First Order Logic and Nonstandard Analysis

Julian Hartman

September 4, 2010

Abstract

This paper is intended as an exploration of nonstandard analysis, and the rigorous use of infinitesimals and infinite elements to explore properties of the real numbers. I first define and explore first order logic, and model theory. Then, I prove the compactness theorem, and use this to form a nonstandard structure of the real numbers. Using this nonstandard structure, it it easy to to various proofs without the use of limits that would otherwise require their use.

Contents

1 Introduction

The founders of modern calculus had a less than perfect understanding of the nuts and bolts of what made it work. Both Newton and Leibniz used the notion of infinitesimal, without a rigorous understanding of what they were. Infinitely small real numbers that were still not zero was a hard thing for mathematicians to accept, and with the rigorous development of limits by the likes of Cauchy and Weierstrass, the discussion of infinitesimals subsided. Now, using first order logic for nonstandard analysis, it is possible to create a model of the real numbers that has the same properties as the traditional conception of the real numbers, but also has rigorously defined infinite and infinitesimal elements.

2 An Introduction to First Order Logic

2.1 Propositional Logic

The most basic form of logic is propositional logic, an it serves a building block for first order logic. Propositional logic takes statements, and links them together with "and" or "or". Basic sentence symbols are regarded as well-formed formulas, and combine with each other in specific ways to make larger well-formed formulas. However, there are no quantifiers in this basic form of logic. First order logic remedies this lack.

2.2 Logical Symbols

Predicate logic serves as a refinement of basic propositional calculus, or sentential logic. First order logic allows for the expression of more complicated conditions such as "for all" or "for some"; it allows for the quantification of the basic propositions first seen in propositional logic. There are a few basic symbols used in first order logical systems that come from propositional logic: the implication arrow ($\Rightarrow$), the negation symbol ($\neg$), the 'and' symbol ($\wedge$), the inclusive 'or' symbol ($\vee$), and the inference symbol ($\vdash$).

In addition to these symbols from propositional logic, we introduce the quantifier symbols $\forall$, the universal quantifier, and $\exists$, the existential quantifier.

2.3 Predicates, Constants and Functions

In first order logic, the equivalent of an english noun or noun phrase is what is called a term.

Definition 1. A term is as follows:
1. A variable, or individual constant is a term.
2. If $t_1, t_2, ..., t_n$ are all terms, $f_j^n(t_1, t_2, ..., t_n)$ is a term, where f is a function letter.
3. An expression is a term if an only if it is a term as defined in 1 or 2.

It is easiest to think of terms in the intuitive sense mentioned above as nouns that are then combined with predicate letters to form atomic formulas. Predicates define relations between different terms.

Definition 2. $P_j^n(t_1, t_2, ..., t_n)$ is an atomic formula, when P is a predicate letter, and t_i is a term $\forall$ i.

These atomic phrases further combine to form well-formed formulas, which are the essential building blocks for a first order logic.
An example of an atomic formula is $< xy$, or x is less than y, where x and y are both terms. The symbol $<$ is a two-place predicate symbol, which means that it defines a relation between the two components x and y.

2.4 Well-Formed Formulas

Well-formed formulas, or 'wffs', are expressions made from atomic formulas, and various quantifiers.

Definition 3. The classification of expressions as a Well-Formed Formula is dependent on three rules:
1. Every atomic formula is a wff.
2. If $\mathcal{A}$ and $\mathcal{B}$ are wfs, and x is a variable, then $(\neg\mathcal{A})$, $\mathcal{A} \Rightarrow \mathcal{B}$, and $((\forall y)\mathcal{A})$ are wffs.
3. An expression is a wff if an only if it is a wff as defined in 1 or 2.

Though by the definition of a well-formed formula, the above example of an atomic formula also serves as an example of a wff, you can make more complicated wffs out of the atomic phrases. For example, the phrase $< xy \Rightarrow \neg \approx xy$ is a wff because "x is less than y" and "x equals y" are both atomic formulas, and the implication and negation are allowed under section two of the definition.

Note: At first, it would seem that the quantifier $\exists$ cannot be in a wff, due to the definition's lack of provision for this quantifier. However, upon further examination we see that $\exists$ is logically equivalent to $\neg((\forall\text{y})\neg\text{X})$. That is, "there exists y such that y has property X" is logically equivalent to the statement "it is not true that for all y, y does not have the property X". Therefore, wffs can use the existential quantifier.

Using well formed phrases and quantifier logic, it becomes much easier to translate various phrases from english to mathematical language that it was using propositional logic. For example, the sentence, "Every person has a mother." can be translated as "$\forall y$, $\exists x$ such that $x\mathcal{M}y$, with $a\mathcal{M}b$ indicating that a is the mother of b.

3 Models

3.1 Structure

A structure of a first order logical language is roughly defined as the universe in which a first order logic operates. All of the parameters of a first order language must be accounted for in the structure.

Definition 4. A structure, Φ, of a first order language consists of a set and functions assigning the set of parameters of the first order language to specific aspects of the structure's set.
1. The structure, Φ assigns the universal quantifier symbol ($\forall$) some nonempty set, ϕ.
2. For each n-place predicate symbol with n different components, P_j^n, Φ assigns an n-place relation in ϕ
3. For each n-place function with n inputs, f_j^n, Φ assigns an n-place operation, $(f_j^n)^\Phi$, from ϕ^n to ϕ.
4. For each constant c, Φ assigns a constant c^Φ in the universe ϕ

For an example of a structure, let us consider a language with only the 2-place predicate relation $a\mathcal{D}b$ meaning a is more delicious than b. Now, we want to assign a structure to this language. If we use a structure with the universe, ϕ being all the different types of fruit, and $\mathcal{D}^\Phi$ being the set of pairs of fruit such that one fruit is more delicious than the other. This satisfies parts one and two of the definition of structure, and as this particular language has no functions, part three is also satisfied.

For an example of the structure of a language with an operation, let us turn to the language of ordered abelian groups. Ordered abelian groups have one 2-place predicate relation, $\leq$, where $a \leq b$ means that a is less than or equal to b. An ordered abelian group has several axioms that it must satisfy, including having a binary operation, and inverses. The operation must be associative, and communative. The group must be closed, and there must be an identity element, such that the binary operation between this element and any other element does not change the latter element. A structure which satisfies this language is the integers, with the operation addition. The $\leq$ relation is preserved in the integers, and addition a the binary, communicative, associative operation as specified. Negative and positive numbers are inverses of each-other, and 0 is the identity element. The integers are also closed under addition.

In the language of well-formed formulas, the integers, variables, and expressions such as $x + y$, or $-y$, or $z + 5$ are considered terms. Expressions such as $5 \leq 4$ or $x + y \leq -6$ are atomic phrases, the most basic well-formulated formulas. These can be combined into more complex wffs, such as $\forall x\ ,\ x + 5 \leq x + 10$

3.2 Truth

A well-formulated formula is not always true, it is just generated in a specific way outlined above.

For example, in the model and structure of abelian groups outlined above, the statement $3 + 5 \leq 1$ is a wff, although it is clearly not true. Similarly, the statement: there exists an integer, x, such that for every integer y, x is less than or equal to y, is a wff, but is not true.

If a sentence is true in a structure Φ, we say that that structure is a model of that sentence. Extending this, a structure Φ is a model for a set of wffs if and only if every wff in the set is true for Φ. However, to make a statement about truth, it is necessary to give a rigorous definition of truth. To this end, we introduce the concept of satisfaction.

3.2.1 Satisfaction

Let Φ be a structure of a language β. ϕ is the domain, or universe, of Φ. Σ is the set of all the sequences in ϕ, the domain of the structure Φ. We know that any n-tuple of numbers in Φ is included in Σ because if you assign the ith dimension of the n-tuple to the ith term of a sequence in ϕ, you can describe any n-tuple as a sequence in ϕ and therefore in Σ. Let s be a sequence in Σ, and $\mathcal{M}$ be a wff in the language β. We will now define what it means for s to satisfy $\mathcal{M}$, or Φ to satisfy $\mathcal{M}$ with s.

Definition 5. To do this we must define s*, a function that essentially takes the language to its structure.
1. s*$(x_j) = s_j$, where x_j is any variable, and s_j is a corresponding term in the universe ϕ.
2. For a constant c_j, s*$(c_j) = (c_j)^\Phi$, the representation of the constant in the universe ϕ.
3. If $t_1, t_2, ..., t_n$ are terms in the language β, then s*$(f_k^n(t_1, t_2, ..., t_n)) = (f_k^n)(s^*(t_1), s^*(t_2), ..., s^*(t_n))$. That is, s* is such that the representation of a function map f in the universe of Φ is the same whether the function maps the terms to an output before the output is mapped to ϕ or the terms are mapped to ϕ, and then mapped to an output by the function f.

Now that we have the language mapped into an arbitrary structure through the function s*, we can define satisfaction inductively for wffs, based on and axiomatic definition for the simplest of wffs, atomic formulas.

Definition 6. The definition of satisfaction is inductive, based on an axiomatic definition of satisfaction for atomic formulas.
1. An atomic formula, $\mathcal{M}$ in the language β is an n-place relation. Let us call that relation $\mathcal{M}_j^n$. It also has a corresponding n-place relation in the structure, Φ. Let's call that relation $(\mathcal{M}_j^n)^\Phi$. Then, the sequence s satisfies $\mathcal{M}$ if and only if $(s^*(t_1), s^*(t_2), ..., s^*(t_n))$ is in the relation $(\mathcal{M}_j^n)^\Phi$. This essentially means that the sequence s satisfies a wff if it is still regulated by the n-place relation even though it has been converted from an abstract set to a set in ϕ.
2. s satisfies $\neg\mathcal{M}$ iff s does not satisfy $\mathcal{M}$. 3. s satisfies $\mathcal{M} \Rightarrow \mathcal{N}$ iff s does not satisfy $\mathcal{M}$ or s satisfies $\mathcal{N}$.
4. s satisfies $(\forall(x_i)\mathcal{M})$ iff all sequences s_n that differ from s at the ith component satisfy $\mathcal{M}$

Now that the notion of satisfaction is formally defined, we can give a formal definition of truth, and of a model.

Definition 7. Truth and Models
1. A wff, $\mathcal{M}$, is true for the structure Φ iff every sequence in Σ satisfies $\mathcal{M}$.
2. A wff, $\mathcal{M}$, is false for the structure Φ iff every sequence in Σ does not satisfy $\mathcal{M}$
3. A structure Φ is a model for a set of wffs, iff every wff in the set is true for Φ

Along with the notion of truth, we get the notion of logical implication.

Definition 8. Logical Implication
For whatever structure for the language this implication is in, and for whatever function mapping the statements to the structure, if a structure Φ and a sequence s satisfy every member of $\mathcal{F}$, then they satisfy M

4 The Compactness Theorem

Now that we have developed the concept of a first order logical language, and concrete definitions for truth and models, it would seem that the only thing left to do is create a model for the real numbers which includes infinite elements. However, since the set of real numbers is an infinite one, it becomes harder to decide whether we can form a nonstandard structure with the same properties as the real numbers. This is where the compactness theorem comes in.

Theorem 1. The Compactness Theorem
1. If a set of wffs, $\mathcal{S}$, implies another wff, $\mathcal{M}$, then there is some finite subset of $\mathcal{S}$, $\mathcal{S}'$, that implies $\mathcal{M}$.
2. If every finite subset of $\mathcal{S}$ is satisfiable, then $\mathcal{S}$ is satisfiable.

To prove this theorem, we need to use the Completeness Theorem, and to that end we will introduce the concepts of soundness and completeness.

4.1 Soundness and Completeness

Definition 9. Theorem
Let $\mathcal{F}$ be a set of formulas. Taking a number of these formulas to be axiomatically true, or true with out proof, a theorem is a new formula that is deductible from the axiomatic formulas in $\mathcal{F}$. The rule of inference that we will use is that from the formulas α and $\alpha \Rightarrow \beta$, we can infer β

Before we prove the soundness theorem, it is necessary to mention a lemma, which we will use to prove the theorem.

Lemma 1. Every logical axiom is valid. This means that any axiom that can be deduced is true logically valid, or true for every structure and every function mapping the language to the structure. There are 6 different types of logical axioms based on the method of deduction.

Theorem 2. The Soundness Theorem
Let $\mathcal{F}$ continue to be a set of formulas. $\mathcal{M}$ is a theorem of $\mathcal{F} \Rightarrow \mathcal{F} \vdash \mathcal{M}$, that it, that $\mathcal{F}$ logically implies $\mathcal{M}$.

Proof. If $\mathcal{M}$ is logical, or already an element of $\mathcal{S}$, then this theorem is obviously true. Otherwise, $\mathcal{M}$ is obtained from $\mathcal{S}$ through the process of deduction, where $\mathcal{M} \Rightarrow \mathcal{T}$ and $\mathcal{T} \vdash \mathcal{M}$. □

Coupled to the soundness theorem is its converse, the completeness theorem. To prove the Completeness theorem, we first prove the following lemma.

Lemma 2. Any logically valid wff $\mathcal{M}$ of $\mathcal{S}$ is a theorem of $\mathcal{S}$

Proof. Suppose that for the set of wffs $\mathcal{S}$, $\mathcal{M}$ is logically valid, but not a theorem. Since $\mathcal{M}$ is not a theorem of $\mathcal{S}$, it can not be proved from $\mathcal{S}$, which means that we can add $\not\mathcal{M}$ to $\mathcal{S}$ as an axiom, and $\mathcal{S}$ is still consistent. If we make a model for $\mathcal{S}'$, which includes $\not\mathcal{M}$, we find that $\mathcal{M}$ is both false, as specified by $\mathcal{S}'$, and true, because it is logical, for $\mathcal{M}$. This is a contradiction. □

Theorem 3. The Completeness Theorem
If $\mathcal{S}$ implies $\mathcal{M}$, then $\mathcal{M}$ is a theorem of $\mathcal{S}$

Proof. Coupled with the soundness theorem, the preceding lemma proves that □

With these two theorems, we can prove the Compactness Theorem.

Theorem 4. The Compactness Theorem
1. If a set of wffs, $\mathcal{S}$, implies another wff, $\mathcal{M}$, then there is some finite subset of $\mathcal{S}$, $\mathcal{S}'$, that implies $\mathcal{M}$.
2. If every finite subset of $\mathcal{S}$ is satisfiable, then $\mathcal{S}$ is satisfiable.

Proof. 1.$\mathcal{S}$ implies $\mathcal{M}$. Then, by the completeness theorem, $\mathcal{M}$ is a theorem of T. Because a deduction must take a finite number of steps, there must be a finite subset of T that $\mathcal{M}$ is a theorem of. Then, we know by the soundness theorem that this subset logically implies $\mathcal{M}$.
2. If every finite subset of $\mathcal{S}$ is satisfiable, then every finite set is consistent, that is, a set doesn't imply $\mathcal{M}$ and $\neg\mathcal{M}$. Since deductions are finite, no deduction in the infinite set $\mathcal{S}$ will be longer than the longest finite subsets. Therefore $\mathcal{S}$ is consistent, which means, by the completeness theorem, that $\mathcal{S}$ is satisfiable. □

5 Nonstandard Analysis

5.1 Making a Nonstandard Structure

The language that we want to use for nonstandard analysis needs to include symbols for all operations on $\mathbb{R}$. We will call the standard structure for this language $\mathcal{R}$. This structure has a universe $\mathbb{R}$, the real numbers. This structure is very familiar, since the real numbers and operations on the real numbers are among the most basic aspects of mathematics.

However, we want to create a nonstandard structure for the language that usually operates on the real numbers. Not only this, but we also want this structure to be a model for the same wffs as $\mathcal{R}$, so we can use this structure to prove theorems that are also true in $\mathcal{R}$.

Let $\mathcal{S}$ be the union of the set of true sentences of $\mathcal{R}$, also known as the theory of $\mathcal{R}$ and constants c_n such that $c_n < r$, c_n being an element of $\mathbb{R}$. By changing r to a large real number, any finite set $\mathcal{S}$ can be satisfied. Therefore, by the compactness theorem, there is an infinite structure Φ, with a universe ϕ and a a constant c in ϕ that satisfies $\mathcal{S}$. This structure, Φ, is a model for the theory of $\mathcal{R}$. This means that $\mathcal{R}$ and Φ are elementarily equivalent, meaning that any wff that is true for one structure is true for the other.

We easily can show that any function or operation that is true in $\mathcal{R}$ is true in Φ, by creating an isomorphic map from $\mathcal{R}$ into Φ. Assuming that the universes $\mathbb{R}$ and ϕ are not originally the same, we can modify the universe ϕ by replacing the elements of ϕ elements of $\mathbb{R}$ mapped to with elements of $\mathbb{R}$. This leaves ϕ with all the elements of the real numbers, and possibly more. Now, since $\mathcal{R}$ and Φ are elementarily equivalent, any theorem, relation, or operation have the same properties in both $\mathbb{R}$ and ϕ. Because of the isomorphic function between these two structures, properties, operations, and relations are directly comparable from $\mathbb{R}$ to ϕ, and back from the image of $\mathbb{R}$ in ϕ to $\mathbb{R}$.

However, this does not yet tell us that this "nonstandard" structure is any different from the standard structure of $\mathcal{R}$, as we have not discerned any distinguishing features. Now we want to show that $\mathcal{R}$ is a substructure of Φ. First of all, we know that there is a point a in ϕ such that if r=a, then $\mathcal{S}$ is satisfied. This element b must not be a real number, because for every real number x, there is a corresponding x* in ϕ, due to the isomorphism between $\mathbb{R}$ and ϕ described above. We know that the structure Φ and the universe ϕ satisfy $\mathcal{S}$. If b were a real number, then only real numbers less than b would be included in the universe ϕ. However, since there is a x* in ϕ that corresponds with every x in $\mathbb{R}$, then b must be an infinite element. Additionally its reciprocal, 1/b, is an infinitesimal element of ϕ.

The most obvious aspect of nonstandard structures applicable to basic proofs in calculus and analysis is the existence of infinitesimal elements. However, before we use infinitesimal elements too freely, we must outline some basic properties of being infinitesimal.

Definition 10. Infinitesimal
1. Let $\mathcal{I}$ denote the set of infinitesimals.
2. $\mathcal{I}$ is the set of x in ϕ such that $x < y$ for all positive y in $\mathbb{R}$
3. x is infinitely close to y ($x \approx y$) if x-y is infinitesimal.

Theorem 5. $\mathcal{I}$ is closed under addition, subtraction, and multiplication from $\mathcal{F}$, the set of finite elements of ϕ.

Proof. Let x and y be infinitesimals. Suppose that z is a positive real element of ϕ. Since real numbers can be divided by other real numbers to yield a third real number, z/2 is real. Since infinitesimals are smaller than any real number, $x < z/2$ and $y < z/2$, for any positive real z. Then, we know that $|x \pm y| < z$, and therefore the infinitesimals are closed under addition and subtraction. Since z is finite, $z < b$, for some finite element b. Then, $x < z/b$, since z/b is finite. Thus, $x * b < z/b * b = z$. Since z an b are any real numbers, we have just proven that $\mathcal{I}$ is closed for finite elements of ϕ. □

In order to use the concept of "infinitely close" to its full potential, we must prove some basic properties of "infinite closeness".

Theorem 6. 1.$\approx$ is an equivalence relation.
2. If $a \approx b$ and $c \approx d$, then $a + c \approx b + d$
3. If $a \approx b$ and $c \approx d$, then $a * c \approx b * d$

Proof. 1. To be an equivalence relation, $\approx$ must be reflexive, symmetric, and transitive. $\approx$ is obviously reflexive, as 0 is an infinitesimal. $\approx$ is also symmetric, because if b is an infinitesimal, so is -b. Finally, it is transitive, because if d is an infinitesimal and a, b and c are real numbers, and $a - b \leq d$ and $b - c \leq d$, then $a - c \leq 2d$. Since 2d is also an infinitesimal (Theorem 5), $\approx$ is transitive. Thus, $\approx$ is an equivalence relation.
2.Let $a \approx b$ and $c \approx d$. Let e be an infinitesimal. Assume, without loss of generality, that $a > b$ and $c > d$. Then, $a - b \leq e$, and $c - d \leq e$. We know that 2e is still infinitesimal, and also that $a - b + c - d \leq 2e \Rightarrow (a + c) - (b + d) \leq 2e \Rightarrow a + c \approx b + d$
3. Assuming that $a * c > b * d$, $a * c \approx b * d$ if $a * c - b * d \leq e * (a + b)$, where a,b,c,d are finite and e is an infinitesimal. Assuming that $a > b$ and $c > d$, since $a \approx b$ and $c \approx d$, $a \leq b + e$ and $c \leq d + e$. Thus, $a * c \leq (b + e)c = b * c + e * c \leq b * (d + e) + e * c = b * d + e(c + b)$. Thus, $a * c - b * d \leq b * d + e(c + b) - b * d = e(c + b)$. □

5.2 Applications of a Nonstandard Structure

Now, let's do some mathematical proofs using the concept of infinitesimals instead of limits.

One of the basic concepts in the study of functions is the concept of continuity. Continuous functions are important in the study of calculus and analysis, and it is easy to redefine the notion of continuity using infinitesimals

Definition 11. Continuity
The traditional definition of continuity is that a function f is continuous if for every ϵ there exists a real $\delta > 0$ such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \epsilon$. We can rephrase that using infinitesimals, by defining f as continuous if when f(x) is infinitely close to f(y), x is infinitely close to y. Of course, we need to show that these two definitions of continuity are equivalent.

Proof. Suppose x and y are infinitely close, $|x - y| < \delta$ is obviously true, for any real $\delta > 0$. If f(x) and f(y) are infinitely close, it is also obviously true that $|f(x) - f(y)| < \epsilon$ for any real $\epsilon > 0$.
Suppose that for every ϵ, $\exists \delta > 0$ such that if $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$. Now suppose that delta is an infinitesimal. Since the infinitesimal delta is smaller than any real delta, for that delta, $|f(x) - f(y)| < \epsilon$ for any real epsilon. Therefore, f(x) and f(y) are infinitely close. □

Instead of taking a limit of a variable going to zero for things such as derivatives, we can just use an infinitesimal.

Definition 12. Limits
4 $\lim_{a \to b} f(a) = c$ traditionally means that for any real $\epsilon > 0$, there exists a δ such that if $|a - b| < \delta$, then $|f(a) - c| < \epsilon$. This is equivalent to the statement that when a and b are infinitesimally close, f(a) and c are infinitesimally close.

Proof. The arguments for the nonstandard definition of limits works in the same way as the arguments for the nonstandard definition of continuity detailed above. □

Definition 13. Derivative
In basic calculus, a derivative is defined as $f'(x) = \lim_{h \to 0} \frac{f(x+h)-f(x)}{h}$. Now that we can use infinitesimals, however, we can define a derivative as follows: Let a be an infinitesimal, x be a variable in ϕ, and f be function. $f'(x) = \frac{f(x+a)-f(x)}{a}$. This is obviously similar to the traditional definition of a derivative, only without the limit notation.

Using this definition, we can prove some basic theorems of calculus.

Theorem 7. Differentiability Implies Continuity
If $f'(a)$ exists, the f is continuous at a. Since f'(x) exists, $f'(x) \approx \frac{f(x+a)-f(x)}{a}$. Since a is an infinitesimal, and f'(x) is finite, $f'(x) * a$ is an infinitesimal, and therefore f(x-a) and f(x) are infinitely close. Thus, f is continuous.

6 Sources

Enderton, Herbert. "A Mathematical Introduction to Logic." New York, Academic Press 1972.
Mendelson, Elliott. "Introduction to Mathematical Logic: Fifth Edition." Boca Raton, CRC 2010.