.*

Proof. First we show (i) $\implies$ (ii). Assume every open cover of E has a finite subcover. Assume for the sake of contradiction $\exists x \in E, x \in E \wedge x^* \notin E$. This implies that $\forall^{\mathsf{s}} y \in E, y \not\simeq x$. Let

$$C = {}^{\mathsf{s}}\{U \in \mathscr{P}(E) : U \text{ is open} \wedge x \notin U\}.$$

Let $\mathcal{U} = \bigcup_{U \in C} U$; this is standard by Dual Transfer. Either C is an open cover for E or it's not.

Suppose C is an open cover for E. Since C is standard, it has a standard finite subcover C_F by Dual Transfer. Since C_F is standard and finite, it contains only standard elements by Theorem 11. By Standardization, the standard elements of C are exactly the open U such that $x \notin U$. Hence, $x \in E$ is not in any element of our finite subcover. This contradicts C_F being a finite subcover for C.

Suppose C is not an open cover for E. This means that $E \nsubseteq \mathcal{U}$. Therefore, $E \setminus \mathcal{U}$ is nonempty. Consider the statement $\exists y \in E, y \in E \setminus \mathcal{U}$. Since E and and $\mathcal{U}$ are standard, Dual Transfer applies to this statement and states $\exists^{\mathsf{s}} y \in E, y \in E \setminus \mathcal{U}$. If for every standard $y \in E$ there is standard open set containing y but not x, this open set must have been in C by Standardization. Therefore we would have $y \in \mathcal{U}$, a contradiction.

Lemma 35. *If $x \not\simeq y$ for a standard $y \in E$, there is a standard open set containing y but not x.*

Suppose not, then every standard open set that contains y also contains x. Given a standard $\varepsilon > 0$, $\{a \in E : |a - y| < \varepsilon\}$ is a standard open set so it contains x. Hence, $\forall^{\mathsf{s}} \varepsilon > 0, |x - y| < \varepsilon$. This is exactly the statement that $x \simeq y$. However, we assumed that $x \not\simeq y$, so this is a contradiction.

Since by assumption $\forall^{\mathsf{s}} y \in E, y \not\simeq x$, using Lemma 35 the case where C is not an open cover for E is a contradiction. Therefore, we have a contradiction in all cases and we conclude $\forall x \in E, x \in E \implies x^* \in E$. Hence, (i) $\implies$ (ii).

Next we show that (ii) $\implies$ (i). Assume E is S-compact and that $C = \{U_\alpha : \alpha \in A\}$ is a standard open cover for E so that $E \subseteq \mathcal{U} = \bigcup_{\alpha \in A} U_\alpha$.

Suppose $x \in E$. Since E is S-compact, this implies $x^* \in E$. Therefore, $x^* \in \mathcal{U}$. This means that

$$\exists U_\alpha \in C, x^* \in U_\alpha.$$

Since C is standard, this is a standard formula of U_α with standard parameters; Dual Transfer applies and furnishes

$$\exists^{\mathsf{s}} U_\alpha \in C, x^* \in U_\alpha.$$

Since this U_α is open and standard, it must also be S-open by Theorem 32. Hence, $x^* \in U_\alpha \implies x \in U_\alpha$. Therefore, any $x \in E$ is contained in a standard U_α. By Corollary 5, we can let $C_F \subseteq C$ be a finite subset of C

containing all its standard elements. Since every $x \in E$ is contained in a standard element of C, C_F is a finite subcover for E. Therefore, every standard open cover of E has a finite subcover. Since this is a classical formula of the open cover with standard parameters, Transfer applies to the formula and tells us every open cover of E has a finite subcover. Therefore, (ii) $\implies$ (i). □

With our NSA notions of topology some classical theorems become immediately obvious.

Theorem 36. *Given $S \subseteq \mathbb{R}$ such that $f: S \to \mathbb{R}$ is a continuous function and $K \subseteq S$ is compact, the restriction $f|_K$ is uniformly continuous.*

Proof. By Transfer, if the theorem is true for every standard f and K, the theorem holds for all f and K. Therefore, assume f and K are standard. Since the restriction of a continuous function is continuous (by a classical result), we know $f|_K$ is continuous. Since it is standard, it is S-continuous by Theorem 27. Hence,

$$\forall^{\mathsf{s}} c \in K, \forall x \in K, x \simeq c \implies f(x) \simeq f(c).$$

In particular, since K is compact and standard it is S-compact by Theorem 34. Given an arbitrary $y \in K$, S-compactness gives $y^* \in K$ so that by above we may write

$$\forall x \in K, x \simeq y^* \implies f(x) \simeq f(y^*).$$

Since f is S-continuous at the standard point y^*, $y \simeq y^* \implies f(y) \simeq f(y^*)$. By the transitivity of infinite closeness (Lemma 19), we have

$$\forall y \in K, \forall x \in K, x \simeq y \implies f(x) \simeq f(y).$$

This is exactly the statement that $f|_K$ is S-uniformly continuous. Since $f|_K$ is standard, $f|_K$ is uniformly continuous by Theorem 29. □

3.3.2 Extreme Value Theorem

The Extreme Value Theorem is a classically difficult problem that admits a relatively straightforward proof with NSA. The NSA proof is particularly attractive because it gives rigor to the intuitive notion of why it should be true. Intuitively, we can think about the problem as dividing up a compact interval and choosing the interval with the maximal function value. With the help of IST, we can actually do this.

Theorem 37 (Extreme Value Theorem)**.** *Given $S \subseteq \mathbb{R}$, let $f\colon S \to \mathbb{R}$ be a continuous function. For any compact $K \subseteq S$, $f|_K$ obtains a maximum and minimum.*

Proof. By Transfer, the theorem is true for all f and K as soon as it is true for all standard f and K; we will prove this case. WLOG, we will only prove f has a maximum.

By Corollary 5, we can let $F \subseteq K$ be a finite subset of K containing all its standard elements. Since F is finite, let $a \in F$ be the point where $f|_F$ has its maximum. Then,

$$\forall x \in F, f(x) \leq f(a).$$

Since F contains all the standard elements of K, we know

$$\forall^{\mathsf{s}} x \in K, f(x) \leq f(a).$$

Now we take the standard part of both sides; this preserves the inequality by Theorem 24. Since f is standard, $f(x)$ is standard when x is. We now have

$$\forall^{\mathsf{s}} x \in K, f(x) \leq f(a)^*.$$

Since K is compact and standard, it is S-compact. Therefore, $a \in K \implies a^* \in K$. Further, since f is standard and continuous, it is S-continuous. Therefore, $a \simeq a^* \implies f(a) \simeq f(a^*)$. Since a^* is standard, $f(a^*)$ is standard by Dual Transfer. Therefore, $f(a^*)$ must have been the standard part of $f(a)$: $f(a^*) = f(a)^*$. Therefore,

$$\forall^{\mathsf{s}} x \in K, f(x) \leq f(a^*).$$

This is a standard formula of x with standard parameters, so Transfer applies and furnishes

$$\forall x \in K, f(x) \leq f(a^*).$$

Since $a^* \in K$, this is exactly the statement that $f|_K$ obtains a maximum at a^*. □

3.3.3 Intermediate Value Theorem

The Intermediate Value Theorem has been proved many ways, but there is a clever NSA proof that again making use of the fact that any set has a finite subset containing its standard elements.

Theorem 38 (Intermediate Value Theorem)**.** *If $f\colon [a,b] \to \mathbb{R}$ is a continuous function and $u \in \mathbb{R}$ satisfies $f(a) < u < f(b)$ or $f(b) < u < f(a)$, then there is a $c \in (a,b)$ such that $f(c) = u$.*

Proof. By Transfer, the theorem is true for all f and u as soon as it is true for all standard f and u defined on a standard $[a,b]$; we will prove this case. WLOG, suppose $f(a) < u < f(b)$.

By Corollary 5, there is a finite set F containing all the standard elements of $[a,b]$. Let $x = \min\{x' \in F : u \leq f(x')\}$. This certainly exists because the set was nonempty; b is standard so $b \in F$ and $u < f(b)$ by assumption. Further, define $y = \max\{y' \in F : y' < x\}$. This set is nonempty since it contains a; x could not have been a since x satisfies $u \leq f(x)$ and $f(a) < u$ by assumption. Further, we know $f(y) < u$ because $f(y) \geq u$ would contradict x being the minimum element with this property.

Note x and y are limited since they are bounded above by the standard b; their standard parts exist. Suppose for the sake of contradiction that $x \not\simeq y$. Therefore, $x - y$ is not infinitesimal and we may define $d = \mathrm{st}(x-y) > 0$. Since $y \simeq y^*$, $\forall^{\mathsf{s}} \varepsilon > 0, |y - y^*| < \varepsilon$. Because $d/2$ is standard (by Dual Transfer), we can let $\varepsilon = d/2$ so that

$$y - y^* \leq |y - y^*| < d/2 \implies y < y^* + d/2.$$

Because $x - y \simeq d$, we know for all standard ε,

$$\begin{aligned} -x + y + d \leq |x - y - d| < \varepsilon \\ \implies y + d - \varepsilon < x. \end{aligned}$$

Adding $y^* - y$ to both sides and using $\forall^{\mathsf{s}}\delta > 0, |y^* - y| < \delta$ we find,

$$\begin{aligned} y^* + d - \varepsilon &< x + (y^* - y) \\ &\leq x + \delta \\ \implies y^* + d - \varepsilon - \delta &< x. \end{aligned}$$

Since $d/4$ is standard (by Dual Transfer), we may let $\varepsilon = \delta = d/4$ so that

$$y < y^* + d/2 < x.$$

Since $y^* + d/2$ is standard by Dual Transfer, $y^* + d/2 \in F$. This is a contradiction since y was chosen to be the largest $y \in F$ satisfying $y < x$. Therefore our assumption was false and we conclude $x \simeq y$.

Because $x^* \simeq x \simeq y$, and x^* is standard, we know f is S-continuous at x^*. Therefore,

$$u \leq f(x) \simeq f(x^*) \simeq f(y) < u.$$

Adding $f(x^*) - f(x)$ (an infinitesimal) to $u \leq f(x)$, we get

$$u + (f(x^*) - f(x)) \leq f(x^*).$$

Noting that the standard part of $f(x^*)$ is itself since it is standard (by Dual Transfer), we can take the standard part of both sides and preserve the inequality by Theorem 24. Therefore, $u \leq f(x^*)$. Similarly for $f(y) < u$ and adding $f(x^*) - f(y)$ we get

$$f(x^*) < u + (f(x^*) - f(y)).$$

Taking the standard part of each side we get $f(x^*) \leq u$. Finally, we have

$$u \leq f(x^*) \leq u.$$

Hence, $f(x^*) = u$. Therefore, we may pick $c = x^*$. □

It's worth note the majority of this proof was spent making the arguments clear. The crux of the proof is simply that $[a, b]$ has a finite subset F containing its standard elements and defining $x = \min\{x \in F : u \leq f(x)\}$ lets us find our $c = x^*$ provided f was continuous.

Bibliography

[1] Augustin-Louis Cauchy. *Cours d'Analyse de l'École Royale Polytechnique. Analyse Algébrique*. 1. de l'Imprimerie Royale, 1821.

[2] Karel Hrbacek. "Axiom of Choice in nonstandard set theory." In: *Journal of Logic and Analysis* 4.8 (2012). DOI: `10.4115/jla.2012.4.8`.

[3] Jiří Lebl. *Basic Analysis: Introduction to Real Analysis*. 2016. ISBN: 9781530256747.

[4] Edward Nelson. "Internal Set Theory: A New Approach to Nonstandard Analysis." In: *Bulletin of the American Mathematical Society* 83.6 (Nov. 1977), pp. 1165–1198. DOI: `10.1090/S0002-9904-1977-14398-X`.

[5] Alain M. Robert. *Nonstandard Analysis*. Dover Books on Mathematics Series. Dover Publications, 2003. ISBN: 9780486432793.

[6] Abraham Robinson. *Non-standard Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1974. ISBN: 9780691044903.