Hello list,

I am looking for references on the following facts about interpreting some sentences involving infinitesimals in filter-powers instead of in ultrapowers...

Notation:
\I is an index set;
\F is a filter on \I;
\U is an ultrafilter on \I, possibly extending \F;
if X is a topological space and x is a point of X, then
\V_x is the filter of neighborhoods of x in X;
Set^\F is a shorthand for the filter-power (Set^\I)/\F;
Set^\U is a shorthand for the ultrapower (Set^\I)/\U;

Terminology:
Set is the "standard universe";
Set^\U is the "non-standard universe";
Set^\F is the "semi-standard universe";
points of Set^\I are called "sequences", or "pre-hyperpoints";
points of Set^\F or Set^\U are called "hyperpoints".

Well, the facts. Here they are:

(1) For any (standard) function f:X->Y from a topological space to another, and for any standard point x in X, the following three statements are equivalent:

(a) f is continuous at x;
(b) for all choices of a triple (\I, \F, x_1), where \I is an index set, \F is a filter on \I, and x_1 is a hyperpoint infinitely close to x, then f(x_1) is infinitely close to f(x);
(c) for the "natural infinitesimal" (\I, \F, x_1) := (X, \V_x, id), the hyperpoint f(x_1) is infinitely close to f(x).

(2) Any filter-infinitesimal (\I, \F, x_1) infinitely close to x factors through the natural infinitesimal (X, \V_x, id) in a unique way.

(3) We can use these ideas to lift proofs done in a certain "strictly calculational fragment" of the language of non-standard analysis to constructions done in a filter-power; and then, if we replace the free variables that stand for infinitesimals in our formulas by natural infinitesimals, we get (by (c)<=>(a) in (1)) a translation of our proof with infinitesimals to a standard proof, in terms of limits and continuity.

On the one hand, I have never seen anything published about filter-infinitesimals, and it took me a long time to find the right formulations for this... on the other hand, ideas similar to these seem to be implicit in many places (see the last sections of the PDF). However, my guess is that at least parts of (3) are new.

Natural Infinitesimals in Filter-Powers

Eduardo Ochs
eduardoochs@gmail.com
http://angg.twu.net/

Preliminary version - 2008jul13

Proofs in non-standard analysis are done by moving back and forth between two universes, the standard universe, **Set**, and the non-standard universe, $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$. Here we will work also with two other universes between **Set** and $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$: $\mathbf{Set}^{\mathbb{I}}$, the universe of ($\mathbb{I}$-indexed) sequences, and $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$, the "semi-standard universe", in which sequences are identified when the set of indices where they coincide is "$\mathcal{F}$-big".

Intuitively, the filter-power $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ is a generalization of $\mathbf{Set}^{\mathbb{N}}/\mathcal{N}$, where $\mathcal{N}$ is the filter of cofinite sets of naturals. Let's use the prefixes "pre-hyper-" and "hyper-" to refer to elements of a $\mathbf{Set}^{\mathbb{I}}$ and of a $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ respectively. Sequences of reals tending to zero are pre-hyperreals (in $\mathbf{Set}^{\mathbb{N}}$) whose corresponding hyperreals in $\mathbf{Set}^{\mathbb{N}}/\mathcal{N}$ — or in a $\mathbf{Set}^{\mathbb{N}}/\mathbb{N}$, where U is a ultrafilter extending $\mathcal{N}$ — behave as infinitesimals.

In a certain sense, ultrapowers are much more well-behaved than filter-powers: the logic of a $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ is two-valued, and we have certain "transfer theorems" that transfer truths from **Set** to a $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ and back. The logic of a non-trivial filter-power, however, is boolean but not two-valued; what we will show here is that certain "purely calculational" proofs involving infinitesimals can be lifted through the quotient $\mathbf{Set}^{\mathbb{I}}/\mathcal{F} \to \mathbf{Set}^{\mathbb{I}}/\mathcal{U}$, yielding proofs in a filter-power $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ that can be immediately reinterpreted as being standard proofs in disguise; "infinitesimality" becomes "continuity".

Important: *I don't know how much of this is new.* This preliminary version has two main intents: (1) to request feedback and pointers to the literature from people who know the subject infinitely more than me, and (2) to give some elementary motivation for topos theory to the regular attendants of the local Logic seminar — the logic of a filter-power, being boolean, is much easier to grasp than the one of a non-boolean topos. Intent (2) made me write this in a very elementary way; I apologize in advance to the "(1)" people for the length of the text, and for the obviousness of some parts.

1 A very quick introduction to filters

A *filter* $\mathcal{F}$ on an index set $\mathbb{I}$ is a family of subsets of $\mathbb{I}$, $\mathcal{F} \subset \mathcal{P}(\mathbb{I})$, such that:

- $\mathbb{I} \in \mathcal{F}$;
- if $I, I' \in \mathcal{F}$ and $I \cap I' \subseteq I'' \subseteq \mathbb{I}$ then $I'' \in \mathcal{F}$.

These two conditions are exactly what it is needed to make

$$(a_i) \sim (b_i) \quad \stackrel{\text{def}}{\Leftrightarrow} \quad \{\, i \in \mathbb{I} \mid a_i = b_i \,\} \in \mathcal{F}$$

an equivalence relation: we want $(a_i) \sim (a_i)$, so $\mathbb{I} \in \mathcal{F}$; and if (a_i) and (b_i) coincide when $i \in I$, and (b_i) and (c_i) coincide when $i \in I'$, then (a_i) and (c_i) coincide at some set of indices $I'' \supset I \cap I'$, so we need the second condition to make '$\sim$' transitive.

This will be our archetypical filter (the "filter of cofinites"):

$$\mathcal{N} := \{\, I \subset \mathbb{N} \mid I \text{ is cofinite (i.e., } \mathbb{N} \backslash I \text{ is finite)} \,\}.$$

We will also need "filters of neighborhoods", "filters of punctured neigborhoods", and "filters of strictly punctured neighborhoods". Fix a topological space X and a point $x \in X$; an *open neighborhood* of x in X is an open set $U \subseteq X$ containg x; a *neighborhood* of x in X is a set $V \subseteq X$ containing some open neighborhood of x; a *punctured neighborhood* of x is a set $V \subset X$ such that $V \cup \{x\}$ is a neighborhood of x; and a *strictly punctured neighborhood* of x is a set $V \subseteq X \backslash \{x\}$ such that $V \cup \{x\}$ is a neighborhood of x.

These filters will also be useful later: if $x \in X$ is a point in a topological space $(X, \mathcal{O}(X))$, then:

$$\begin{array}{rcl} \mathcal{V}_x & := & \{\, V \mid \exists U \in \mathcal{O}(X).\, x \in U \subseteq V \subseteq X \,\} \\ \mathcal{V}_x^- & := & \{\, V \mid \exists U \in \mathcal{O}(X).\, x \in U \subseteq (V \cup \{\, x \,\}) \subseteq X \,\} \end{array}$$

$\mathcal{V}_x$ is the *filter of neighborhoods* of x; $\mathcal{V}_x^-$ is the *filter of punctured neighborhoods* of x. Sometimes "$\mathcal{V}_x^-$" will stand for a filter on X, sometimes for a filter on $X \backslash \{\, x \,\}$.

(By the way: $\mathcal{N}$ is the filter of punctured neighborhoods of ∞ in the one-point compactification of $\mathbb{N}$.)

Let's define two operations on families of subsets of $\mathbb{I}$:

$$\begin{array}{rcl} \uparrow\!\mathcal{A} & := & \{\, A' \subseteq \mathbb{I} \mid \exists A \in \mathcal{A}. A \subseteq A' \subseteq \mathbb{I} \,\} \\ \bigcap_{\text{fin}} \mathcal{A} & := & \{\, A_1 \cap \ldots \cap A_n \mid n \in \mathbb{N}, A_i \in \mathcal{A} \,\} \end{array}$$

In $\bigcap_{\text{fin}} \mathcal{A}$ we consider that when $n = 0$ the intersection "$A_1 \cap \ldots \cap A_n$" is $\mathbb{I}$.

Fact: for any family $\mathcal{A}$ of subsets of $\mathbb{I}$, the family $\bigcap_{\text{fin}} \uparrow\!\mathcal{A}$ (that is equal to $\uparrow \bigcap_{\text{fin}} \mathcal{A}$) is a filter on $\mathbb{I}$ — the "filter generated by $\mathcal{A}$".

Let's call a filter $\mathcal{F}$ on $\mathbb{I}$ *bad* when $\mathcal{F} = \mathcal{P}(\mathbb{I})$.

A non-bad filter $\mathcal{F}$ divides the sets in $\mathcal{P}(\mathbb{I})$ in three classes:

- The *$\mathcal{F}$-big* sets are the ones in $\mathcal{F}$;

- The $\mathcal{F}$-*small* sets sets are the ones whose complements are $\mathcal{F}$-big;
- The $\mathcal{F}$-*medium* sets are the other sets in $\mathcal{P}(\mathbb{I})$ - the ones that are neither $\mathcal{F}$-big nor $\mathcal{F}$-small.

The cofinite subsets of $\mathbb{N}$ are $\mathcal{N}$-big; the finite subsets are $\mathcal{N}$-small; and for any $k \in \{ 2, 3, 4, \dots \}$ the set of multiples of k ("$k\mathbb{N}$") is $\mathcal{N}$-medium.

An *ultrafilter* $\mathcal{U}$ on $\mathbb{I}$ is a (non-bad) filter on $\mathbb{I}$ that divides the subsets of $\mathbb{I}$ in just $\mathcal{U}$-big and $\mathcal{U}$-small subsets; there are no $\mathcal{U}$-medium subsets.

We will reserve the notation '$\mathcal{U}$' for ultrafilters.

A non-bad filter $\mathcal{F}$ on $\mathbb{I}$ is said to be *principal* when $\bigcap \mathcal{F} \neq \emptyset$. For each non-empty set $I \subset \mathbb{I}$, $\uparrow\{I\}$ is a principal filter; $\uparrow\{I\}$ is an ultrafilter iff I has exactly one element. $\mathcal{N}$ is not principal. Filters and ultrafilters of the form $\uparrow\{I\}$ induce trivial equivalence relations; we are more interested in non-principal filters and ultrafilters.

All naïve attempts to construct non-principal ultrafilters explicitly — say, by enlarging filters until no more medium sets are left — are bound to fail. The following argument does not *prove* this (which is very hard; reference?) but it is easy enough, and quite enlightening.

> Take a denumerable set $\mathcal{A} = \{ A_1, A_2, A_s, \dots \}$ of generators; form the sequence $A'_1 := A_1$, $A'_2 := A_1 \cap A_2$, $A'_3 := A_1 \cap A_2 \cap A_3$, ..., and then form the sequence $A''_1, A''_2, A''_3, \dots$ by removing the repetitions from $A'_1, A'_2, A'_3, \dots$. Clearly, $\uparrow \bigcap_{\text{fin}} \mathcal{A} = \uparrow \bigcap_{\text{fin}} \mathcal{A}' = \uparrow\mathcal{A}' = \uparrow\mathcal{A}''$, where $\mathcal{A}'' := \{ A''_1, A''_2, \dots \}$; if $\mathcal{A}''$ is finite, $\mathcal{A}'' = \{ A''_1, \dots, A''_n \}$, then $\uparrow\mathcal{A}'' = \uparrow\{ A''_n \}$; if $A''_n = \emptyset$ then $\uparrow\mathcal{A}''$ is bad, and if $A''_n \neq \emptyset$ then $\uparrow\mathcal{A}''$ is principal.
>
> If $\mathcal{A}''$ is infinite, take the sequence of "differences" $D_1 = A''_1 \backslash A''_2$, $D_2 = A''_2 \backslash A''_3$, ...; both $D_1 \cup D_3 \cup D_5 \cup \dots$ and $D_2 \cup D_4 \cup D_6 \cup \dots$ are $\uparrow\mathcal{A}''$-medium sets, and so $\uparrow\mathcal{A}''$ is not an ultrafilter.

2 A very quick introduction to NSA

More terminology:

Set is the *standard universe*.
$\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ is the *non-standard universe*.
$\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ (here $\mathcal{F}$ can be just a filter) is the *semi-standard universe*.
A *pre-hyperpoint* is a point of $\mathbf{Set}^{\mathbb{I}}$, i.e., an $\mathbb{I}$-indexed sequence.
A *hyperpoint* is a point of $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ or $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$, i.e., the image of a sequence by the quotients '$/\mathcal{F}$' or '$/\mathcal{U}$' — i.e., an equivalence class of sequences.
A *standard element* of $\mathbf{Set}^{\mathbb{I}}$ or $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ is a constant sequence (modulo the equivalence relation, maybe).

The obvious maps

$$\begin{array}{ccccc} \mathbf{Set} & \longrightarrow & \mathbf{Set}^{\mathbb{I}} & \longrightarrow & \mathbf{Set}^{\mathbb{I}}/\mathcal{F} \\ & & & \searrow & \vdots \\ & & & & \mathbf{Set}^{\mathbb{I}}/\mathcal{U} \end{array}$$

will not usually be named. When $\mathcal{F} \subset \mathcal{U}$ the map $\mathbf{Set}^{\mathbb{I}}/\mathcal{F} \to \mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ exists; it takes each equivalence class of sequences into a bigger equivalence class.

Now take this (pre-)hyperreal: $\epsilon := (1, \frac{1}{2}, \frac{1}{3}, \ldots)$. For any standard interval containing 0, say, $U := (-\frac{1}{4}, \frac{1}{4})$, the sentence "$\epsilon \in U$" is true in $\mathbf{Set}^{\mathbb{N}}/\mathcal{N}$: the sequence of truth-values

$$(1 \in U, \tfrac{1}{2} \in U, \tfrac{1}{3} \in U, \ldots)$$

is false for the first indices, but true from some point on, and so it coincides with $(\top, \top, \top, \ldots)$ in a big set of indices.

We will say that a pre-hyperpoint x' is *infinitely close* to a standard point x when for any standard open set U containing x the sentence "$x' \in U$" is true in a big set of indices.

Following these ideas, $\epsilon := (1, \frac{1}{2}, \frac{1}{3}, \ldots)$ is infinitely close to 0 — but $\epsilon > 0$.

Typical proofs in NSA work a part of the time in **Set** and part in a $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$, using certain "transfer theorems" to "transfer truths" between the standard and the non-standard universe; the details are complex, and not relevant now. This observation, however, is crucial:

> In a $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ every ($\mathbb{I}$-indexed) sequence of truth-values is either equivalent to 'true' or to 'false'; but in a $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ we may have more truth-values. For example, $(\top, \bot, \top, \bot, \top, \bot, \ldots)$ only coincides with $(\top, \top, \top, \ldots)$ and with $(\bot, \bot, \bot, \ldots)$ in medium sets of indices.

So, in a '$\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$' we have transfer theorems, and a two-valued logic. Everything seems to indicate that '$\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$'s are better than '$\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$'s, but...

3 Ultrafilters are evil (in a sense)

The existence of non-principal ultrafilters is a consequence of the Axiom of Choice, and it is independent of ZF, but the proof of this is quite hard; see [Hal64]. What matters here is: *ultrafilters are a source of non-intuitiveness in the semantics*; proofs done in $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ may be hard to translate to standard proofs in **Set**. On the other hand, proofs done in a $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ — especially proofs in $\mathbf{Set}^{\mathbb{N}}/\mathcal{N}$ — are easy to translate; when the filters are explicilty presented, they are standard proofs in disguise.

One of the reasons why NSA never gained very wide acceptance was because it has been proved that by using NSA one "cannot prove anthing new" (this is only true for some

One of the reasons why NSA never gained very wide acceptance was because it was believed that by using NSA one "cannot prove anthing new" (actually this is only true for some classes of formulas; see [HK86]). Proofs using infinitesimals may be clearer and shorter, but infinitesimals are like objects that we are not supposed to play with (because they are "unnecessary"?!), and we may need to hide them quickly when the grown-ups approach... From this point of view, one class of proofs in $\mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ is especially interesting: the ones that "lift" to a filter-power $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$, from where they can be translated quickly to standard proofs.

I don't know how to characterize the full class of "liftable" proofs yet — but it seems that "purely calculational" proofs can be lifted. I am going to give a definition of "purely calculational" proofs — that may be overly restrictive —, and show how to lift and translate proofs of that form. We will focus on a single example, that seems to be rich enough:

4 A proof with infinitesimals

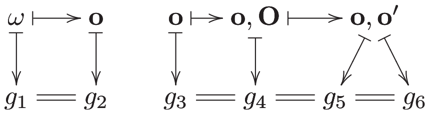
Definition (tentative, and sketchy): a *purely calculational proof* with infinitesimals is one made of a series of steps of the forms:
"$\forall x_1 \sim x_0.f(x_1) = g(x_1)$" or
"$\forall x_1 \sim x_0.\exists! y_1 \sim y_0.f(x_1) = h(x_1, y_1)$",
where x_0, y_0, f, and g are standard.

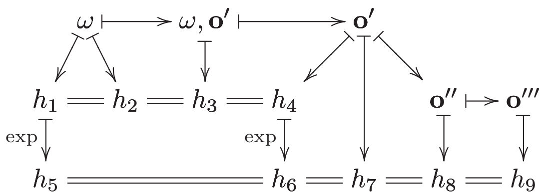
These steps can be composed in several ways. Let's look at an example.

Suppose that we want to prove that $\lim_{n\to+\infty}(1+\frac{a}{n})^n = e^a$; it is enough to prove that for any infinitely big natural number ω, $(1+\frac{a}{\omega})^\omega \sim e^a$. The calculations are the ones below, at the left; the right side shows some abbreviations. Note that we write just "g_1" for "$g_1(\omega)$", "h_3" for "$h_3(\omega, \mathbf{o}')$", etc.

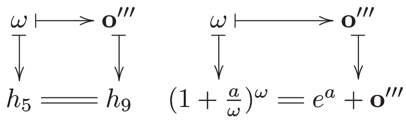
At some steps new symbols — $\mathbf{o}, \mathbf{o}', \mathbf{o}''$ — are introduced. Their names ("little 'o'"s) imply that they are infinitesimals, and there are implicit quantifiers: "there is a unique value for $\mathbf{o}$ (or $\mathbf{o}'$, or $\mathbf{o}''$) here making the equality hold". At g_4 we introduce an '$\mathbf{O}$', that stands for a "finite hyperreal". Readers who are not familiar with this concept ([SL76], sec.4.4.1) should just skip this step.

$$
\begin{array}{rclrcl}
\frac{a}{\omega} & = & \mathbf{o} & g_1 & = & g_2 \\
f(b+\mathbf{o}) & = & f(b) + f'(b)\,\mathbf{o} + \mathbf{O}\mathbf{o}^2 & g_3 & = & g_4 \\
 & = & f(b) + f'(b)\,\mathbf{o} + \mathbf{o}'\mathbf{o} & & = & g_5 \\
 & = & f(b) + (f'(b) + \mathbf{o}')\mathbf{o} & & = & g_6 \\
\log(1+\mathbf{o}) & = & (1+\mathbf{o}')\mathbf{o} & g_7 & = & g_8 \\
\log\,(1+\frac{a}{\omega})^{\omega} & = & \omega\ \log\,(1+\frac{a}{\omega}) & h_1 & = & h_2 \\
 & = & \omega\ ((1+\mathbf{o}')\ \frac{a}{\omega}) & & = & h_3 \\
 & = & (1+\mathbf{o}')\ a & & = & h_4 \\
(1+\frac{a}{\omega})^{\omega} & = & e^{((1+\mathbf{o}')\ a)} & h_5 & = & h_6 \\
 & = & e^{(a+\mathbf{o}'a)} & & = & h_7 \\
 & = & e^{(a+\mathbf{o}'')} & & = & h_8 \\
 & = & e^{a} + \mathbf{o}'''. & & = & h_9
\end{array}
$$





When we compose all cells we get this:

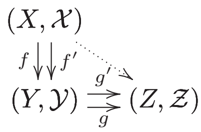


We will see how to lift this proof to a standard proof.

5 Filtered spaces

Definitions: a *filtered space* is a pair $(X, \mathcal{X})$ where X is a set and $\mathcal{X}$ is a filter over X; a *function* from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ is a function from an $\mathcal{X}$-big subset of X to Y; a *total function* from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$ is a function defined on the whole set X. We say that a function $f : (X, \mathcal{X}) \to (Y, \mathcal{Y})$ is *continuous* when the inverse image by f of each $\mathcal{Y}$-big set is an $\mathcal{X}$-big set, and we say that two functions $f, g : (X, \mathcal{X})$ to $(Y, \mathcal{Y})$ are *equivalent* when they coincide on some $\mathcal{X}$-big set.

If $f, f' : (X, \mathcal{X}) \to (Y, \mathcal{Y})$ are equivalent and continuous, and $g, g' : (Y, \mathcal{Y}) \to (Z, \mathcal{Z})$ are also equivalent and continuous, then $(f; g)$ and $(f'; g')$ are also equivalent and continuous.



A topological space $(X, \mathcal{O}(X))$ with a chosen point $x_0 \in X$ has a natural filtered space structure: take $\mathcal{X} := \mathcal{X}_{x_0}$, the filter of neighborhoods of x_0.

A function $f : X \to Y$ between two topological spaces can be seen as a (total) function between filtered spaces. If f takes x_0 to y_0, then $f : X \to Y$ is continuous at x_0 iff $f : (X, \mathcal{X}_{x_0}) \to (Y, \mathcal{Y}_{y_0})$ is continuous.

Now fix an index set $\mathbb{I}$ and a filter $\mathcal{F}$ on it; our semi-standard universe will be $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$.

A *pre-hyperpoint* of $(X, \mathcal{X})$ is a function from $(\mathbb{I}, \mathcal{F}) \to (\mathcal{X}, \mathcal{X})$. A *total pre-hyperpoint* of $(X, \mathcal{X})$ is a function whose domain is the whole $\mathbb{I}$. Two pre-hyperpoints $(\mathbb{I}, \mathcal{F}) \to (X, \mathcal{X})$ are *equivalent* when they coincide in an $\mathcal{F}$-big set of indices. *Hyperpoints* are pre-hyperpoints modulo equivalence.

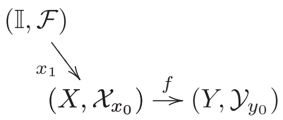
A *pre-infinitesimal* in $(X, \mathcal{X}_{x_0})$, for us, will be a pre-hyperpoint $x_1 : (\mathbb{I}, \mathcal{F}) \to (X, \mathcal{X}_{x_0})$ "infinitely close to the chosen point" x_0, in the sense that for each standard neighborhood $X' \ni x_0$ the x_1 belongs to X' for a big set of indices — formally, $\forall X' \in \mathcal{X}_{x_0}. x_1^{-1}(X') \in \mathcal{F}$. *Infinitesimals* are pre-infinitesimals modulo equivalence.

Note that as the chosen point doesn't need to be 0, "infinitesimal" becomes an umbrella term that can mean "infinitely small" ($x_0 = 0$), "infinitely close to", and even "infinitely big" ($x_0 = \infty$).

These definitions were all chosen to make this work:

Key idea: *infinitesimality is the same as continuity at the chosen point.*

When we look at things in this way then it becomes obvious that a standard continuous function $f : X \to Y$ taking x_0 to y_0 takes any infinitesimal $x_1 \sim x_0$ to an infinitesimal $f(x_1) \sim y_0$.

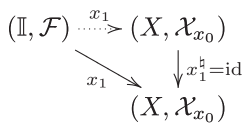


As this holds for and index set $\mathbb{I}$, any filter $\mathcal{F}$ on $\mathbb{I}$, and any infinitesimal $x_1 \sim x_0$, we can do much more.

6 Natural Infinitesimals

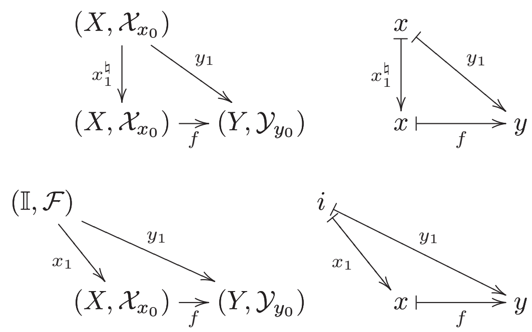
Definition: the *natural infinitesimal* on a (standard) filtered space $(X, \mathcal{X}_{x_0})$, that we will denote by $x_1^\natural \overset{\natural}{\sim} x_0$, is the identity function $x_1^\natural = \mathrm{id} : (X, \mathcal{X}_{x_0}) \to$

$(X, \mathcal{X}_{x_0})$; seen as an infinitesimal, it lives in $\mathbf{Set}^X/\mathcal{X}_{x_0}$. As it corresponds to the identity map, any other infinitesimal $x_1 \sim x_0$ — in the diagram below we take an x_1 living in $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ — factors through $x_1^\natural$ it in a unique way; this suggests that there is a kind of "change of base" operation between filter-powers.



Now, for any $f : (X, \mathcal{X}_{x_0}) \to (Y, \mathcal{Y}_{y_0})$ taking x_0 to y_0, this holds:

Key theorem:
(i) f is continuous at x_0
$\Leftrightarrow$ (ii) for $(\mathbb{I}, \mathcal{F}) := (X, \mathcal{X}_{x_0})$, $x_1^\natural \overset{\natural}{\sim} x_0$, we have $f(x_1^\natural) \sim f(x_0)$
$\Leftrightarrow$ (iii) for all $(\mathbb{I}, \mathcal{F})$ and $x_1 \sim x_0$, we have $f(x_1) \sim f(x_0)$.



Proof: (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) are obvious from what we've seen before — that the composite of continuous maps between filtered spaces is continuous. For $\neg$(i) $\Rightarrow$ $\neg$(ii), as f is not continuous at x_0, we can choose a $Y' \in \mathcal{Y}_{y_0}$ such that $f^{-1}(Y') \notin \mathcal{X}_{x_0}$; but then $y_1^{-1}(Y') = {x_1^\natural}^{-1}(f^{-1}(Y')) \notin \mathcal{X}_{x_0}$, and $f(x_1^\natural) \not\sim f(x_0)$. For $\neg$(i) $\Rightarrow$ $\neg$(iii), take $(\mathbb{I}, \mathcal{F}) := (X, \mathcal{X}_{x_0})$, $x_1 := x_1^\natural$, and reuse the proof of $\neg$(i) $\Rightarrow$ $\neg$(ii).

In texts about Non-Standard Analysis the infinitesimal characterization of continuity is presented in another form:

(i) f is continuous at x_0
$\Leftrightarrow$ (iv) for all $(\mathbb{I}, \mathcal{U})$ and $x_1 \sim x_0$, we have $f(x_1) \sim f(x_0)$.

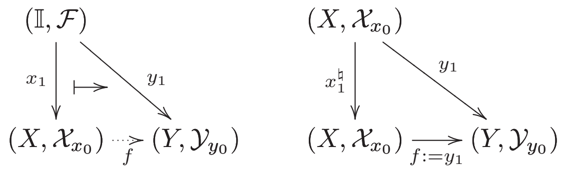
Clearly, (iii)$\Rightarrow$(iv); but to show that (iv) implies the rest we need to be in a universe with enough ultrafilters; [Ban83] is probably a good reference.

[B. Banaschewski, The Power of the Ultrafilter Theorem, Journal of the London Mathematical Society (2) 27, 193–202, 1983.]

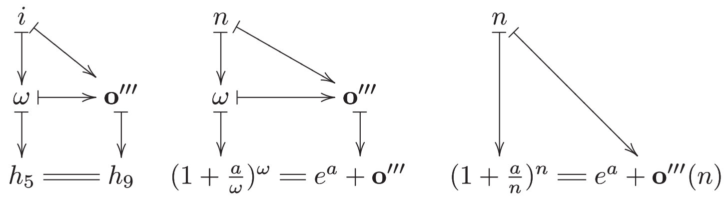
7 Translating a proof with infinitesimals

Each of the cells in the diagram in sec. 5 is an instance of the key theorem — maybe slightly disguised. For example, to prove that $g(b+\mathbf{o}) = (g'(b)+\mathbf{o}')\mathbf{o}$ we may start with $\frac{g(b+\mathbf{o})}{\mathbf{o}} - g'(b) = \mathbf{o}'$, for an infinitesimal $\mathbf{o} \neq 0$, i.e., $\lim_{\epsilon\to 0} \frac{g(b+\mathbf{o})}{\mathbf{o}}$.

What really matters, when we look at the diagrams, is that for any $(\mathbb{I}, \mathcal{F})$ and for any infinitesimal $x_1 : (\mathbb{I}, \mathcal{F}) \to (X, \mathcal{X}_{x_0})$ — maybe obeying some condition, like $\mathbf{o} \neq 0$ — there is a unique "adequate" infinitesimal $y_1 : (\mathbb{I}, \mathcal{F}) \to (Y, \mathcal{Y}_{y_0})$; we want to "represent" the operation $x_1 \mapsto y_1$ as a function $f : (X, \mathcal{X}_{x_0}) \to (Y, \mathcal{Y}_{y_0})$, and we can do that trivially by setting $(\mathbb{I}, \mathcal{F}) := (X, \mathcal{X}_{x_0})$, $x_1 := x_1^\natural$; then we can take $f := y_1$, and the f obtained in this way works in the general case.



Applying this idea to the composite of all cells in the example in sec. 5, we get this:



where $i \in (\mathbb{I}, \mathcal{F})$, $n, \omega \in (\mathbb{N}, \mathcal{N})$, and all the other "points" live in $(\mathbb{R}, \mathcal{R}_0)$. Note that the '$\mapsto$' arrows in this diagram do not stand for functions in the usual sense, but for functions between filtered spaces (not necessarily total). Incidentally, all of them are continuous.

8 Future directions

Let **Filt** be the category of filtered spaces (in **Set**). For each filtered space $(\mathbb{I}, \mathcal{F})$, the filter-power $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ can be seen as the fiber of some bigger category (a fibration, of course!), over the object $(\mathbb{I}, \mathcal{F})$ of the base category **Filt**.

I have not attempted (yet) to define precisely this fibration, or to study it — but it seems to be the right place to interpret "change of base" steps, like the passage from $h_5 = h_6$ to $h_6 = h_8$ n the example. By the way, the quotient $\mathbf{Set}^{\mathbb{I}}/\mathcal{F} \to \mathbf{Set}^{\mathbb{I}}/\mathcal{U}$ described in [...] (for $\mathcal{F} \subset \mathcal{U}$) is associated to seeing $\mathrm{id} : \mathbb{I} \to \mathbb{I}$ as a continuous function $(\mathbb{I}, \mathcal{U}) \to (\mathbb{I}, \mathcal{F})$.

Take a filter $\mathcal{F}$ over $\mathbb{I}$, and add the empty set to it: this new set of subsets, $\mathcal{F} \cup \emptyset \subset \mathcal{P}(\mathbb{I})$, is a topology on $\mathbb{I}$ — a funny one, in which there are no "divisors of zero": the intersection of two non-empty open sets is always another non-empty set. Now regard the topology $\mathcal{F} \cup \{\emptyset\}$ as a Heyting algebra; the double-negation operation takes $\emptyset$ to itself, and all non-empty open sets to X. I have the impression that if we sheafify $\mathbf{Set}^{(\mathbb{I},\mathcal{F}\cup\{\emptyset\})}$ (using the '¬¬' modality?) we get $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ — and then geometric morphims, etc. But (as my terminology reveals!) I know far less about sheaves than I should...

Define the "diameter" of a set in $\mathbb{R}$ in the obvious way. A standard function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a iff for any infinitesimal $\mathbf{o}$, $\mathrm{diam}(f([a-\mathbf{o}, a+\mathbf{o}]))$ is an infinitesimal. It is easy to adapt this idea to a set-valued function $F : \mathbb{R} \to \mathcal{P}(\mathbb{R})$, or to a relation $R \subset \mathbb{R} \times \mathbb{R}$. We may then consider a variation of the set-valued function F, which will take a real x and a "hint", and then return an element of $F(x)$; the hint selects which one. If we move the hint parameter to the index set we get a point-valued non-standard function that in some senses represents the behavior of the standard set-valued function F.

Take some function $g : \mathbb{R}^2 \to \mathbb{R}^2$ that has two derivatives at a point $a \in \mathbb{R}^2$. For infinitesimals o in $\mathbb{R}^2$, we have that $g(a + \mathbf{o}) = g(a) + (g'(a) + \mathbf{o}')\mathbf{o}$ — i.e., there exists an infinitesimal non-standard function $\mathbf{o}' : \mathbb{R}^2 \to \mathbb{R}^2$ that makes the equation hold... but this $\mathbf{o}'$ is by no means unique!

I have only described cases where the new infinitesimals are related to the previous ones by "unique existentials", as in $\forall\omega \sim \infty.\exists!\mathbf{o} \sim 0.R(\omega, \mathbf{o})$. I have the impression that it should be possible to deal with steps that introduce new infinitesimals with just '∃', like:

$$\forall x_1 \sim x_0.\ \exists y_1 \sim y_0.\ R(x_1, y_1)\text{—}$$

a sentence like that one is exactly when for any (standard) neighborhood $Y' \in \mathcal{Y}_{y_0}$ of y_0 its inverse image

$$X' := \{\, x \in X \mid \exists y \in Y'.\ R(x, y) \,\}$$

is a neighborhood of x_0 — i.e., this operation $Y' \mapsto X'$ is continuous... it seems that in these cases the natural infinitesimal should have $R \subseteq X \subseteq Y$ as its index set.

9 Related work

[Remember: this is a preliminary version...]

If I remember correctly, [Dav77] proves "(iv)⇒(i)" from sec. 6 by starting with a filter of neighborhoods, and then enlarging it to an ultrafilter on the same index set; I got the inspiration for trying to work with filters from his "Concurrence Theorem" — and from my difficulties in developing a good intuition about NSA's infinitesimals.

In [BW05] filters are used in a way that seems to be closely related to the one that I use in this paper; and in sec. 8 of [Bee95] there's a procedure for "elimination of infinitesimals" in a "neighborhood semantics".

In the papers [Nel77], [Nel88], Edward Nelson discusses — among many other things — automatic ways to translate non-standard proofs to standard; his setting and his techniques are very different from the ones that I use here. In [Rob73] (§11) Abraham Robinson states the need for tools to analyze lengths of proofs.

Jonas Eliasson uses a category of filters in his [Eli01], and its bibliography points to several papers by Palmgren and Blass that look interesting, but that I have not yet been able to get.

Also: [But99], [AH02]

In Johnstone's "Topos Theory" book ([Joh77]) there's a section about the "filter-power construction" (9.4), and another one about sheaves as categories of fractions (3.4). I need to thank Peter Arndt for helping me to understand them; I still don't understand them as much as I should (far from it!), but I can assure him his efforts have not been in vain...

I described the example in sec. 4 briefly in my MsC thesis (1999) and in a presentation that I gave about it in 2000; the text and the slides for the presentations are available from `http://angg.twu.net/math-b.html#MsC`, but they are in Portuguese, and at that point I didn't know how to generalize the method, or how to characterize "purely calculational proofs".

10 A final note

The idea of filter-infinitesimals is not new, the proof of the key theorem is obvious, and "infinitesimals as sequences tending to zero" should be one of the basic intuitive ways to understand infinitesimals, especially if we use some kind of "physicists' notation" and we omit the index variable... also, somewhere in his [Bel98] *(oh no, I lost the page again... 8-()* Bell says something along the lines of: "all proofs in Calculus are constructive". It should be clear, then, that the non-two-valuedness of $\mathbf{Set}^{\mathbb{I}}/\mathcal{F}$ shouldn't be an obstacle, and that most proofs in Calculus, being constructive, should lift through the quotient $\mathbf{Set}^{\mathbb{I}}/\mathcal{F} \to \mathbf{Set}^{\mathbb{I}}/\mathcal{U}$...

Why can't I find the ideas in this paper published anywhere else? Am I missing something, or are they new?

References

[AH02] Jeremy Avigad and Jeffrey Helzner. Transfer principles in nonstandard intuitionistic arithmetic. *Archive for Mathematical Logic*, 41:581–602, 2002.

[Bee95] Michael Beeson. Using nonstandard analysis to verify the correctness of computations. *International Journal of Foundations of Computer Science*, 6(3):299–338, 1995.

[Bel98] John. L. Bell. *A Primer of Infinitesimal Analysis*. Cambridge University Press, 1998.

[But99] Carsten Butz. The filter construction revisited. Preprint, 1999.

[BW05] Michael Beeson and Freek Wiedijk. The meaning of infinity in calculus and computer algebra systems. *Journal of Symbolic Computation*, 39(5):523–538, 2005.

[Dav77] Martin Davis. *Applied Nonstandard Analysis*. John Wiley, New York, 1977.

[Eli01] Jonas Eliasson. Ultrapowers as sheaves on a category of ultrafilters. Technical report, Uppsala University, 2001.

[Hal64] J. D. Halpern. The independence of the axiom of choice from the boolean prime ideal theorem. *Fundamenta Mathematicae*, pages 57–66, 1964.

[HK86] C. Ward Henson and H. Jerome Keisler. On the strength of nonstandard analysis. *The Journal of Symbolic Logic*, 51(2):377–386, 1986.

[Joh77] P. T. Johnstone. *Topos Theory*. Academic Press, 1977.

[Nel77] Edward Nelson. Internal set theory; a new approach to nonstandard analysis. *Bulletion of the American Mathematical Society*, 83:1165–1198, 1977.

[Nel88] Edward Nelson. The syntax of nonstandard analysis. *Annals of Pure and Applied Logic*, 38:123–134, 1988.

[Rob73] Abraham Robinson. Metamathematical problems. *Journal of Symbolic Logic*, 38(3):500–516, 1973.