

Nonstandard analysis as a completion of standard analysis

27 November, 2010 in [expository](#), [math.CA](#), [math.LO](#) | Tags: [Bolzano-Weierstrass theorem](#), [correspondence principle](#), [countable saturation](#), [nonstandard analysis](#), [szemerédi regularity lemma](#), [ultralimit analysis](#) | by [Terence Tao](#)

Many structures in mathematics are *incomplete* in one or more ways. For instance, the field of rationals \mathbf{Q} or the reals \mathbf{R} are *algebraically incomplete*, because there are some non-trivial algebraic equations (such as $x^2 = 2$ in the case of the rationals, or $x^2 = -1$ in the case of the reals) which could *potentially* have solutions (because they do not imply a necessarily false statement, such as $1 = 0$, just using the laws of algebra), but do not *actually* have solutions in the specified field.

Similarly, the rationals \mathbf{Q} , when viewed now as a metric space rather than as a field, are also *metrically incomplete*, because there exist sequences in the rationals (e.g. the decimal approximations $3, 3.1, 3.14, 3.141, \dots$ of the irrational number π) which could *potentially* converge to a limit (because they form a [Cauchy sequence](#)), but do not *actually* converge in the specified metric space.

A third type of incompleteness is that of *logical incompleteness*, which applies now to formal theories rather than to fields or metric spaces. For instance, [Zermelo-Frankel-Choice \(ZFC\) set theory](#) is logically incomplete, because there exist statements (such as the consistency of ZFC) which could *potentially* be provable by the theory (because it does not lead to a contradiction, or at least so we believe, just from the axioms and deductive rules of the theory), but is not *actually* provable in this theory.

A fourth type of incompleteness, which is slightly less well known than the above three, is what I will call *elementary incompleteness* (and which model theorists call the failure of the *countable saturation property*). It applies to any structure that is describable by a first-order language, such as a field, a metric space, or a universe of sets. For instance, in the language of ordered real fields, the real line \mathbf{R} is elementarily incomplete, because there exists a sequence of statements (such as the statements $0 < x < 1/n$ for natural numbers $n = 1, 2, \dots$) in this language which are *potentially* simultaneously satisfiable (in the sense that any finite number of these statements can be satisfied by some real number x) but are not *actually* simultaneously satisfiable in this theory.

In each of these cases, though, it is possible to start with an incomplete structure and *complete* it to a much larger structure to eliminate the incompleteness. For instance, starting with an arbitrary field k , one can take its algebraic completion (or [algebraic closure](#)) \bar{k} ; for instance, $\mathbf{C} = \bar{\mathbf{R}}$ can be viewed as the algebraic completion of \mathbf{R} . This field is usually significantly larger than the original field k , but contains k as a subfield, and every element of \bar{k} can be described as the solution to some polynomial equation with coefficients in k . Furthermore, \bar{k} is now [algebraically complete](#) (or algebraically closed): every polynomial equation in \bar{k} which is potentially satisfiable (in the sense that it does not lead to a contradiction such as $1 = 0$ from the laws of algebra), is actually satisfiable in \bar{k} .

Similarly, starting with an arbitrary metric space X , one can take its [metric completion](#) \bar{X} ; for instance, $\mathbf{R} = \bar{\mathbf{Q}}$ can be viewed as the metric completion of \mathbf{Q} . Again, the completion \bar{X} is usually much larger than the original metric space X , but contains X as a subspace, and every element of \bar{X} can be described as the limit of some Cauchy sequence in X . Furthermore, \bar{X} is now a complete metric space: every sequence in \bar{X} which is potentially convergent (in the sense of being a Cauchy sequence), is now actually convergent in \bar{X} .

In a similar vein, we have the [Gödel completeness theorem](#), which implies (among other things) that for

any consistent first-order theory T for a first-order language L , there exists at least one *completion* \overline{T} of that theory T , which is a consistent theory in which every sentence in L which is potentially true in \overline{T} (because it does not lead to a contradiction in \overline{T}) is actually true in \overline{T} . Indeed, the completeness theorem provides at least one model (or [structure](#)) \mathfrak{U} of the consistent theory T , and then the completion $\overline{T} = \text{Th}(\mathfrak{U})$ can be formed by interpreting every sentence in L using \mathfrak{U} to determine its truth value. Note, in contrast to the previous two examples, that the completion is usually not unique in any way; a theory T can have multiple inequivalent models \mathfrak{U} , giving rise to distinct completions of the same theory.

Finally, if one starts with an arbitrary structure \mathfrak{U} , one can form an *elementary completion* $^*\mathfrak{U}$ of it, which is a significantly larger structure which contains \mathfrak{U} as a substructure, and such that every element of $^*\mathfrak{U}$ is an elementary limit of a sequence of elements in \mathfrak{U} (I will define this term shortly). Furthermore, $^*\mathfrak{U}$ is elementarily complete; any sequence of statements that are potentially simultaneously satisfiable in $^*\mathfrak{U}$ (in the sense that any finite number of statements in this collection are simultaneously satisfiable), will actually be simultaneously satisfiable. As we shall see, one can form such an elementary completion by taking an [ultrapower](#) of the original structure \mathfrak{U} . If \mathfrak{U} is the *standard universe* of all the *standard* objects one considers in mathematics, then its elementary completion $^*\mathfrak{U}$ is known as the *nonstandard universe*, and is the setting for [nonstandard analysis](#).

As mentioned earlier, completion tends to make a space much larger and more complicated. If one algebraically completes a finite field, for instance, one necessarily obtains an infinite field as a consequence. If one metrically completes a countable metric space with no isolated points, such as \mathbb{Q} , then one necessarily obtains an uncountable metric space (thanks to the [Baire category theorem](#)). If one takes a logical completion of a consistent first-order theory that can model [true arithmetic](#), then this completion is no longer describable by a recursively enumerable schema of axioms, thanks to [Gödel's incompleteness theorem](#). And if one takes the elementary completion of a countable structure, such as the integers \mathbb{Z} , then the resulting completion $^*\mathbb{Z}$ will necessarily be uncountable.

However, there are substantial benefits to working in the completed structure which can make it well worth the massive increase in size. For instance, by working in the algebraic completion of a field, one gains access to the full power of [algebraic geometry](#). By working in the metric completion of a metric space, one gains access to powerful tools of real analysis, such as the [Baire category theorem](#), the [Heine-Borel theorem](#), and (in the case of Euclidean completions) the [Bolzano-Weierstrass theorem](#). By working in a logically and elementarily completed theory (aka a *saturated model*) of a first-order theory, one gains access to the branch of model theory known as *definability theory*, which allows one to analyse the structure of definable sets in much the same way that algebraic geometry allows one to analyse the structure of algebraic sets. Finally, when working in an elementary completion of a structure, one gains a *sequential compactness* property, analogous to the Bolzano-Weierstrass theorem, which can be interpreted as the foundation for much of nonstandard analysis, as well as providing a unifying framework to describe various correspondence principles between finitary and infinitary mathematics.

In this post, I wish to expand upon these above points with regard to elementary completion, and to present nonstandard analysis as a completion of standard analysis in much the same way as, say, complex algebra is a completion of real algebra, or real metric geometry is a completion of rational metric geometry.

— 1. Elementary convergence —

In order to understand the concept of a metric completion of a metric space $X = (X, d)$, one needs to know about the distinction between a Cauchy sequence and a convergent sequence. Similarly, to talk about the elementary completion of a structure \mathfrak{U} , one needs the notion of an elementarily Cauchy sequence and an elementarily convergent sequence.

Let us set out some notation. We assume that we have some [first-order language](#) L , which allows one to

form sentences involving the first-order logical symbols ($\forall, \exists, \vee, \wedge, \neg, \implies$, etc.), variables of one or more types, the equality symbol $=$, some constant symbols, and some operations and relations. For instance:

- L could be the language of multiplicative groups, in which there is only one type of object (a group element), a constant symbol e , a binary operation \cdot from pairs of group elements to group elements, and a unary operation $()^{-1}$ from group elements to group elements.
- L could be the language of real ordered fields, in which there is one type of object (a field element), constant symbols $0, 1$, binary operations $+, \cdot$, and unary operations $-, ()^{-1}$ (with the latter only being defined for non-zero elements), and the order relation $<$.
- L could be the language of (formal) metric spaces, in which there are two types of objects (points in the space, and real numbers), the constants, operations and relations of a real ordered field, and a metric operation d from pairs of points in the space to real numbers.
- L could be the language of sets, in which there is one type of object (a set) and one relation \in .
- etc., etc.

We assume that the language has at most countably many types, constants, operations, and relations. In particular, there are at most countably many sentences in L .

A **structure** \mathfrak{U} for a language L is a way of interpreting each of the object classes in L as a set, and each of the constants, operations, and relations as an elements, functions, and relations on those sets respectively. For instance, a structure for the language of groups would be a set G , together with a constant symbol $e \in G$, a binary function $\cdot : G \times G \rightarrow G$, and a unary operation $()^{-1} : G \rightarrow G$. In particular, groups are structures for the language of groups, but so are many non-groups. Each structure \mathfrak{U} can be used to *interpret* any given sentence S in L , giving it a truth value of true or false. We write $\mathfrak{U} \models S$ if S is interpreted to be true by \mathfrak{U} . For instance, the axioms of a group can be expressed as a single sentence A , and a structure \mathfrak{U} for the language of groups is a group if and only if $\mathfrak{U} \models A$.

Now we introduce the notion of elementary convergence.

Definition 1 (Elementary convergence) Let \mathfrak{U} be a structure for a language L , and let x_1, x_2, \dots be a sequence of objects in \mathfrak{U} (all of the same type). Let x be another object in \mathfrak{U} of the same type as the x_n .

- We say that the sequence x_n is *elementarily Cauchy* if, for every predicate $P(x)$ that takes one variable of the same type as the x_n as input, the truth value of $P(x_n)$ becomes eventually constant (i.e. either $P(x_n)$ is true for all sufficiently large n , or $P(x_n)$ is false for all sufficiently large n). We write this eventual truth value as $\lim_{n \rightarrow \infty} P(x_n)$.
- We say that the sequence x_n is *elementarily convergent* to x if we have $\lim_{n \rightarrow \infty} P(x_n) = P(x)$ for every predicate $P(x)$ that takes one variable of the same type as the x_n or x as input.

Remark 1 One can view the predicates P (or more precisely, the sets $\{x \in \mathfrak{U} : P(x) \text{ true}\}$) as generating a topology on \mathfrak{U} (or more precisely, on the domain of one of the object types of L in \mathfrak{U}), in which case elementary convergence can be interpreted as convergence in this topology. Indeed, as there are only countably many predicates, this topology is metrisable.

To give an example, let us use the language of ordered fields L , with the model \mathbf{R} , and pick a transcendental number x , e.g. $x = \pi$. Then the sequence $x + \frac{1}{n}$ is elementarily convergent to x . The reason for this is that the language L is fairly limited in nature, and as such it can only define a fairly small number of sets; in particular, if P is a predicate of one variable, then the [Tarski-Seidenberg](#)

[theorem](#) tells us that the set $\{y \in \mathbf{R} : P(y) \text{ true}\}$ cut out by that set has to be a [semi-algebraic set](#) over the algebraic reals, i.e. a finite union of (possibly unbounded) intervals (which can be open, closed, or half-open) whose endpoints are algebraic reals. In particular, a transcendental number x , if it lies in such a set, lies in the interior of such a set, and so $x + \frac{1}{n}$ will also lie in such a set for n large enough, and similarly if x lies outside such a set.

In contrast, if one picks an algebraic number for x , such as $x = \sqrt{2}$, then $x + \frac{1}{n}$ does *not* converge in an elementary sense to x , because one can find a predicate such as $P(y) := (y^2 = 2)$ which is true for x but not true for any of the $x + \frac{1}{n}$. So the language L has sufficiently “poor vision” that it cannot easily distinguish a transcendental number such as π from nearby numbers such as $\pi + \frac{1}{n}$, but its vision is significantly better at algebraic numbers, and in particular can distinguish $\sqrt{2}$ from $\sqrt{2} + \frac{1}{n}$ easily. So we see that elementary convergence is, in this case, a slightly stronger concept than the usual topological or metric notion of convergence on \mathbf{R} .

In the case of the real model \mathbf{R} of ordered fields, elementary limits are unique, but this is not the case in general. For instance, in the language of fields, and using the complex model \mathbf{C} , any given complex number z is elementarily indistinguishable from its complex conjugate \bar{z} , and so any sequence z_n of complex numbers that would converge elementarily to z , would also converge elementarily to \bar{z} . (In fact, there is an enormous Galois group $\text{Gal}(\mathbf{C}/\mathbf{Q})$, the action of which is completely undetectable with regards to elementary convergence.)

A related problem is that the operations on a structure \mathfrak{U} are not necessarily continuous with respect to these elementary limits. For instance, if x_n, y_n are sequences of real numbers that converge elementarily to x, y respectively, it is not necessarily the case that $x_n + y_n$ converge to $x + y$ (consider for instance the case when $x_n = \pi + 1/n$ and $y_n = -\pi + 1/n$).

One way to partially resolve these problem is to consider the convergence not just of sequences of individual objects x_n , but of sequences of families $(x_{n,\alpha})_{\alpha \in A}$ of objects:

Definition 2 (Joint elementary convergence) Let \mathfrak{U} be a structure a language L , let A be a set, and for each natural number n , let $(x_{n,\alpha})_{\alpha \in A}$ be a tuple of of elements in \mathfrak{U} , and let $(x_\alpha)_{\alpha \in A}$ be another tuple in \mathfrak{U} , with each $x_{n,\alpha}$ having the same type as x_α .

- We say that the tuples $(x_{n,\alpha})_{\alpha \in A}$ are *jointly elementarily Cauchy* if, for every natural number m , every predicate $P(y_1, \dots, y_m)$ of m variables in L of the appropriate type, and every $\alpha_1, \dots, \alpha_m \in A$, the truth value of $P(x_{n,\alpha_1}, \dots, x_{n,\alpha_m})$ is eventually constant.
- We say that the tuples $(x_{n,\alpha})_{\alpha \in A}$ are *jointly elementarily convergent* to $(x_\alpha)_{\alpha \in A}$ if, for every natural number m , every predicate $P(y_1, \dots, y_m)$ of m variables in L of the appropriate type, and every $\alpha_1, \dots, \alpha_m \in A$, the truth value of $P(x_{n,\alpha_1}, \dots, x_{n,\alpha_m})$ converges to the truth value of $P(x_{\alpha_1}, \dots, x_{\alpha_m})$ as $n \rightarrow \infty$.

For instance, using the complex model \mathbf{C} of the language of fields, if z_n converges elementarily to (say) i , then we cannot prevent z_n from also converging elementarily to $-i$. (Indeed, it is not hard to see that z_n converges elementarily to i if and only $z_n \in \{-i, +i\}$ for all sufficiently large n .) But if we ask that (z_n, i) jointly converges to (i, i) , then (z_n, i) will not also jointly converge to $(-i, i)$ (though it does jointly converge to $(-i, -i)$).

In a similar fashion, if x_n, y_n are reals that converge *jointly* elementarily to x, y , then $x_n + y_n$ will converge elementarily to $x + y$ also.

Now we give a more sophisticated example. Here, L is the language of set theory, and \mathfrak{U} is a model of ZFC. In ZFC set theory, we can of course construct most of the objects we are used to in modern

mathematics, such as a copy $\mathbf{R}_{\mathfrak{U}}$ of the real line, a copy $\mathbf{N}_{\mathfrak{U}}$ of the natural numbers, and so forth. Note that \mathfrak{U} 's interpretation $\mathbf{N}_{\mathfrak{U}}$ of the natural numbers may be different from the “true” natural numbers \mathbf{N} ; in particular, in non-standard models of set theory, $\mathbf{N}_{\mathfrak{U}}$ may be much larger than \mathbf{N} (e.g. it may be an ultrapower ${}^*\mathbf{N}$ of \mathbf{N}). Because of this, we will be careful to subscript \mathfrak{U} 's copies of such objects in order to distinguish them from their true counterparts, though it will not make much difference for this immediate example.

We can also define in \mathfrak{U} all the formal apparatus needed for probability theory, such as a probability space $(\Omega, \mathcal{B}, \mathbf{P})$ and a real-valued random variable $X : \Omega \rightarrow \mathbf{R}_{\mathfrak{U}}$ on that space.

Now suppose that inside \mathfrak{U} we have a sequence $(\Omega_n, \mathcal{B}_n, \mathbf{P}_n)$ of probability spaces, and a sequence $X_n : \Omega_n \rightarrow \mathbf{R}_{\mathfrak{U}}$ of random variables on these probability spaces. Now suppose the quintuple $(\Omega_n, \mathcal{B}_n, \mathbf{P}_n, X_n, \mathbf{R}_{\mathfrak{U}})$ is jointly elementarily convergent to a limit $(\Omega, \mathcal{B}, \mathbf{P}, X, \mathbf{R}_{\mathfrak{U}})$. The axioms of being a probability space can be encoded inside the first order language of set theory, so the limit $(\Omega, \mathcal{B}, \mathbf{P})$ is also a probability space (as viewed inside \mathfrak{U}). Similarly, $X : \Omega \rightarrow \mathbf{R}_{\mathfrak{U}}$ is a random variable on this probability space.

Now let q, r be rational numbers. If $\mathbf{P}(X > q) > r$, then by the definition of elementary convergence (and the fact that rational numbers can be defined using an expression of finite length in the language L), we see that $\mathbf{P}_n(X_n > q) > r$ holds for all sufficiently large n . From this, one can deduce that the X_n [converge in distribution](#) to X . Thus we see that in this case, joint elementary convergence is at least as strong as convergence in distribution, though (much as with the example with elementary convergence in \mathbf{R} using the language of ordered fields) the two notions of convergence are not equivalent.

We included $\mathbf{R}_{\mathfrak{U}}$ in the quintuple due to the use of real numbers such as $\mathbf{P}(X > q)$ in the above discussion, but it is not strictly necessary, because one can construct $\mathbf{R}_{\mathfrak{U}}$ uniquely in \mathfrak{U} from the axioms of set theory by using one of the standard constructions of the real numbers. But note that while we may use the set $\mathbf{R}_{\mathfrak{U}}$ of real numbers in the above elementary convergence, one cannot invoke specific real numbers unless they are “constructible” in the sense that they can be uniquely specified in the language L . If one wished to be able to use arbitrary real numbers as constants, one would not only place $\mathbf{R}_{\mathfrak{U}}$ into the quintuple, but also place in every element x of $\mathbf{R}_{\mathfrak{U}}$ into the tuple (thus making the tuple quite large, and most likely uncountable, though note from [Skolem's paradox](#) that it is possible for $\mathbf{R}_{\mathfrak{U}}$ to be (externally) countable even as it is uncountable from the internal perspective of \mathfrak{U}).

As we see from the above discussion, joint elementary convergence is a useful notion even when some of the elements in the tuple are constant. We isolate this case with another definition:

Definition 3 (Joint relative elementary convergence) Let \mathfrak{U} be a structure for a language L , let A, B be sets, and for each natural number n , let $(x_{n,\alpha})_{\alpha \in A}$ be a tuple of elements in \mathfrak{U} , and let $(x_\alpha)_{\alpha \in A}$ and $(c_\beta)_{\beta \in B}$ be further tuples in \mathfrak{U} . We say that the tuples $(x_{n,\alpha})_{\alpha \in A}$ are *jointly elementarily convergent to $(x_\alpha)_{\alpha \in A}$ relative to the constants $(c_\beta)_{\beta \in B}$* if the disjoint union $(x_{n,\alpha})_{\alpha \in A} \uplus (c_\beta)_{\beta \in B}$ is jointly elementarily convergent to $(x_\alpha)_{\alpha \in A} \uplus (c_\beta)_{\beta \in B}$.

We define the notion of $(x_{n,\alpha})_{\alpha \in A}$ are *jointly elementarily Cauchy relative to the constants $(c_\beta)_{\beta \in B}$* similarly.

Informally, if $(x_{n,\alpha})_{\alpha \in A}$ is jointly elementarily convergent to $(x_\alpha)_{\alpha \in A}$ relative to $(c_\beta)_{\beta \in B}$ if the language L is unable to asymptotically distinguish the $x_{n,\alpha}$ from the x_α , even if it “knows” about the constants c_β .

— 2. Elementary completion —

Not every sequence in a structure is elementarily Cauchy or elementarily convergent. However, we have the following simple fact:

Proposition 4 (Arzelá-Ascoli) Let \mathfrak{U} be a structure for a language L , and let x_n be a sequence of elements in \mathfrak{U} (all of the same type). Then there is a subsequence of the x_n that is elementarily Cauchy.

Proof: There are at most countably many predicates $P_1(x), P_2(x), \dots$ of a single variable of the right type in L . By the infinite pigeonhole principle, we can find a subsequence $x_{n_{1,1}}, x_{n_{1,2}}, \dots$ of the x_n such that $P_1(x_{n_{1,i}})$ is eventually constant. We can find a further subsequence $x_{n_{2,1}}, x_{n_{2,2}}, \dots$ of that sequence for which $P_2(x_{n_{2,i}})$ is eventually constant. We continue extracting subsequences $x_{n_{k,1}}, x_{n_{k,2}}, \dots$ of this nature for each $k = 1, 2, \dots$, and then the diagonal sequence $k \mapsto x_{n_{k,k}}$ is elementarily Cauchy, as desired. \square

The same argument works when considering countably many variables and countably many constants (as there are still only countably many predicates to deal with):

Proposition 5 (Arzelá-Ascoli) Let \mathfrak{U} be a structure for a language L , let A, B be at most countable sets, and for each natural number n , let $(x_{n,\alpha})_{\alpha \in A}$ be a tuple of elements in \mathfrak{U} , and let $(c_\beta)_{\beta \in B}$ be another tuple in \mathfrak{U} . Then there is a subsequence $(x_{n_j,\alpha})_{\alpha \in A}$ which is jointly elementarily Cauchy relative to $(c_\beta)_{\beta \in B}$.

As in the metric case, not every elementarily Cauchy sequence is elementarily convergent, and not every sequence has an elementarily convergent subsequence. For instance, in the language of ordered fields, using the structure \mathbf{R} , the sequence $\frac{1}{n}$ has no elementarily convergent subsequence (because any limit of such a subsequence would be positive but also less than $1/n$ for arbitrarily large n , contradicting the [Archimedean property](#) of the reals). From Proposition 4, we conclude that the reals are elementarily incomplete; there must exist some subsequence of $1/n$ that is elementarily Cauchy, but not elementarily convergent.

However, we can always complete any structure \mathfrak{U} by passing to the [ultrapower](#) ${}^*\mathfrak{U}$. This concept has been discussed in previous blog posts, so we give only a quick review here.

For the rest of this post, we fix a single [non-principal ultrafilter](#) $\alpha_\infty \in \beta\mathbf{N} \setminus \mathbf{N}$ on the (standard) natural numbers \mathbf{N} . (See this [previous blog post](#) for some basic discussion of what non-principal ultrafilters are, and how they are used in non-standard analysis.) A property $P(\alpha)$ of a natural number α is said to hold *for all α sufficiently close to α_∞* if the set of α for which $P(\alpha)$ holds lies in the ultrafilter α_∞ .

Definition 6 (Ultrapower) Let \mathfrak{U} be a structure for some language L . Given two sequences $(x_\alpha)_{\alpha \in \mathbf{N}}$ and $(y_\alpha)_{\alpha \in \mathbf{N}}$ of objects in \mathfrak{U} , we say that the sequences are *equivalent* if one has $x_\alpha = y_\alpha$ for all α sufficiently close to α_∞ . The equivalence class associated to a given sequence $(x_\alpha)_{\alpha \in \mathbf{N}}$ will be called the *ultralimit* of the x_α and denoted $\lim_{\alpha \rightarrow \alpha_\infty} x_\alpha$. The *ultrapower* ${}^*\mathfrak{U}$ of \mathfrak{U} is the collection of all ultralimits $\lim_{\alpha \rightarrow \alpha_\infty} x_\alpha$ of sequences of objects in \mathfrak{U} . By identifying ${}^*x := \lim_{\alpha \rightarrow \alpha_\infty} x$ with x for every object x in \mathfrak{U} , we see that every object in \mathfrak{U} can be identified with an object in ${}^*\mathfrak{U}$. We refer to elements of \mathfrak{U} as *standard objects*, and elements of ${}^*\mathfrak{U}$ as *non-standard objects*.

Every relation and operation in \mathfrak{U} can be extended to ${}^*\mathfrak{U}$ by taking ultralimits. For instance, given a k -ary relation $R(y_1, \dots, y_k)$, and non-standard objects $x_i = \lim_{\alpha \rightarrow \alpha_\infty} x_{i,\alpha}$ for $i = 1, \dots, k$, we say that $R(x_1, \dots, x_k)$ holds in ${}^*\mathfrak{U}$ if and only if $R(x_{1,\alpha}, \dots, x_{k,\alpha})$ holds in \mathfrak{U} for all α sufficiently close to α_∞ . Similarly, given a k -ary operation $f(y_1, \dots, y_k)$ and non-standard objects $x_i = \lim_{\alpha \rightarrow \alpha_\infty} x_{i,\alpha}$, we define the non-standard object $f(x_1, \dots, x_k)$ to be the ultralimit $\lim_{\alpha \rightarrow \alpha_\infty} f(x_{1,\alpha}, \dots, x_{k,\alpha})$ of the standard objects $f(x_{1,\alpha}, \dots, x_{k,\alpha})$.

Ultrapowers are also discussed in more detail in [this previous blog post](#). A fundamental [theorem of Los asserts](#) that the ultrapower ${}^*\mathfrak{U}$ is [elementarily equivalent](#) to \mathfrak{U} : any sentence in \mathcal{L} which is true in \mathfrak{U} , is also true in ${}^*\mathfrak{U}$, and vice versa; this fact is also known as the [transfer principle](#) for nonstandard analysis. For instance, the ultrapower of an ordered field is an ordered field, the ultrapower of an algebraically closed field is an algebraically closed field, and so forth. One must be slightly careful, though, with models \mathfrak{U} that involve standard objects such as a copy \mathfrak{N} of the natural numbers, or a copy $\mathfrak{R}_{\mathfrak{U}}$ of the real numbers; the ultrapower ${}^*\mathfrak{U}$ will have their own non-standard copy $\mathfrak{N}_{*}\mathfrak{U} = {}^*\mathfrak{N}_{\mathfrak{U}}$ and $\mathfrak{R}_{*}\mathfrak{U} = {}^*\mathfrak{R}_{\mathfrak{U}}$ of these objects, which are considerably larger than their standard counterparts, in the sense that they contain many more elements. Thus, for instance, if one is taking the ultrapower of a standard probability space $(\Omega, \mathcal{B}, \mathbf{P})$, in which the probability measure $\mathbf{P} : \mathcal{B} \rightarrow \mathbf{R}$ takes values in the standard reals, the ultrapower $({}^*\Omega, {}^*\mathcal{B}, {}^*\mathbf{P})$ is a non-standard probability space, in which the non-standard probability measure ${}^*\mathbf{P} : {}^*\mathcal{B} \rightarrow {}^*\mathbf{R}$ now takes values in the non-standard reals.

One can view the ultrapower ${}^*\mathfrak{U}$ as the completion of \mathfrak{U} , in much the same way as the reals are a completion of the rationals:

Theorem 7 (Elementary completeness) Every elementarily Cauchy sequence x_n in an ultrapower ${}^*\mathfrak{U}$ is elementarily convergent.

This property is also known as [countable saturation](#).

Proof: We can write $x_n = \lim_{\alpha \rightarrow \alpha_\infty} x_{n,\alpha}$ for each natural number $n \in \mathbf{N}$ and a sequence $x_{n,\alpha}$ of standard objects in \mathfrak{U} . As before, we enumerate the predicates P_1, P_2, P_3, \dots of one variable. For each natural number $m \in \mathbf{N}$, the truth value of $P_m(x_n)$ becomes eventually constant; we will call this constant $\lim_{n \rightarrow \infty} P_m(x_n)$.

Now let M be a standard natural number. By construction, there exists an n_M such that

$$P_m(x_{n_M}) = \lim_{n \rightarrow \infty} P_m(x_n)$$

for all $1 \leq m \leq M$. As x_{n_M} is the ultralimit of the $x_{n_M,\alpha}$, there thus exists a set $E_M \in \mathcal{p}$ such that

$$P_m(x_{n_M,\alpha}) = \lim_{n \rightarrow \infty} P_m(x_n)$$

for all $\alpha \in E_M$. By replacing each E_M with $\bigcap_{M' \leq M} E_{M'}$ if necessary, we may assume that the E_M are decreasing: $E_1 \supset E_2 \supset \dots$

For each $\alpha \in \mathbf{N}$, let M_α be the largest integer in $\{0, \dots, \alpha\}$ such that $\alpha \in E_{M_\alpha}$, or $M_\alpha = 0$ if no such integer exists. By construction, we see that for any $m \in \mathbf{N}$, we have

$$P_m(x_{n_{M_\alpha},\alpha}) = \lim_{n \rightarrow \infty} P_m(x_n)$$

whenever $\alpha \in E_m$ and $\alpha \geq m$. If we then set x to be the non-standard object $x := \lim_{\alpha \rightarrow \alpha_\infty} x_{n_{M_\alpha},\alpha}$, we thus have

$$P_m(x) = \lim_{n \rightarrow \infty} P_m(x_n)$$

for each $m \in \mathbf{N}$, and thus x_n converges elementarily to x as required. \square

Combining this theorem with [Proposition 4](#) we conclude an analogue of the Bolzano-Weierstrass theorem for ultrapowers:

Corollary 8 (Bolzano-Weierstrass for ultrapowers) In an ultrapower ${}^*\mathfrak{U}$, every sequence x_n of non-standard objects in ${}^*\mathfrak{U}$ has an elementarily convergent subsequence x_{n_j} .

The same argument works (but with more complicated notation) for countable families of objects, and with countably many constants:

Theorem 9 (Bolzano-Weierstrass for ultrapowers, II) Let ${}^*\mathfrak{U}$ be an ultrapower, let A, B be at most countable, let $n \mapsto (x_{n,\alpha})_{\alpha \in A}$ be a sequence of tuples of nonstandard objects in ${}^*\mathfrak{U}$, and let $(c_\beta)_{\beta \in B}$ be another sequence of tuples of nonstandard objects. Then there is a subsequence $(x_{n_j,\alpha})_{\alpha \in A}$ which converges jointly elementarily to a limit $(x_\alpha)_{\alpha \in A}$ relative to the constants $(c_\beta)_{\beta \in B}$.

The proof of this theorem proceeds almost exactly as in the single variable case, the key point being that the number of predicates that one has to stabilise remains countable.

Remark 2 If one took the ultrafilter α_∞ over a larger set than the natural numbers \mathbb{N} , then one could make the sets A, B larger as well. Such larger saturation properties, beyond countable saturation, are useful in model theory (particularly when combined with the use of large cardinals, such as inaccessible cardinals), but we will not need them here.

Conversely, every nonstandard object can be viewed as the elementary limit of standard objects:

Proposition 10 Let $x \in {}^*\mathfrak{U}$ be a nonstandard object. Then there is a sequence $x_n \in \mathfrak{U}$ of standard objects that converges elementarily to x .

Proof: Let P_1, P_2, \dots be an enumeration of the predicates of one variable. For any natural number n , there exists a nonstandard object $y \in {}^*\mathfrak{U}$ such that $P_i(y)$ has the same truth value as $P_i(x)$ for all $i = 1, \dots, n$, namely $y = x$. By transfer, there must therefore exist a standard object $x_n \in \mathfrak{U}$ such that $P_i(x_n)$ has the same truth value as $P_i(x)$. Thus x_n converges elementarily to x , and the claim follows. \square

Exercise 1 If x is the ultralimit of a sequence x_n of standard objects, show that there is a subsequence x_{n_j} that converges elementarily to x .

Exercise 2 (Heine-Borel theorem for structures) Given any structure \mathfrak{U} , show that the following four statements are equivalent:

- (Countable saturation) If $P_1(x), P_2(x), \dots$ are a countable family of predicates, such that if any finite number of P_i are simultaneously satisfiable in \mathfrak{U} (i.e. for each n there exists $x_n \in \mathfrak{U}$ such that $P_i(x_n)$ holds for all $i = 1, \dots, n$), then the entire family of P_i are simultaneously satisfiable (i.e. there exists $x \in \mathfrak{U}$ such that $P_i(x)$ holds for all i).
- (Countable compactness) Every countable cover of \mathfrak{U} by sets of the form $\{x \in \mathfrak{U} : P(x) \text{ true}\}$ for some predicate P , has a finite subcover.
- (Elementary completeness) Every elementarily Cauchy sequence in \mathfrak{U} has an elementarily convergent subsequence.
- (Bolzano-Weierstrass property) Every elementary sequence in \mathfrak{U} has an elementarily convergent subsequence.

From Proposition 10 and Theorem 7 we see that ${}^*\mathbb{U}$ can be viewed as an elementary completion of \mathbb{U} , though the analogy with metric completion is not perfect because elementary limits are not unique.

The Bolzano-Weierstrass theorem for ultrapowers can then be used to derive the foundational properties of nonstandard analysis. For instance, consider the standard natural numbers $1, 2, 3, \dots$ in \mathbb{N} , and hence in ${}^*\mathbb{N}$. Applying the Bolzano-Weierstrass theorem for ultraproducts, we conclude that some subsequence n_j of natural numbers will converge elementarily to a non-standard natural number $N \in {}^*\mathbb{N}$. For any standard natural number $m \in \mathbb{N}$, we have $n_j > m$ for all sufficiently large j , and hence on taking elementary limits we have $N > m$ (since m is constructible). Thus we have constructed an *unbounded* nonstandard natural number, i.e. a number which is larger than all standard natural numbers.

In a similar spirit, we can also use the Bolzano-Weierstrass theorem construct *infinitesimal* nonstandard real numbers $0 < \epsilon = o(1)$ which are positive, but less than every standard positive real number (and in particular less than $1/n$ for any standard n).

More generally, we have the overspill principle: if $P(n, c_1, \dots, c_k)$ is a predicate involving some non-standard constants c_1, \dots, c_k , such that $P(n, c_1, \dots, c_k)$ is true for arbitrarily large standard natural numbers $n \in \mathbb{N}$, then it must also be true for at least one unbounded nonstandard natural number $n \in {}^*\mathbb{N}$. Indeed, one simply takes a sequence $n_j \in \mathbb{N}$ of standard natural numbers for which $P(n_j, c_1, \dots, c_k)$, and extracts a subsequence of these n_j which converges elementarily to a non-standard limit n (relative to c_1, \dots, c_k), which must then be unbounded. Contrapositively, if $P(n, c_1, \dots, c_k)$ holds for all unbounded n , then it must also hold for all sufficiently large standard n .

Similarly, we have the *underspill principle*: if a predicate $P(x, c_1, \dots, c_k)$ is true for arbitrarily small positive standard real x , then it must also be true for at least one infinitesimal positive non-standard real x ; and contrapositively, if it is true for all infinitesimal positive non-standard real x , then it is also true for all sufficiently small standard real x .

A typical application of these principles is in the nonstandard formulation of continuity:

Proposition 11 (Nonstandard formulation of continuity) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a standard function, which can then be extended by ultralimits to the nonstandard completion ${}^*f : {}^*\mathbb{R} \rightarrow {}^*\mathbb{R}$. Let $x \in \mathbb{R}$. Then the following are equivalent:

- f is continuous at x .
- One has ${}^*f(y) = f(x) + o(1)$ whenever y is a nonstandard real such that $y = x + o(1)$.

Proof: If f is continuous, then the “epsilon-delta” definition implies that whenever $y = x + o(1)$ (so that $|y - x| < \delta$ for every standard $\delta > 0$), one has $|{}^*f(y) - f(x)| < \epsilon$ for every standard ϵ (by transfer), and thus ${}^*f(y) = f(x) + o(1)$. Conversely, if f is discontinuous at x , then there exists a sequence x_n of standard reals converging to x such that $|f(x_n) - f(x)| > \epsilon$ for some standard $\epsilon > 0$; taking ultralimits using Bolzano-Weierstrass to extract a subsequence that is elementarily convergent relative to f , we thus have a non-standard $y = x + o(1)$ with $|{}^*f(y) - f(x)| > \epsilon$, so that ${}^*f(y) \neq f(x) + o(1)$. \square

Exercise 3 With the notation as above, show that f is uniformly continuous if and only if ${}^*f(y) = {}^*f(x) + o(1)$ whenever $x, y \in {}^*\mathbb{R}$ are such that $y = x + o(1)$.

Exercise 4 With the notation as above, show that f is differentiable at a standard real x with derivative $f'(x)$ if and only if ${}^*f(x + h) = f(x) + hf'(x) + o(h)$ for all nonstandard reals

$$h = o(1).$$

— 3. The correspondence principle —

One can use the Bolzano-Weierstrass theorem for ultrapowers to establish various versions of the correspondence principle, as [discussed previously](#) on this blog. A simple example occurs when demonstrating the equivalence of colouring theorems, such as the following:

Theorem 12 (van der Waerden theorem, infinitary version) Suppose the integers are coloured by finitely many colours. Then there exist arbitrarily long monochromatic arithmetic progressions.

Theorem 13 (van der Waerden theorem, finitary version) For every c and k there exists N such that whenever $\{-N, \dots, N\}$ is coloured by c colours, there exists a monochromatic arithmetic progression of length k .

It is clear that Theorem 13 implies Theorem 12. To deduce Theorem 12 from Theorem 13, we can argue as follows. Suppose Theorem 13 fails, then there exists c and k , and arbitrarily large standard natural number N for which there exists a c -colouring of $\{-N, \dots, N\}$ without any monochromatic arithmetic progressions of length k . Applying the overspill principle (or the Bolzano-Weierstrass theorem), there must then also exist an unbounded nonstandard natural number for which there exists a c -colouring of $\{-N, \dots, N\}$ without any monochromatic arithmetic progressions of length k . But the nonstandard interval $\{-N, \dots, N\}$ contains the standard integers \mathbb{Z} as a subset, thus the integers can now also be c -coloured without any monochromatic arithmetic progressions of length k , contradicting Theorem 12.

As another example, we can relate qualitative and quantitative results in algebraic geometry. For instance, [as noted previously on this blog](#), the following basic result in algebraic geometry,

Theorem 14 (Qualitative decomposition into varieties) Every algebraic set over an algebraically closed field can be decomposed into finitely many algebraic varieties.

is equivalent to the following more quantitative version:

Theorem 15 (Quantitative decomposition into varieties) Every algebraic set A of complexity at most M over an algebraically closed field can be decomposed at most C_M algebraic varieties, each of complexity at most C_M , where C_M depends only on M .

(For a definition of complexity see [the previous blog post](#).)

Clearly Theorem 15 implies Theorem 14. To show the converse implication, suppose that Theorem 15 failed, then there exists M such that for every standard natural number N there exists an algebraic set A_N of complexity at most M over some field k_N that cannot be decomposed into fewer than N algebraic varieties. We use the Bolzano-Weierstrass theorem for ultrapowers to extract a subsequence A_{N_j}, k_{N_j} that converges jointly elementarily to some limit A, k . As each of the k_{N_j} are algebraically closed fields, the elementary limit k is also. As the A_{N_j} were algebraic sets over k_{N_j} of uniformly bounded complexity, A is an algebraic set over k , and thus by Theorem 14 is decomposable into at most N_0 algebraic varieties for some finite N_0 . The property of being decomposable into at most N_0 algebraic varieties can be phrased in an elementarily open manner, i.e. as the disjunction of sentences in

first-order logic; this is essentially established in [the previous blog post](#), with the key point being that any top-dimensional component of an algebraic set has a lesser degree than that of the original set, and so has a uniform bound on complexity. Thus, we see that for all sufficiently large j , A_{N_j} must also be decomposable into at most N_0 algebraic varieties, a contradiction.

Our final example, namely the Furstenberg correspondence principle, is a bit more sophisticated. Here, we are demonstrating the equivalence of the following two statements:

Theorem 16 (Furstenberg recurrence theorem) Let (X, \mathcal{B}, μ, T) be a measure-preserving system, and let $A \subset X$ have positive measure. Let $k \geq 1$. Then there exists $r > 0$ such that $A \cap T^r A \cap \dots \cap T^{(k-1)r} A$ has positive measure.

Theorem 17 (Szemerédi's theorem) Every set of integers of positive upper density contains arbitrarily long arithmetic progressions.

It is easy to use Theorem [17](#) to show Theorem [16](#), so we focus on the reverse inclusion. Suppose that Theorem [17](#) failed, then there exists a standard integer $k \geq 1$, a standard real $\delta > 0$, and a standard set $A \subset \mathbb{Z}$ of integers of positive upper density at least δ that has no arithmetic progressions of length k . In particular, for arbitrarily large standard N , there exists a subset of $\{-N, \dots, N\}$ of density at least δ without any progressions of length k . Applying the overspill principle, there thus exists an unbounded nonstandard N and a nonstandard subset A of $\{-N, \dots, N\}$ of density at least δ without any progressions of length k .

Let \mathcal{A} be the collection of all nonstandard subsets of $\{-N, \dots, N\}$. (Note that not every subset of a nonstandard set remains nonstandard; it may instead merely be an external subset. See [this previous blog post](#) for further discussion.) This is a Boolean algebra. From countable saturation we see that this Boolean algebra has the following special property: if any set E in this Boolean algebra is partitioned into an (externally) countable family $(E_n)_{n \in \mathbb{N}}$ of further elements in this Boolean algebra, then all but finitely many of the E_n are empty. For if this were not the case, then by the axiom of choice, one could find a subsequence n_j and a set of elements x_{n_j} of E_{n_j} . Passing to a further subsequence, we can assume that the x_{n_j} converge elementarily to a limit x . But then this limit lies in E but not in any of the E_n , a contradiction.

From the above property, we see that any (external) finitely additive measure $\mu : \mathcal{A} \rightarrow [0, 1]$ on \mathcal{A} is automatically a [premeasure](#), and thus by the [Hahn-Kolmogorov extension theorem](#), can be extended to a countably additive measure on the [measure-theoretic completion](#) $\overline{\langle \mathcal{A} \rangle}$ of the (external) σ -algebra $\langle \mathcal{A} \rangle$ generated by \mathcal{A} .

In particular, if we consider the nonstandard normalised counting measure

$$E \mapsto \frac{\#E}{2N + 1}$$

on \mathcal{A} and take its standard part,

$$\mu : E \mapsto \text{st}\left(\frac{\#E}{2N + 1}\right)$$

this is a finitely additive probability measure on \mathcal{A} , and hence extends to a probability measure in $\mathcal{B} := \overline{\langle \mathcal{A} \rangle}$, which we will continue to call μ . This measure is known as the *Loeb measure* on the nonstandard set $\{-N, \dots, N\}$. Observe that any nonstandard subset of $\{-N, \dots, N\}$ of infinitesimal density will have Loeb measure zero. On the other hand, the set A had density at least δ , and so will have Loeb measure at least δ also.

Next, we define the shift $T : \{-N, \dots, N\} \rightarrow \{-N, \dots, N\}$ by $Tx := x + 1$, leaving T undefined for $x = N$. But observe that $\{N\}$ has an infinitesimal density, hence has Loeb measure zero. So T is defined almost everywhere, and is easily seen to be measurable and measure-preserving; it has an (almost everywhere defined) inverse that is also measurable and measure-preserving. Thus $\{-N, \dots, N\}$ with Loeb measure and the shift T becomes a measure-preserving system. Applying Theorem 16, we can thus find a standard $r > 0$ such that $A \cap T^r A \cap \dots \cap T^{(k-1)r} A$ has positive Loeb measure, so A contains a k -term arithmetic progression, a contradiction.

Remark 3 As stated, the measure space structure on $\{-N, \dots, N\}$ is not separable (i.e. countably generated) or regular (coming from a metric space). However, this can be fixed by restricting attention to the much smaller σ -algebra generated by A and its shifts (after dealing with the null sets on which T is not defined, e.g. by cutting out the $o(N)$ neighbourhood of $\{-N, N\}$). We omit the details.

— 4. The Szemerédi regularity lemma —

Finally, we use nonstandard analysis to give a proof of the [Szemerédi regularity lemma](#), which we phrase as follows:

Lemma 18 (Szemerédi regularity lemma) Let $G = (V, E)$ be a finite graph, and let $\epsilon > 0$. Then there exists a vertex partition $V = V_1 \cup \dots \cup V_m$ with $m \leq C(\epsilon)$ such that for all pairs (i, j) outside of a bad set S with $\sum_{(i,j) \in S} |V_i||V_j| \leq \epsilon|V|^2$, there exists a density d_{ij} such that

$$|E(A, B) - d_{ij}|V_i||V_j|| \leq \epsilon|V_i||V_j|$$

for all $A \subset V_i$ and $B \subset V_j$, where $E(A, B) := |E \cap (A \times B)|$, viewing E as a symmetric subset of $V \times V$.

Here we do not make the cells V_i in the partition of equal size, as is customary, but it is not too difficult to obtain this additional property from the above formulation of the lemma. The ability of nonstandard analysis to establish regularity lemmas was first observed by [Elek and Szegedy](#).

An application of the Bolzano-Weierstrass theorem for ultraproducts shows that this lemma is equivalent to the following nonstandard version:

Lemma 19 (Szemerédi regularity lemma, nonstandard formulation) Let $G = (V, E)$ be a nonstandard finite graph, and let $\epsilon > 0$. Then there exists a vertex partition $V = V_1 \cup \dots \cup V_m$, where m is a standard natural number and V_1, \dots, V_m are nonstandard subsets of V such that for all pairs (i, j) outside of a bad set S with $\sum_{(i,j) \in S} |V_i||V_j| \leq \epsilon|V|^2$, there exists a standard density d_{ij} such that

$$|E(A, B) - d_{ij}|V_i||V_j|| \leq \epsilon|V_i||V_j|$$

for all $A \subset V_i$ and $B \subset V_j$.

To see why Lemma 19 implies Lemma 18, suppose that Lemma 18 failed. Then there is a standard $\epsilon > 0$ such that for every standard m one could find a standard finite graph $G_m = (V_m, E_m)$ which could not be regularised into m or fewer cells as required by the lemma. Applying the Bolzano-Weierstrass theorem for ultraproducts, we can assume that G_m converges elementarily (relative to ϵ) to a limit $G = (V, E)$, which is then a nonstandard finite graph $G = (V, E)$ which cannot be regularised

into any standard finite number of cells. But this contradicts Lemma 19.

It remains to prove Lemma 19. Let $\mu_V : \mathcal{B}_V \rightarrow [0, 1]$ be Loeb measure on V (as constructed in the previous section), and $\mu_{V \times V} : \mathcal{B}_{V \times V} \rightarrow [0, 1]$ be Loeb measure on $V \times V$. It is easy to see that $\mathcal{B}_{V \times V}$ contains the product σ -algebra $\mathcal{B}_V \times \mathcal{B}_V$ as a subalgebra, and that the product measure $\mu_V \times \mu_V$ is the restriction of $\mu_{V \times V}$ to $\mathcal{B}_V \times \mathcal{B}_V$. The edge set E , viewed as a symmetric nonstandard subset of $V \times V$, is measurable in $\mathcal{B}_{V \times V}$, but is not necessarily measurable in $\mathcal{B}_V \times \mathcal{B}_V$. One can then form the conditional expectation $f := \mathbf{E}(1_E | \mathcal{B}_V \times \mathcal{B}_V)$, which is a $\mathcal{B}_V \times \mathcal{B}_V$ -measurable function that is defined up to $\mu_V \times \mu_V$ -almost everywhere equivalence, and takes values in $[0, 1]$.

The σ -algebra $\mathcal{B}_V \times \mathcal{B}_V$ is generated by product sets $A \times B$ of \mathcal{B}_V -measurable functions, which in turn can be approximated in measure to arbitrary accuracy by product sets of nonstandard sets. As f is $\mathcal{B}_V \times \mathcal{B}_V$ -measurable, we can approximate it to less than ϵ^2 in $L^1(\mu_V \times \mu_V)$ norm by a finite linear combination of indicator functions of products of nonstandard sets. Organising these products, we thus see that

$$\|f - \sum_{i=1}^m \sum_{j=1}^m d_{ij} 1_{V_i \times V_j}\|_{L^1(\mu_V \times \mu_V)} < \epsilon^2$$

for some finite partition of V into nonstandard sets V_1, \dots, V_m and some standard real numbers $d_{ij} \in [0, 1]$. By Markov's inequality, we thus see that

$$\|f - d_{ij}\|_{L^1(V_i \times V_j)} < \epsilon \mu_V(V_i) \mu_V(V_j)$$

for all (i, j) outside of a bad set S with

$$\sum_S \mu_V(V_i) \mu_V(V_j) \leq \epsilon.$$

Now let $A \subset V_i$ and $B \subset V_j$ be nonstandard sets, with (i, j) outside of S . Then

$$\|f 1_{A \times B} - d_{ij} 1_{A \times B}\|_{L^1(V_i \times V_j)} < \epsilon \mu_V(V_i) \mu_V(V_j).$$

On the other hand, $1_E - f$ is orthogonal to all $\mathcal{B}_V \times \mathcal{B}_V$ functions, and in particular to $1_{A \times B}$, and thus

$$\int_{V \times V} 1_E 1_{A \times B} d\mu_{V \times V} = \int_{V \times V} f 1_{A \times B} d\mu_{V \times V}.$$

Since

$$\int_{V \times V} 1_E 1_{A \times B} d\mu_{V \times V} = \frac{|E(A, B)|}{|V|^2}$$

and

$$\int_{V \times V} d_{ij} 1_{A \times B} d\mu_{V \times V} = d_{ij} \frac{|A||B|}{|V|^2}$$

and

$$\mu_V(V_i) = |V_i|/|V|; \mu_V(V_j) = |V_j|/|V|$$

we thus see that

$$\left| \frac{|E(A, B)|}{|V|^2} - d_{ij} \frac{|A||B|}{|V|^2} \right| < \epsilon \frac{|V_i||V_j|}{|V|^2}.$$

Thus G has been regularised using a finite number of cells, as required.

Share this: [Print](#) [Email](#)

Possibly related posts: (automatically generated)

- [Ultralimit analysis, and quantitative algebraic geometry](#)
- [Infinite fields, finite fields, and the Ax-Grothendieck theorem](#)
- [Radio Stations interrupted](#)
- [Abuse of Force](#)

Like Be the first to like this post.

36 comments

[Comments feed for this article](#)

[27 November, 2010 at 2:10 am](#)

none



Thanks for this great post, it demystifies a topic I've wondered about for while. There are several typos that are a bit hard for me to enter here since I can only have one large window on this screen. But it says R_M in a few places where it should say $R_{\mathfrak{U}}$, I think.

“(In fact, there is an enormous Galois group $\{\hbox{Gal}(\overline{\mathbf{C}}/\mathbf{Q})\}$, the action of which is completely undetectable with regards to absolute elementary convergence is completely unable to detect.)” is garbled.

One or two others that I remember but am not seeing now. I'll try to find them on a re-read.

[Corrected, thanks - T.]

0 0 Rate This

[Reply](#)

[27 November, 2010 at 3:18 am](#)

[gowers](#)



I'm very much enjoying the post. A very minor typo: in the fourth paragraph you meant to say slightly *less* well known.

[Corrected, thanks - T.]

0 0 Rate This

[Reply](#)

[27 November, 2010 at 7:04 am](#)

Anonymous



how do we know that $3, 3.1, 3.14, \dots$ is not converging to a rational number?

[Because π is known to be irrational, and limits in the real numbers are unique - T.]

0 0 Rate This

[Reply](#)

[28 November, 2010 at 6:30 am](#)

Greg Graviton



Concerning Furstenberg's recurrence theorem, I take it that you define the topology on X as follows: a basis of open sets is given by the nonstandard subsets of $[-N, N]$ (actually, these sets are also closed).

But: a measure preserving system is usually assumed to be second countable, whereas there are uncountably many nonstandard subsets of $[-N, N]$.

If I remember correctly, your usual remedy to this problem was to restrict attention to the sigma-algebra generated by the shifts of A .

0 0 Rate This

[Reply](#)

[28 November, 2010 at 9:28 am](#)

[Terence Tao](#)



This is correct: at present the measure space is neither separable nor metrisable, but as you say this can be fixed by cutting the sigma algebra down to size. Technically, the statement of the Furstenberg recurrence theorem is valid for arbitrary measure-preserving systems, though for technical reasons, the proof of the theorem usually proceeds by first reducing to the case of regular measure-preserving systems in which the sigma algebra can be identified with the Borel sigma algebra of a Polish space, in order to use tools such as disintegration. I've added a remark to indicate this.

0 0 Rate This

[Reply](#)

[29 November, 2010 at 12:00 pm](#)

[Concentration compactness via nonstandard analysis « What's new](#)



[...] in this previous blog post (and see also this later post on ultralimit analysis, as well as the most recent post on this topic). Very briefly, we will need to fix a non-principal ultrafilter on the natural numbers. Once one [...]

0 0 Rate This

[Reply](#)

[30 November, 2010 at 10:26 am](#)

Jonathan Weinstein



I had some trouble understanding your non-standard definition of uniform continuity (Exercise 3.) I tried to write down my own non-standard versions of continuity and uniform continuity: Let $I \subset {}^*\mathbf{R}$ be the set of non-negative infinitesimals (less than every positive standard real.) Then continuity is, in shortest form,

$$|x - y| \in I \Rightarrow |f(x) - f(y)| \in I$$

or in longer form

$$\forall \epsilon \in I : \forall x : \exists \delta \in I : |x - y| < \epsilon \Rightarrow |f(x) - f(y)| < \delta$$

where the epsilon and delta quantifiers come in a non-standard, but perhaps more intuitive order: "If x only moves a tiny bit, so does $f(x)$."

Now, if I did this right, uniform continuity would arise by switching the order of quantifiers, similarly to standard except the delta instead of epsilon quantifier gets switched with x :

$$\forall \epsilon \in I : \exists \delta \in I : \forall x : |x - y| < \epsilon \Rightarrow |f(x) - f(y)| < \delta$$

By contrast, I can't see where your definition distinguishes the order of quantifiers. I understand $f(y) = f(x) + o(1)$ to be shorthand for $|f(y) - f(x)| \in I$ but I feel like this shorthand is obscuring what the infinitesimal is allowed to depend on. Assuming your intended meaning is correct, maybe there is just some convention here I don't know; I have read about non-standard analysis but never worked with it. I will highly appreciate hearing about any mistakes I made here.

0 0 Rate This

[Reply](#)

[30 November, 2010 at 10:39 am](#)

[Terence Tao](#)



In general, nonstandard analysis largely eliminates the need to carefully order one's quantifiers as is the case in standard analysis, but in return, one must now carefully distinguish between what objects are standard and what objects are nonstandard.

Continuity is equivalent to the property that $|x - y| \in I \implies |f(x) - f(y)| \in I$ for all *standard* reals $x \in \mathbf{R}$ and non-standard reals $y \in {}^*\mathbf{R}$. Uniform continuity is equivalent to the same property $|x - y| \in I \implies |f(x) - f(y)| \in I$, but now where $x \in {}^*\mathbf{R}$ and $y \in {}^*\mathbf{R}$ are *both* permitted to be non-standard.

It may help to think of a nonstandard real x as an (ultra)limit of a sequence of standard reals x_n , with a non-negative infinitesimal in I being the (ultra)limit of a sequence of non-negative standard reals that converges (in the standard sense) to zero. In standard analysis, continuity is equivalent to the claim that if y_n is a sequence of (standard) real numbers and x is a (standard) real number such that $|x - y_n| \rightarrow 0$ as $n \rightarrow \infty$, then $|f(x) - f(y_n)| \rightarrow 0$ as $n \rightarrow \infty$. Uniform continuity is equivalent to the claim that if x_n, y_n are two sequences of (standard) real numbers such that $|x_n - y_n| \rightarrow 0$ as $n \rightarrow \infty$, then $|f(x_n) - f(y_n)| \rightarrow 0$ as $n \rightarrow \infty$. Viewed in this way, we see that the nonstandard definitions of continuity and uniform continuity are basically the ultralimits of the standard definitions.

More generally, a rough first approximation to nonstandard analysis can be obtained by introducing an additional parameter n , and interpreting "standard" as "independent of n " and "nonstandard" as "allowed to depend on n ". (Similarly, "infinitesimal" becomes "going to zero as $n \rightarrow \infty$ ", and so forth.) This is only a rough approximation because one also needs an ultrafilter to decide what sets of n are more "important" than others in order to decide whether a given nonstandard object has a given property. (e.g. is the quantity $(-1)^n$ positive or negative in the nonstandard universe? It depends on whether the even numbers or the odd numbers are contained in the ultrafilter.)

In any case, it is an excellent exercise to try to formally establish the equivalence of the standard and nonstandard formulations of both concepts from first principles (i.e. using the definition of ultralimit); using underflow or countable saturation is the slickest way to proceed, but it is instructive to do it “with bare hands” as well, to see how the quantifiers are being automatically absorbed by the ultralimit process.

2 0 Rate This

[Reply](#)

[30 November, 2010 at 11:46 am](#)

Jonathan Weinstein



Thanks, that's extremely helpful! Part of the reason my intuition was messed up was that I forgot non-standard reals can be infinite as well as infinitesimal. So I didn't notice the key idea that a prototypical non-uniformly continuous function like $\exp(x)$ is actually not even continuous at infinite non-standard reals. Likewise, I see that $f(x)=1/x$ restricted to $(0,1]$ is continuous at standard reals but discontinuous at infinitesimals.

I look forward to trying to read the rest of the article...

0 0 Rate This

[Reply](#)

[30 November, 2010 at 11:54 am](#)

Terence Tao



Yes, this is correct. On a compact (metric space) domain, every nonstandard point is infinitesimally close to a standard point (this is basically the Heine-Borel theorem), and so uniform continuity and continuity become the same concept on such domains (as is already well known in standard analysis); but on non-compact domains this is no longer the case.

(One little subtlety though: while functions such as $1/x$ and $\exp(x)$ are not uniformly continuous in the sense that $|x - y| \in I \implies |f(x) - f(y)| \in I$ holds for all nonstandard x, y , they are still continuous in the *internal* sense, in that for every nonstandard $\varepsilon > 0$ and nonstandard x there exists a nonstandard $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever y is nonstandard with $|x - y| \leq \delta$. Indeed, from the transfer principle, any continuous function f will become internally continuous when transferred to the nonstandard world. But this notion of continuity is distinct from the more “external” notion of continuity that involves the external concept of an infinitesimal.)

0 0 Rate This

[Reply](#)

[2 December, 2010 at 11:07 am](#)

katz



Thanks for a great post. There is a small detail I don't follow: you require countability of the `_constants_` which appears to make the construction inapplicable to \mathbb{R} , an important case :)

0 0 Rate This

[Reply](#)[2 December, 2010 at 11:23 am](#)**[Terence Tao](#)**

Unfortunately I believe a restriction of this form is necessary in order to get elementary completeness (which, once one allows uncountably many constants, is subtly different from countable saturation, which does indeed hold with arbitrary constants). For instance, consider the sequence $x_n = n$ in the hyperreals ${}^*\mathbf{R}$. With at most countably many constants, one can extract a subsequence x_{n_j} that is elementarily convergent to some (unbounded) hyperreal limit x . But if one allows every single hyperreal as a constant, then it is impossible for any subsequence x_{n_j} to converge elementarily to a limit x , because the predicate $P(y) := (y = x)$ (which uses x as a constant) would be false for each of the x_{n_j} but true for x .

(I do not know, though, if the hyperreals are elementarily complete if one allows all the standard reals (but not the hyperreals) as constants.)

0 0 Rate This

[Reply](#)[2 December, 2010 at 7:38 pm](#)**katz**

Can one view the discussion of elementary completeness as taking place with \mathbf{Q}^* as the backdrop, or must the constants be left unspecified and dependent on the application? Incidentally, as per the discussion of continuity versus uniform continuity, there is a parallel distinction between convergence and uniform convergence which happens to be the background necessary to understand Cauchy's 1853 paper on the sum theorem.

0 0 Rate This

[Reply](#)[3 December, 2010 at 1:20 pm](#)**[John Rood](#)**

Can you say anything useful about nilpotent infinitesimals, aka differential forms etc?

http://en.wikipedia.org/wiki/Differential_%28infinitesimal%29

Maybe these are simply outside your development here?

0 0 Rate This

[Reply](#)[3 December, 2010 at 1:35 pm](#)**[Terence Tao](#)**

One can certainly use the hyperreals to construct rings with nilpotent infinitesimals via a quotient

space construction. For instance, starting with a nonstandard infinitesimal dx in the hyperreals ${}^*\mathbf{R}$, we can construct the space $o(dx)$ of hyperreals of magnitude bounded by εdx for every standard $\varepsilon > 0$. This is an ideal of $O(\mathbf{R})$, the space of bounded hyperreals, and so one can form the quotient space $\mathbf{R}[dx] := O(\mathbf{R})/o(dx)$, which is a commutative ring containing the standard reals as a subring for which $dx^2 = 0$. Any standard differentiable function $f : \mathbf{R} \rightarrow \mathbf{R}$ extends to a function $f : \mathbf{R}[dx] \rightarrow \mathbf{R}[dx]$, obeying the Newton approximation $f(x + tdx) = f(x) + tf'(x)dx$ for any standard x, t . One can use this as a rigorous basis for differential infinitesimal calculus if desired.

0 0 Rate This

[Reply](#)

[3 December, 2010 at 2:46 pm](#)

[John Rood](#)



Hmmm. Very good. But let me stick my neck out a bit farther.

Are we then talking about something like embedding the exterior algebra into some multivariable version of the nonstandard reals?

One point I wish to make is that (cf the wikipedia link in my previous post) the standard topos models my involve intuitionistic logic in which, say, the law of the excluded middle fails. I presume you don't have such models in mind ... ?

0 0 Rate This

[Reply](#)

[5 December, 2010 at 10:03 am](#)

katz



John: the hyperreals are commutative, so you won't get very far with the exterior algebra here. Also, the hyperreals are built in the context of ZFC with classical logic, and in this sense is conservative. Lawvere's approach, while it has its advantages, necessitates jettisoning traditional set theory as well as classical logic, but results in a surprisingly "clean" presentation of certain topics whose classical treatment is more cumbersome.

0 0 Rate This

[Reply](#)

[6 December, 2010 at 6:41 am](#)

[John Rood](#)



Thank you for this reply. Of course they're commutative, sorry.

0 0 Rate This

[Reply](#)

[6 December, 2010 at 7:20 am](#)

[John Rood](#)



Dear katz, thinking about this a bit more, I find I still have a question. But, this leads me to a question of “protocol” on this blog, namely, is there some way I can take this conversation out of this context, say by contacting you directly by email or something. Or you can contact me (my email address is known to the blog, of course).

It occurs to me that one should (?) (or might?—where does the following go wrong) have vectors over the nonstandard reals, then vector spaces, then algebras like the tensor algebra, Lie algebras AND the exterior algebra. All of these things would be carrying along with them “little differentials”, i.e. infinitesimals. Does this make any sense? Then one might hope to recover the “Lie Bracket” by a standard geometric argument (about going around the sides of an “infinitesimal” parallelogram which doesn’t commute) which produces a new vector (field) from two input vector (fields) as a sort of first order approximation (infinitesimally) of this geometric non-commutativity.

As I say, I’d be interested to hear what you have to say / know about this, and where it goes wrong ... ?

0 0 Rate This

[Reply](#)

[6 December, 2010 at 7:53 am](#)

katz



I think all of the items you mentioned should go through without problem. The hyperreal approach to Lie groups and algebras is already in Robinson I think. Actually I was unable to find your email address via this site. You can contact me by looking for first name Mikhail. I prefer not to leave an email address here so as to minimize robot-generated mail. Otherwise we can continue here; I am sure Terry will let us know if we overuse his hospitality :)

0 0 Rate This

[Reply](#)

[6 December, 2010 at 10:43 am](#)

[John Rood](#)



Ok, tyvm. Let me mention that I do now have your email address with a tip o’ the hat to Terry and I presume that you also have mine.

So (and please forgive me if I am making superficial errors here) the “obvious” question is: If all this Lie Theory goes through using Robinson’s nonstandard analysis, what does one get from Lawvere’s approach? I am aware that Lie Groups are about as well behaved as one can expect any spaces to be. Maybe more general spaces ... bla bla bla ... ?

Is there a simple answer to this question of what is Lawvere up to? (Even a non-simple answer might be of interest ...)

0 0 Rate This

[Reply](#)

[6 December, 2010 at 10:51 am](#)

katz

Lawvere's approach allows for a certain additional elegance (some would say, simplicity) due to the presence of nilsquare infinitesimals. To take an elementary example, if $y=f(x)$ is differentiable then for nilsquare dx one would have $dy=f'(x)dx$ without a higher order error term. Similarly, one can define the tangent plane as everything proportional to nilsquare infinitesimal displacements. A planetary orbit can be thought of literally as made up of infinitesimal straight line segments.

As I see it, the disadvantage of the approach is an absence of the full power of the transfer principle which is available in the hyperreal framework. I hereby challenge John L. Bell to produce an infinitesimal proof of the invariant subspace conjecture :)

0 0 Rate This

[Reply](#)[6 December, 2010 at 3:00 pm](#)**[John Rood](#)**

Ok, very good. But I find I still have a point of confusion. Are all infinitesimals' squares zero? Do Leibniz' higher order infinitesimals make sense? Which kind is dx ?

0 0 Rate This

[Reply](#)[6 December, 2010 at 7:25 pm](#)**katz**

No, some infinitesimals are invertible and therefore not nilsquare. To have $dy=f'(x)dx$ we must start with a nilsquare dx . Leibniz would not hear of nilsquare ones. Rather, they are in the spirit of his critic Nieuwentijdt. A great account of all this is in John L. Bell's article in the Stanford Encyclopedia of Philosophy.

0 0 Rate This

[Reply](#)[7 December, 2010 at 5:55 am](#)**[John Rood](#)**

Yes! Very nice way into this topic, apparently:

<http://plato.stanford.edu/entries/continuity/>

Here is a quotation from the section about Smooth Infinitesimal Analysis:

for all infinitesimal ε , not $\varepsilon \neq 0$ [endquote]

Evidently not compatible with the Law of the Excluded Middle ...

Amazing how the “intuitive” history of mathematics (analysis) can be formalized! And the internet makes all of this easier to access.

katz, thank you very much for your discussion here.

0 0 Rate This
[Reply](#)

[14 December, 2010 at 11:43 am](#)

none



So I wonder to what extent this nonstandard stuff breaks the “standard” proofs in calculus, that depend on the least-upper-bound axiom, which fails in this non-Archimedean version of the reals. The nonstandard reals are elementarily equivalent to the standard reals, i.e. they satisfy the same first-order sentences, but the least-upper-bound property is second-order and can tell the two models apart. I didn't know any logic when I took calculus but I guess this means that my “rigorous” calculus course was actually done in second-order logic, which means(?) that its proofs relied on a tacit, informal version of set theory under the covers. That gap got taken care of later in logic class, where we went over an explicit encoding of the reals in ZFC, which meant that the calculus proofs worked because we could treat second-order statements about reals as first-order statements about sets. But to develop calculus with nonstandard reals, it looks like you have to do all the proofs over again; you don't get to add new, easier proofs while still keeping the old ones. Does that look right? I've never really thought about this before.

0 0 Rate This
[Reply](#)
[15 December, 2010 at 10:05 pm](#)

katz



In response to the comment by “none”: second order statements remain valid if interpreted as applying to internal sets, in particular natural extensions of real sets. Thus, the least upper bound will exist, etc. For instance, the natural extension of the open interval $(0,1)$ will contain all decimal expansions starting with infinitely many 9s, and the least upper bound is 1.

0 0 Rate This
[Reply](#)

[15 December, 2010 at 11:57 am](#)

Albanius



It would be very helpful to see this as a pdf.

My attempts to print from this page or copy to MS-Word did not capture the symbols.

0 0 Rate This
[Reply](#)
[15 December, 2010 at 12:23 pm](#)

Jonathan Weinstein

Albanius,

Prof. Tao provides valuable, informative and entertaining content, with higher quality than most published material, for free. I would hate to complain about the file format.

0 0 Rate This

[Reply](#)

[15 December, 2010 at 1:03 pm](#)

Albanius

It is precisely because the content is so valuable that I would like to be able to print it out to study.

0 0 Rate This

[Reply](#)

[15 December, 2010 at 4:00 pm](#)

Anonymous

Albanius, the AMS has been publish Prof. Tao's in printed book form once per year. See for example here: <http://www.ams.org/bookstore-getitem/item=MBK-59>

0 0 Rate This

[Reply](#)

[18 December, 2010 at 10:10 am](#)

anonymous2

In Firefox choose File > Print and it formats the blog post nicely and includes the "special" characters (i.e. TeX). WordPress does that by default.

0 0 Rate This

[Reply](#)

[18 December, 2010 at 10:24 am](#)

Albanius

Newbie mistake, sorry.

After clicking the PRINT button between the essay and the comments, the next window had an option for PDF, which created a pdf file I could save and print, just what I wanted.

0 0 Rate This

[16 December, 2010 at 9:44 am](#)

alex



Terry Tao thank you for all the beautiful mathematics, I am learning a lot reading your blog.

You said that the model of real numbers for the language of ordered fields can see algebraic numbers precisely but see transcendental numbers fuzzily. I was wondering if anyone has applied this method to prove/construct a number transcendental?

0 0 Rate This

[Reply](#)

[17 December, 2010 at 7:37 pm](#)

Quinn Culver



Typo in the proof of Proposition 4?:

“...is elementarily convergent, as desired.” Should be “is elementarily Cauchy, as desired.”?

[Corrected, thanks - T.]

0 0 Rate This

[Reply](#)

[17 December, 2010 at 8:20 pm](#)

Quinn Culver



Another potential typo: in Definition 6, “...for all n sufficiently close to p .” should be “...for all α sufficiently close to α_∞ .”

(If you'd prefer I not point out minor typos like this, let me know.)

[Corrected, thanks - T.]

0 0 Rate This

[Reply](#)