A theory of objects and visibility. A link between relative analysis and alternative set theory

Richard O'Donovan

N° d'Ordre: D.U. 2145

UNIVERSITÉ BLAISE PASCAL

U.F.R. Sciences et Technologies

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES

N° 680

THÈSE

présentée pour obtenir le grade de

DOCTEUR D'UNIVERSITÉ

Spécialité : Mathématiques Pures

par O'DONOVAN Richard
Master

A Theory of Objects and Visibility

A link between Relative Analysis and Alternative Set Theory

Soutenue publiquement le 7 juillet 2011, devant la commission d'examen.

JURY :

Président:	Zlatoš Pavol	Professeur, Université Commenius, Bratislava
Examinateurs:	van den Berg Imme	Professeur, Université de Evóra
	Péraire Yves	Professeur, Université Blaise Pascal, Clermont-Ferrand. Directeur de thèse

RAPPORTEURS :

Zlatoš Pavol	Professeur, Université Commenius, Bratislava
Di Nasso Mauro	Professeur, Université U. Dini, Pise
van den Berg Imme	Professeur, Université de Evóra

Contents

Acknowledgements

My greatest thanks go to Yves Péraire who accepted to guide me and advise me in spite of the three hundred and fifty kilometers between our offices. He provided provocative criticism and good humour all along.

These greatest thanks must be shared with Karel Hrbacek, Pavol Zlatoś and Mauro Di Nasso. Without Karel Hrbacek's continued helpful comments and feedback, I would never have studied nonstandard analysis in such depth. His reading of an earlier version of this work was invaluable. Pavol Zlatoś provided extremely fruitful insights into the technical and philosophical aspects of the alternative set theory. Mauro Di Nasso was one of the first to recognise the challenge to make the link between research and teaching. He was kind enough to allow me to present the pedagogical approach and related problems in mathematical congresses that he organised.

Thanks to their friendly welcome, I have also come to know Clermont-Ferrand, New York, Bratislava and Pisa.

Thanks also to Imme van den Berg for accepting to be in the jury and to Youcef Amirat, Yannick Heurteaux and Pierre Henrard for receiving me in the UFR of Mathematics at Blaise Pascal University.

I would also like to thank my colleagues Olivier Lessmann for his encouragements and help, John Kimber for the endless mathematical discussions by the bike-shed and my headmaster, Roland Jeannet, for the timetables.

Eliane O'Donovan is probably one of the only English teachers who can make a presentation about nonstandard analysis. I am not quite certain she wants to thank me for that.

Résumé

La théorie présentée ici est issue d'années d'enseignement de l'analyse au niveau pré-universitaire en utilisant d'abord le concept d'infiniment petit, tel que défini dans l'analyse nonstandard de Robinson, puis ensuite d'ultrapetit, tel que défini dans notre travail en collaboration avec Hrbacek et Lessmann et présenté en annexe. Á la suite de ces recherches, s'est posée la question : Si l'on a à disposition des quantités finies mais ultragrandes, est-il possible de se passer de quantité dites infinies ?

La théorie alternative des ensembles de Vopěnka est une théorie avec des ensembles finis et des classes qui, elles, peuvent être infinies. La théorie des objets est le résultat d'un mélange de certains axiomes de Vopěnka avec des axiomes déterminant des niveaux de visibilité tels que dans l'analyse relative.

On s'est donné comme premier principe : $x \subseteq y \Rightarrow x \sqsubseteq y$ qui spécifie que si l'objet x est inclus dans l'objet y, alors x "paraît" au niveau de y. Cette affirmation serait fausse avec des quantités infinies ; elle est néanmoins une caractérisation des ensembles finis : cela est bien connu en analyse nonstandard. L'introduction de ce principe comme point de départ est donc une affirmation forte que les objets devront être finis au sens habituel de ce terme. L'autre axiome fondateur ici est le schéma d'axiomes d'induction de Gordon et Andreev : Si Φ est une formule, et si $\Phi(\emptyset)$ est vrai et que $\Phi(x)$ et $\Phi(y)$ impliquent $\Phi(x \cup \{y\})$, alors $\Phi(x)$ est vrai pour tout x. Un accent particulier est mis sur le concept de formules dites contextuelles. Ce concept est une de nos contributions à l'analyse relative de Hrbacek et détermine les formules bien formées.

On montre que le système qui en résulte est relativement cohérent avec la théorie FRIST de Hrbacek et la théorie RIST de Péraire qui sont elles-mêmes des extensions conservatives de ZFC. La théorie des objets est une extension de la théorie des ensembles de Zermelo et Fraenkel sans axiome du choix et négation de l'axiome de l'infini. Les nombres entiers et rationnels sont définis et ces derniers sont munis de relations d'ultra-proximité. Une ébauche d'une construction de "grains numériques" est présentée : ces nombres pourraient avoir des propriétés suffisamment semblables aux nombres réels pour permettre de faire de l'analyse.

Mots clés : analyse nonstandard, analyse contextuelle, théorie alternative des ensembles, fondations, fini, niveau, IST, RIST, FRIST.

Mathematics Subject Classification : Primaire : 26E35, 03E35. Secondaire 03E65, 03E70, 03F25.

Abstract

The theory presented here stemmed from years of teaching analysis at pre-university level first using the concept of infinitesimal as defined in nonstandard analysis by Robinson, then the concept of ultrasmall as defined in our joint work with Hrbacek and Lessmann presented in the appendix. This research led to the question: If one has finite yet ultralarge quantities, is it possible to avoid infinite quantities?

The alternative set theory of Vopěnka is a theory of finite sets including classes that can be infinite. The theory of objects is a merger of certain axioms of Vopěnka with axioms that determine levels of visibility as in relative analysis.

We took as first principle: $x \subseteq y \Rightarrow x \sqsubseteq y$, which specifies that if object x is included in object y, then x "appears" at the level of y. This statement would be false with infinite quantities and is in fact a characterisation of finite sets: this is a well-known theorem of nonstandard analysis.

The introduction of this principle as starting point is making a strong point that all objects will be finite – in the usual sense of the word. The other founding axiom is Gordon and Andreev's axiom schema: If Φ is a formula, and if $\Phi(\emptyset)$ is true and that $\Phi(x)$ and $\Phi(y)$ imply $\Phi(x \cup \{y\})$, then $\Phi(x)$ is true for all x. An emphasis is made on the concept of contextual formulae. This concept is one of our contributions to relative analysis of Hrbacek and determines an equivalence to well-formed formulae.

We show that the resulting system is relatively consistent with Hrbacek's FRIST and Péraire's RIST which are conservative extensions of ZFC. The theory of objects extends set theory of Zermelo and Fraenkel without choice and with negation of the infinity axiom. Integers and rationals are defined and endowed with an ultraproximity relation. A draft of a construction of "numeric grains" is presented: these numbers could prove to have properties sufficiently similar to real numbers to allow to perform analysis.

Keywords: nonstandard analysis, contextual analysis, alternative set theory, foundations, finite, level, IST, RIST, FRIST.

Mathematics Subject Classification: Primary: 26E35, 03E35. Secondary 03E65, 03E70, 03F25.

Introduction

Important questions may have naive origins: high school students will ask about the meaning of things mathematical such as: "Do limits really exist if we never reach them?" or "Is infinity *the* infinity?" Far less naive is the question posed by Vopěnka [Benešová et al., 1989]: "How is it possible that some results of classical infinite mathematics are applicable in the real world and others are not?" His claim is that "the existence of such [infinite] sets in the real world is, to say the least, dubious." Yet mathematics may produce results which are applicable, hence his reformulation: "How is it that some real situations admit models – often relatively faithful – in classical infinite mathematics?"

The foundational aspects of students' questions combined with the pedagogical difficulties attached to the traditional presentation of analysis made us consider other mathematical theories as background for teaching analysis in Geneva high school.

Working on theoretical approaches which could make more pedagogical sense, "Analysis with Ultrasmall Numbers" was developed in collaboration with Karel Hrbacek and Olivier Lessmann. A presentation of this approach is given in appendix A (published in [Hrbacek et al., 2010a]). An assessment of its use in class is given in appendix C (published in [O'Donovan, 2009]). Appendix D is a short article discussing misunderstandings that appear concerning real numbers when presented in informal manner (published in [O'Donovan, 2010]). Appendix B is an excerpt from a book to appear containing some proofs used in chapter 7 [Hrbacek et al., 2010b].

Some pedagogical aspects also led to the question of whether "infinity" is necessary for analysis.

The Greeks considered two sorts of infinity. *Potential infinity* is the infinity of never ending processes. For any whole number, it is possible to find one which is greater. Counting never ends. This infinity is not deniable. The other is *actual infinity*: infinity understood as a completed whole. There is "something" that is greater than all whole numbers. The existence of this sort of infinity requires an act of faith. It is not the generalisation of something that can be observed.

It will be considered here whether actual infinity can be avoided. The existence of the actually infinite is neither denied nor acknowledged. On this question, we remain agnostic. Yet, as will be discussed, the question of what "infinity" stands for, and whether some subcollections of objects could be considered to have a flavour of infinity, depends on how this concept is defined (see page 35).

After studying versions of nonstandard analysis for several years, we encountered the Alternative Set Theory as developed by Vopěnka (also presented by Sochor and Zlatoś [Vopěnka, 1979, Benešová et al., 1989]) in which all sets are finite. Proper classes are accepted, but there is a drastic difference between methods applicable to sets and those applicable to proper classes.

We thus considered a merger of Alternative Set Theory (AST) and relative analysis.

The Theory of Objects (TO) is developed from axioms upwards. Instead of starting with ZFC axioms and adding extra axioms as is done with RIST (Relative Internal Set Theory) or FRIST (Fully Relativised Internal Set Theory), the idea was to go the other way: consider axioms that introduce the concept of "ultralarge" (numbers with a flavour of "infinity") and add only those axioms which seem necessary.

From relative analysis, the concepts of levels and of contextual formulae are taken. These will be considered the basic building blocks of the language. From the alternative set theory finiteness of sets will be kept. Vopěnka gives an image of alternative set theory as a tree with a common trunk and branches which could even be mutually incompatible. The following theory of objects is thus a branch of alternative set theory, only it stems out of the trunk far lower than Vopěnka expected as it separates before the definition of proper classes. The theory of objects retains most of the philosophical views expressed by the alternative set theory, in particular some dynamic interpretations.

To pay our debt, the following can thus be viewed as:

A Contextual Alternative Set Theory

Chapter 1

Axioms

The following axiomatic construction is classical: axioms are given that characterise collections called objects and a first order universe of discourse is developed. "Sitting on the edge of the world" one can see that these objects are finite collections. Yet from the inside of this universe of discourse, everything being finite, the word "finite" itself becomes meaningless. Hence the use of the word "object" rather than a contextually meaningless adjective. The point of view is syntactical: axioms specify which objects the language can talk about. When an axiom or a statement stipulates that something exists, it must be understood as meaning that it exists in the universe of discourse. Whether there is a stronger form of existence to these mathematical objects will not be addressed here.

Comments and remarks about the theory – comparisons with other theories – not being part of the theory itself, will appear using a different font, such as this one.

The undefined predicates are infixed "$=$", "$\in$" and "$\sqsubseteq$" . As would be expected, $x = y$ is interpreted as meaning that x and y are identical. The predicate $\in$ is classically interpreted for $x \in y$ as meaning "object x belongs to object y" and $x \sqsubseteq y$ is interpreted as meaning "x is at least as visible as y" or "x is observable when y is observable". It may be said that "x is y-observable".

The "$\sqsubseteq$" symbol is interpreted the same way as in RIST and FRIST.

The intuitive image is that the universe of discourse about objects is stratified according to visibility, where visibility can also be interpreted as a generalisation of the concept of different scales of observation as in physics.

Formulae are constructed inductively.

Definition 1.

(1) If x and y are variables then $x = y$, $x \in y$ and $x \sqsubseteq y$ are formulae.

(2) If Φ and Ψ are formulae then $(\Phi \wedge \Psi)$ and $\neg\Phi$ are formulae.

(3) If x is a variable and Φ is a formula then $(\exists x)\Phi$ is a formula.

A closed formula, or formula with no free variables, is a statement.

A formula constructed as above, but nowhere using the $\sqsubseteq$ symbol, is an $\in$-formula.

If $x \sqsubseteq y$ and $\neg(y \sqsubseteq x)$ then the notation $x \sqsubset y$ is used (x is strictly more visible than y). If $x \sqsubseteq y$ and $y \sqsubseteq x$, then the notation $x \sqsupseteq\sqsubseteq y$ is used (x and y have same visibility).

The notation $x \notin y$ is shorthand for $\neg(x \in y)$; and as usual, $(\forall x)P(x)$ is shorthand for $\neg((\exists x)\neg P(x))$.

The notation $\Phi \vee \Psi$ is shorthand for $\neg(\neg\Phi \wedge \neg\Psi)$. The notation $\Phi \Rightarrow \Psi$ is shorthand for $(\neg\Phi) \vee \Psi$ and $\Phi \Leftrightarrow \Psi$ is shorthand for $(\Phi \Rightarrow \Psi) \wedge (\Psi \Rightarrow \Phi)$.

Following common usage, parentheses may be omitted for the sake of simplicity when the meaning is clear.

Axiom 1 (Atomic object)**.**

$$(\exists x)(\forall y)\ (y \notin x)$$

Axiom 2 (Extensionality)**.**

$$(\forall x)(\forall y)[(\forall z)\ (z \in x \Leftrightarrow z \in y) \Leftrightarrow (x = y)]$$

An object satisfying axiom 1 (having no element), by extensionality, is unique. This is called the atomic object or the empty object. It is denoted by $\emptyset$.

Axiom 3 (Successor object)**.**

$$(\forall x)(\forall y)(\exists z)(\forall u)\ [(u \in z) \Leftrightarrow (u \in x \vee u = y)\ ]$$

The object successor expresses the fact that for any objects x and y, there is an object z containing y and the elements of x.

The union of two objects a and b, if it exists, is the collection of all elements of a and all elements of b, noted $a \cup b$ and $x \in a \cup b \Leftrightarrow (x \in a \vee x \in b)$. The successor object axiom thus states that given objects x and y, the object $x \cup \{y\}$ exists. The existence of the union for any two arbitrary objects is shown by proposition 5.

These three axioms are the only ones that refer directly to objects with no mention of levels. The next axiom describes the properties of levels (the "visibility" of objects), then there are three axiom schemata which enable the construction of more complex objects.

Axiom 4 (Levels of observation)**.**

(1) (Transitivity of levels)

$$(\forall x)(\forall y)(\forall z)\ [\ (x \sqsubseteq y) \wedge (y \sqsubseteq z) \Rightarrow (x \sqsubseteq z)\ ]$$

(2) (Discreteness)

$$(\forall x)(\forall y)[(x \sqsubset y) \Rightarrow [(\exists z)(x \sqsubset z \sqsubseteq y) \wedge \neg(\exists u)(x \sqsubset u \sqsubset z) \\ \wedge (\exists w)(x \sqsubseteq w \sqsubset y) \wedge \neg(\exists v)(w \sqsubset v \sqsubset y)]]$$

The first item of this axiom is similar to the definition of levels in (F)RIST[1].

The first part of the discreteness axiom states that for any object, there is an object of immediate finer level. The second part of the axiom states the existence of an immediate coarser level for objects which are not of coarsest level. It seems necessary to state also this property in order to disprove the existence of a fine "limit level" (in analogy with limit ordinals: a level which has no predecessor and is finer than those such that all of their elements are of levels which have predecessor levels). Such a limit level would probably turn out to be contradictory with the rest of the theory of objects.

1.1 Notation

The notation $(\exists^a x)\ P(x)$ (resp. $(\forall^a x)\ P(x)$) stands for $(\exists x)\ (x \sqsubseteq a) \wedge P(x)$ (resp. $(\forall x)\ (x \sqsubseteq a) \Rightarrow P(x)$). These are called **level quantifiers**.

$(\exists^{a,b} x) P(x)$ stands for $(\exists x)[(x \sqsubseteq a \vee x \sqsubseteq b) \wedge P(x)]$.

$x \sqsubseteq a_1, \ldots, a_k$ stands for $x \sqsubseteq a_1 \vee \cdots \vee x \sqsubseteq a_k$ and this can be written $(\exists^{a_1,\ldots,a_k} t)(t = x)$. Here, $a_1, \ldots, a_k$ are **level variables**.

$a_1, \ldots, a_k \sqsubseteq x$ stands for $a_1 \sqsubseteq x \wedge \cdots \wedge a_k \sqsubseteq x$.

Note that level variables can be free or bound.

As shown above, references to levels can be rewritten using a level quantifier. The following definition requires that all references to levels be made by level quantifiers. The main reason for this is that it offers a direct way to check certain conditions on levels.

In the following, a **parameter** is a free variable. This may not be the most widespread usage but helps clarify the issues involved with determining levels. It is used in this way in [Hrbacek et al., 2010a].

Definition 2. *A* ***contextual formula*** Φ *(or briefly:* ***c****-formula) is a formula where all references to levels are made by level quantifiers and all parameters of* Φ *are among the level variables for each occurrence of level quantifiers.*

In particular, every formula where the symbol "$\sqsubseteq$*" never appears is a* ***c****-formula.*

[1] The acronym RIST stands for Relative Internal Set Theory, developed by Péraire. FRIST stands for Fully Relativised Set Theory, developed by Hrbacek. We write (F)RIST to refer to both theories simultaneously.

c-statements (closed formulae) follow the same rule as **c**-formulae and, in general, any noun preceded by "**c**-" denotes a reference to context level.

If $\Phi(x)$ is a **c**-formula, then the context depends also on x, hence the context level of a formula may well be variable.

1.2 Contextual formulae: Examples

- The first three axioms do not refer to levels. They are closed formulae, hence they are contextual statements.
- Axiom 4 is about levels. Each part of the axiom is a closed formula, hence references to levels do not refer to parameters. The axiom can be rewritten with level quantifiers. The rewriting of the first part is shown. $(\forall x)(\forall y)(\forall z)(x \sqsubseteq y) \wedge (y \sqsubseteq z) \Rightarrow x \sqsubseteq z)$ is $(\forall x)(\forall y)(\forall z)[(\exists^y t)(t = x) \wedge (\exists^z u)(u = y) \Rightarrow (\exists^z v)(x = v)]$. Since the statement has no free variable, the statement is contextual. The rewriting of the other part of this axiom is similar.
- $(\exists^x t)(t \neq x) \wedge (\exists^{x,z} t)(t = x)$ has free variables x and z. The first level quantifier does not refer to z, hence this is not a **c**-statement.
- $(\exists^{x,z} t)(t \neq x) \wedge (\exists^{x,z} t)(t = x)$ is a **c**-statement. Its free variables are x and z.
- $(\exists^{x,z,v} t)(t \neq x) \wedge (\exists^{x,z} v)(v = x)$ is a **c**-statement. Its free variables are x and z. They are in all lists of level variables.
- $(\forall z)[(\exists^x t)(t \neq x) \wedge (\exists^{x,z} t)(t = x)]$ is a **c**-statement. Its free variable is x.

  This example shows that it is possible to quantify over variables that appear as level variables only (in this case: z).
- Anticipating on the definition of functions, in $f(x)$, classically, x is free, hence $f(x)$ can be contextual only if level quantifiers used in the definition of f have x in the list. The context level will depend on x and is thus not constant over the domain of the function.

1.3 Notation

If $a_1, \ldots a_k$ are all the parameters of Φ and if $a_1, \ldots a_k \sqsubseteq b$, then the shorthand $\Phi \sqsubseteq b$ may be used. (It does *not* imply that we assume Φ to be an object, it is a simplified way of stating a property about its parameters.)

If Φ is contextual, then Φ^v is obtained from Φ by adding v to each list of level variables i.e., each occurrence of $\exists^{a_1,\ldots,a_k}$ is replaced by $\exists^{a_1,\ldots,a_k,v}$ and each occurrence of $\forall^{a_1,\ldots,a_k}$ is replaced by $\forall^{a_1,\ldots,a_k,v}$. The formula Φ is **relativised** to the level of v. Since Φ is assumed to be contextual, either v is a free variable which is not already in Φ, then Φ^v is also contextual, or v is a free variable which is already in Φ – in which case the extra reference is irrelevant and Φ^v is contextual and is the same as Φ, or v is bound in Φ – again, Φ^v is contextual.

1.4 Relativisation: Examples

- If $\Phi(x)$ is $(\exists^x z)(z = x)$, then $\Phi^v(x)$ is $(\exists^{x,v} z)(z = x)$. It is clear that $\Phi(x)$ is contextual and so is $\Phi^v(x)$.
- If $\Phi(x)$ is $(\exists z)(z = x)$, then $\Phi^v(x)$ is $(\exists z)(z = x)$. This is because $\Phi(x)$ has no level quantifier, hence there is no reference to levels which can be refined.
- If $\Phi(x)$ is $(\exists^{x,u} z)(z = x)$, then $\Phi^v(x)$ is $(\exists^{x,u,v} z)(z = x)$.
- If v is observable at the context level of Φ, then Φ^v refers to the same levels as Φ and is thus equivalent to Φ. The context level may vary with x but if v is of the level of a parameter, then adding v to the list of level variables is irrelevant.

  If $\Phi(x)$ is $(\exists^x z)(z = x)$ and $v = \emptyset$, then $\Phi^v(x)$ is $(\exists^{x,\emptyset} z)(z = x)$. This is $(\exists z)(z \sqsubseteq x \vee z \sqsubseteq \emptyset) \wedge (z = x)$. Since $\emptyset \sqsubseteq x$, by transitivity (axiom 4-1) $z \sqsubseteq \emptyset$ implies $z \sqsubseteq x$, hence the formula can be rewritten: $(\exists z)(z \sqsubseteq x \vee z \sqsubseteq x) \wedge (z = x)$.

The contextual statements of TO are interpreted as internal statements of FRIST as shown by metatheorem 3, page 57.

Since the theory of objects is a first order theory, any letter preceded by a quantifier stands for a variable.

Axiom 5 (Induction). *Let Φ be a **c**-formula.*

$$(\forall x)(\forall y)[\Phi(\emptyset) \wedge (\Phi(x) \wedge \Phi(y) \Rightarrow \Phi(x \cup \{y\}))] \Rightarrow (\forall x)\Phi(x)$$

The only parameter outside of Φ is $\emptyset$ which is of coarsest level. Since Φ is assumed to be a **c**-formula, if it refers to levels, it automatically refers to the coarsest level also, hence this axiom is contextual.

Axiom 6 (Refinement). *Let Φ be a **c**-formula, with free variables x and y and possibly other free variables $x_1, \ldots x_n$.*

For all v such that $u \sqsubset v$ and $x_1, \ldots, x_n \sqsubseteq v$

$$(\forall^u a)(\exists y)(\forall x \in a)\Phi^v(x, y) \Leftrightarrow (\exists y)(\forall^u x)\Phi^v(x, y)$$

It is not required that $x_1, \ldots x_n \sqsubseteq u$, only that $x_1, \ldots x_n \sqsubseteq v$. The main interest in using this axiom is to show that there is a y which is *not* u-observable, having the required property. If Φ does not refer to levels then Φ^v is Φ since v would be added only to level quantifiers.

The contrapositive form is often used:

$$(\exists^u a)(\forall y)(\exists x \in a)\Phi^v(x, y) \Leftrightarrow (\forall y)(\exists^u x)\Phi^v(x, y)$$

The refinement principle is the transposition of the idealisation axiom of (F)RIST. This axiom enables to talk about objects of finer levels of observation. Note that in (F)RIST, the axiom starts with $(\forall^{a,fin} u)$ restricting to "finite" sets. In the theory of objects, this restriction is not given. It will be shown, however, that objects are "finite" in the usual sense.

The free variables of this axiom are u and $x_1, \ldots x_n$ and also v. It also refers to u which is not necessarily finer than the level of the parameters. Hence this axiom is not contextual. It is the only non contextual axiom. It allows specifically to elaborate a discourse about objects in a finer level than a given level.

Axiom 7 (Transfer principle)**.** *Let $\Phi(x_1, \ldots, x_n)$ be a **c**-formula and v be a free variable.*

$$\Phi^v(x_1, \ldots, x_n) \Leftrightarrow \Phi^{v'}(x_1, \ldots, x_n)$$

Note that the above formula is equivalent to

$$\Phi(x_1, \ldots, x_n) \Leftrightarrow \Phi^v(x_1, \ldots, x_n)$$

since it is also true if v is not finer than a parameter already in the list.

This axiom is clearly a contextual statement.

Axiom 7 is the transposition of transfer as given in FRIST. It is a homogeneity axiom stating that no level has a specific role. Any level can be considered a context level provided it is fine enough.

The list of axioms of TO is quite small – even though the three axiom schemata are in fact rules for infinitely many axioms. It is small in the sense that only a small portion of the usual properties of set theory are included in the list: union, intersection, the collection of all subobjects, pairing, choice, etc. are proved to exist due essentially to the induction axiom. Specification is a theorem schema.

1.5 Similarities with the Alternative Set Theory

The existence of the empty set and extensionality are formulated the same way in the theory of objects, in the Alternative Set Theory and in ZFC.

The set successor axiom does not exist in ZFC and might be considered the characteristic of AST. It goes with the induction axiom of AST which proves what in ZFC are axioms: the pair axiom, the union set, the power set and the replacement schema. No other axiom about sets are introduced in AST. (In one version [Vopěnka, 1979], the induction axiom is given in the form of proposition 1 but then an axiom of regularity or foundation is added. In another [Andreev and Gordon, 2006], it is given as axiom 5.)

Clearly, the infinity axiom and axiom of choice cannot be introduced in AST. With the infinity set and the powerset, countable and uncountable transfinite cardinals are proven to exist in ZFC. The induction axiom cannot be used to prove, say, the existence of the union of two uncountable sets.

The theory of objects is thus philosophically closer to AST than to ZFC. Yet AST also considers classes. The universe of sets is said to be *extended*. The axiom of the existence of classes [Vopěnka, 1979] is:

For each property $\varphi(x)$ of sets from the universe of sets, the extended universe contains the class $\{x \mid \varphi(x)\}$.

This axiom is followed by axioms which stipulate properties of classes. Classes encompass some collections which in ZFC are sets, such as the class of whole numbers, but there are also classes such as the universal class V, where $x \in V \Leftrightarrow x = x$. Note that set theorists often consider the class of all sets even though set theory does not define such classes. AST avoids Russel's antinomy and other logical difficulties by specifying rules for working with classes and other rules for working with sets.

Then semisets are defined as being subclasses of sets. Proper semisets are semisets which are not sets.

In TO, the convenient assumption that every definable property can be used to define a collection of all objects having that property, is not made. There are many properties in mathematics that characterise potentially infinitely many objects. The axioms of AST proves the existence of proper classes that are the extensions of such properties; TO stems from its trunk before this point.

It will be shown below (chapter 4) that semisets of AST bare similarities with incomplete objects of TO – which are due to the existence of levels – thus allowing to introduce in TO a concept comparable to potential infinity.

1.6 Absence of standardisation

When performing analysis, one of the key results of nonstandard analysis is the existence of the standard part of a real number that is not infinitely large. In (F)RIST this transposes to the existence of the observable part of a real number that is not ultralarge (relative to some observation level). The name of the axiom comes from IST. In IST with two levels (standard and nonstandard – with the wording corresponding to this theory): If x is a nonstandard real number not infinitely large, then $A = \{u \in \mathbb{R} \mid u \leq x\}$ is a nonstandard set. Standardisation ensures that there is a standard set A' such that standard elements of A are the standard elements of A'. This standard set has a least upper bound c and by closure, this least upper bound is standard and $c \approx x$. This is the standard part of x.

In RIST, α-external formulae are formulae which do not refer to levels or where reference is in the form $(\exists^\beta x)P(x)$ with β of a finer level than α (where $P(x)$ is an α-external formula).

RIST Standardisation: [Péraire, 1996]

For α-external formula P:

$$(\forall^\alpha y)(\exists^\alpha z)(\forall^\alpha t)[t \in z \Leftrightarrow (t \in y \wedge P(t))]$$

For any set $y \sqsubseteq a$, and any formula P which either does not refer to levels or refers to levels finer than the level of a, there is an a-observable set containing all a-observable elements of A satisfying P. If A is $\mathbb{N}$ and $P(x)$ is the formula $x \in \mathbb{N} \wedge 0 < x < n$ for 1-large N (in the sense of a-largeness given in RIST), the unique 1-standardisation of $\mathcal{N} = \{x \in \mathbb{N} \mid P(x)\}$ is $\mathbb{N}$ itself. The standardisation of a finite set can be an infinite set.

FRIST Standardisation: [Hrbacek et al., 2010b]

Given **V**, A, $x_1, \ldots, x_k$:

$$(\exists B \in \mathbf{V})\ (\forall y \in \mathbf{V})\ (y \in B \Leftrightarrow y \in A\ \ \wedge\ \mathcal{P}(y, x_1, \ldots, x_k; \mathbf{V}))$$

Here, the formula P does not refer to levels finer than **V**.

In the theories FRIST and RIST one considers families of formulae which are not quite the same; however, differences are not important here. The example given above for RIST would also produce – in FRIST – an infinite set for the standardisation of the finite set $1, 2 \ldots, N$.

Of course, the infinite sets given by standardisation are subsets of referential sets which are already infinite. Standardisation does not "produce" such sets, it only extracts them from pre-existing sets.

In TO if the referential is an object, then the element theorem shows that it contains no element of a finer level than that of the object: hence standardisation of the object would be the object itself.

Since there is no object containing all "finite" ordinals, there is no object of coarsest level containing $1, 2, \ldots, N$ for ultralarge N. Consequences of standardisation in its classical forms must be avoided in TO.

Chapter 2

The Theory of Objects

In this chapter, consequences of the axioms are drawn. As mentioned above, the induction axiom allows to deduce many properties which in ZFC require extra axioms, such as union, specification, "powerset", and others. It is also shown that ordered objects have minimal elements. These first results do not use the concept of levels.

Consequences which are closer to nonstandard properties are also drawn: overflow and also the fact that levels are not objects.

The two next results provide schemata for proofs which are immediate consequences of axiom 5.

Proposition 1. *Let* Φ *be a* ***c****-formula.*

$$(\forall x)(\forall y)[\Phi(\emptyset) \wedge (\Phi(x) \Rightarrow \Phi(x \cup \{y\}))] \Rightarrow (\forall x)\Phi(x)$$

Proof. In general and for any A, B and C, the following holds: $[A \Rightarrow C] \Rightarrow [(A \wedge B) \Rightarrow C]$.

Assume $\Phi(\emptyset)$ and $(\forall x)(\forall y)[\Phi(x) \Rightarrow \Phi(x \cup \{y\})]$. By the observation above, this implies $(\forall x)(\forall y)[\Phi(x) \wedge \Phi(y) \Rightarrow \Phi(x \cup \{y\})]$ and by induction, this implies $(\forall x)\Phi(x)$. □

The formulation of proposition 1, restricted to $\in$-formulae, is found in [Vopěnka, 1979] given as an axiom. Then axiom 5 is not given. In that case, some form of foundation is also given.

The inclusion $x \subseteq y$ is shorthand as usual for $u \in x \Rightarrow u \in y$. If $x \subseteq y$ and $\neg(y \subseteq x)$, then the inclusion is strict, written $x \subset y$. If $x \subseteq y$ then x is said to be a subobject of y.

It is possible to use induction on subobjects to prove that a given property holds for the object.

Corollary 1 (induction on subobjects)**.** *Let* $\Phi(x)$ *be a* ***c****-formula.*

If

$$\Phi(\emptyset) \wedge (\forall v)[v \subseteq x \wedge \Phi(v) \Rightarrow (\forall y \in x)(\Phi(v \cup \{y\})]$$

then $\Phi(x)$*.*

Proof. By proposition 1. Let $\Psi(v)$ be the formula

$$(v \subseteq x \wedge \Phi(v)) \vee (v \nsubseteq x)$$

As $\Phi(\emptyset)$ is assumed, there holds $\Psi(\emptyset)$.

Assume $\Psi(v)$ and let y be arbitrary. If $y \in x$ and $v \subseteq x$, then $v \subseteq x \wedge \Phi(v) \Rightarrow (\forall y \in x)(\Phi(v \cup \{y\})$. If $y \notin x$ or $v \nsubseteq x$, then $v \cup \{y\} \nsubseteq x$. Hence $\Psi(v \cup \{y\})$. By induction, $(\forall v)\Psi(v)$.

In particular, $(x \subseteq x \wedge \Phi(x)) \vee (x \nsubseteq x)$. Hence $\Phi(x)$. □

Theorem 1 (Foundation).

$$(\forall x)(\forall u)(x \notin x).$$

Proof. By induction. Let $\Phi(x)$ be the statement $(x \notin x) \wedge [(\forall u)(u \in x \Rightarrow u \notin u)]$. (It is a **c**-formula.) Then obviously $\Phi(\emptyset)$.

Assume $\Phi(x)$ and $\Phi(y)$. Then $u \in x \cup \{y\} \Rightarrow (u \in x) \vee (u = y)$. If $u \in x$ then $u \notin u$ and if $u = y$ then also $u \notin u$. If $x \cup \{y\} \in x \cup \{y\}$ then $(x \cup \{y\} \in x) \vee (x \cup \{y\} = y)$. If $x \cup \{y\} \in x$ then because $\Phi(x)$, there would hold $x \cup \{y\} \notin x \cup \{y\}$: a contradiction. And if $x \cup \{y\} = y$, there would hold $y \in y$ which is impossible since $\Phi(y)$; also a contradiction. Therefore $x \cup \{y\} \notin x \cup \{y\}$ and $\Phi(x \cup \{y\})$. By induction, $(\forall x)\Phi(x)$ which is $(\forall x)(x \notin x) \wedge [(\forall u)(u \in x \Rightarrow u \notin u)]$. □

There are other classical formulations of a foundation axiom. One is $(\forall x)[x \neq \emptyset \Rightarrow (\exists y)(y \in x \wedge y \cap x = \emptyset]$. Another one, in [Hrbacek and Jech, 1999], states in words, that every set is well founded. In turn, a set is well founded if its transitive closure is transitive and well founded which in turns implies that there is a binary relation R on the set such that every nonempty subset has an R-minimal element.

The foundation axiom as given here is a consequence of both formulations. It will be shown that the transitive closure of an object is an object which is finite in the classical sense. It is also a theorem of TO that all objects can be linearly ordered (corollary 6) and that for each order, an object has a minimal element (theorem 2). Hence objects satisfy these two formulations of foundation.

Some of the theorems follow arguments from [Vopěnka, 1979] and also from arguments in [Andreev and Gordon, 2006] with the additional considerations on levels and contextual formulae. Others are close to classical nonstandard properties restricted to objects. The resulting merger belongs to neither, especially theorems such as theorem 2 (below page 14) which state properties true for any object.

Proposition 2 (Specification schema [2]). *Let $\Phi(u)$ be a **c**-formula.*

$$(\forall x)(\exists z)(\forall u)[u \in z \Leftrightarrow u \in x \wedge \Phi(u)]$$

[2] In ZFC, this axiom is also called "comprehension schema". The name "specification" is chosen here because it is felt that it conveys well the fact that a subobject can be defined by specifying a property.

Proof. By induction on subobjects (corollary 1)

Let $\Psi(w)$ be the formula $(w \subseteq x) \wedge (\exists z \subseteq w)(\forall u)(u \in z \Leftrightarrow \Phi(u))$. Then, obviously, $\Psi(\emptyset)$. Assume $\Psi(w)$ and let $y \in x$. If $\Phi(y)$ then set $z' = z \cup \{y\}$ otherwise $z' = z$. Hence $(\exists z' \subseteq w \cup \{y\})(\forall u)(u \in z' \Leftrightarrow \Phi(u))$. Hence $\Psi(w \cup \{y\}$. By corollary 1, this shows that $\Psi(x)$ i.e. $(\exists z \subseteq x)(\forall u)(u \in z \Leftrightarrow \Phi(u))$. □

This object is in fact unique by extensionality:

Corollary 2.

$$(\forall x)(\exists! z)(\forall u)[u \in z \Leftrightarrow u \in x \wedge \Phi(u)]$$

Proof. The existence of such an object is proven by proposition 2. Assume there exist z and z' satisfying the property. Then $u \in z' \Leftrightarrow u \in z$, hence by extensionality $z = z'$. □

By specification (as in ZFC) the following are immediate.

Proposition 3. $(\forall x)(\forall y)(\exists! z)(\forall u)(u \in z \Leftrightarrow (u \in x \wedge u \in y))$

This unique object is called: "the intersection of x and y" and is denoted by $x \cap y$.

Proposition 4.

$$(\forall x)(\forall y)(\exists! z)(\forall u)(u \in z \Leftrightarrow u \in x \wedge \neg(u \in y))$$

This unique object is denoted by $x \setminus y$.

The inclusion $x \subseteq y$ is defined as usual by $u \in x \Rightarrow u \in y$. If $x \subseteq y$ and $\neg(y \subseteq x)$, then the inclusion is strict, written $x \subset y$. If $x \subseteq y$ then x is said to be a subobject of y.

Proposition 5 (Existence of a union object)**.**

$$(\forall x)(\forall y)(\exists! z)(\forall u)[u \in z \Leftrightarrow u \in x \vee u \in y]$$

Proof. By induction on subobjects (corollary 1). Let $\Phi(v)$ be the statement

$$(v \subseteq y) \wedge (\exists! z)(\forall u)[u \in z \Leftrightarrow (u \in x) \vee (u \in v)]$$

There holds $\Phi(\emptyset)$ (take $z = x$). Assume $\Phi(v)$, hence there is a unique z that satisfies $(\forall u)[u \in z \Leftrightarrow (u \in x) \vee (u \in v)]$ and let $t \in y$. Then $u \in z \cup \{t\} \Leftrightarrow (u \in x) \vee (u \in v \cup \{t\})$, hence $\Phi(v \cup \{t\})$ and by corollary 1, $\Phi(v)$. □

Since this object is unique, it can be given a name: "the union object" and is denoted by $x \cup y$.

Proposition 6 (Replacement schema)**.** *Let Φ be a **c**-formula.*

$$(\forall u)(\exists! v)\Phi(u, v) \Rightarrow [(\forall x)(\exists z)(\forall v)(v \in z \Leftrightarrow (\exists u \in x) \wedge \Phi(u, v)]$$

Replacement ensures the existence of an object containing all Φ-images of elements of x – for contextual formula Φ. It could be called: "contextual replacement schema".

Proof. By proposition 1 (induction).

Let $\Psi(x)$ be the formula $(\exists z)(\forall v)(v \in z \Leftrightarrow (\exists u \in x) \wedge \Phi(u,v))$. First, $\Psi(\emptyset)$ is trivially true.

Assume $\Psi(x)$ and let y be arbitrary. By initial assumption, there is a unique w such that $\Phi(y,w)$. Let z satisfy $(\forall v)(v \in z \Leftrightarrow (\exists u \in x) \wedge \Phi(u,v)$.

Then $(\forall v)(v \in z \cup \{w\} \Leftrightarrow (\exists u \in x \cup \{y\}) \wedge \Phi(u,v))$, hence $\Psi(x \cup \{y\})$ and $(\forall x)\Psi(x)$. □

An object z containing all subobjects of x is said to be the (unique) object containing all parts of x.

Proposition 7 (Existence of an object containing all parts)**.**

$$(\forall x)(\exists! z)(\forall y)(y \in z \Leftrightarrow y \subseteq x)$$

Proof. By proposition 1 (induction).

Let $\Phi(x)$ be $(\exists! z)(\forall u)(u \in z \Leftrightarrow u \subseteq x)$.

$\Phi(\emptyset)$ is obvious. Assume $\Phi(x)$ and let y be arbitrary. Let z be the unique object satisfying $(\forall u)(u \in z \Leftrightarrow u \subseteq x$. There holds $(\forall u)(\exists! v)(v = u \cup \{y\})$, hence by replacement principle there is an s such that $(\forall v)(v \in s \Leftrightarrow (\exists u \in z) \wedge (v = u \cup \{y\}))$. Then $(\forall u)(u \in z \cup \{s\} \Leftrightarrow u \subseteq x \cup \{y\})$ which implies $\Phi(x \cup \{y\})$. By proposition 1, $(\forall x)(\exists! z)(\forall u)(u \in z \Leftrightarrow u \subseteq x)$. □

The unique object whose elements are exactly all subobjects of an object x is denoted by $\mathcal{P}(x)$.

The definition of ordering is as in other theories: a relation that is reflexive, antisymmetric and transitive. An order is contextual (a **c**-order) if it is defined by a **c**-formula. (Ordering using references to levels for example, would not necessarily define objects. Consider the order $a \prec b \Leftrightarrow a \sqsubset b$. Then, as will be seen, $\{x \mid x \prec b\}$ would be a level and this is not an object. It is nonetheless possible that an ordering defined by a non contextual formula could be redefined by a contextual formula.)

Theorem 2 (Existence of a minimal element and of a maximal element)**.** *Let $\preceq$ be a **c**-ordering of nonempty object u.*

$$(\exists y \in u)(\forall x \in u)\ (x \preceq y \Rightarrow y \preceq x)$$

$$(\exists z \in u)(\forall x \in u)\ (z \preceq x \Rightarrow x \preceq z)$$

Proof. By induction: For $v \subseteq u$, let $\Phi(v)$ be the formula $(\exists y \in u)(\forall x \in v)(x \preceq y \Rightarrow y \preceq x)$. Obviously, $\Phi(\emptyset)$. Assume $\Phi(v)$ i.e., $(\exists y \in u)(\forall x \in v)(x \preceq y \Rightarrow y \preceq x)$ and let $z \in u$ be arbitrary. If $z \preceq y$ then set $y' = z$ and $(\exists y')(\forall x \in v \cup \{z\})(x \preceq y' \Rightarrow y' \preceq x)$ otherwise leave $y' = y$. Then $(\exists y')(\forall x \in v \cup \{z\})(x \preceq y' \Rightarrow y' \preceq z)$, hence $\Phi(x \cup \{z\})$. Therefore $(\forall v)\Phi(v)$ and in particular $\Phi(u)$.

By reversing the order, any ordered object also has a maximal element. □

By the object successor axiom, if x and y are objects then $\{x\}$ and $\{y\}$ are also objects. Then by the union theorem $\{x\} \cup \{y\} = \{x, y\}$ is an object as well.

As $\{\{x\}\}$ and $\{\{x, y\}\}$ are objects then $\{\{x\}\} \cup \{\{x, y\}\} = \{x, \{x, y\}\}$ is an object.

Let (x, y) be the ordered pair $\{\{x\}, \{x, y\}\}$.

Proposition 8 (Existence of the Cartesian product)**.**

$$(\forall x)(\forall y)(\exists! z)((u, v) \in z \Leftrightarrow (u \in x \wedge v \in y))$$

Proof. For $u \in x$ and $v \in y$, the object $\{u, v\}$ is in $\mathcal{P}(x \cup y)$, hence the pair (u, v) is in $\mathcal{P}(\mathcal{P}(x \cup y))$.

Then, by specification, $z = \{(u, v) \in \mathcal{P}(\mathcal{P}(x \cup y)) \mid u \in x \wedge v \in y\}$. Uniqueness is immediate. □

Let A and B be two objects. Then the collection $\{(a, b) \mid (a \in A) \wedge (b \in B)\}$ is an object called the Cartesian product of A and B denoted by $A \times B$.

> The proof given above is classical in the sense that it holds in ZFC or any system where the power set exists, either by axiom or by proof.
>
> The existence of the Cartesian product can also be shown by double induction. This proof is not available in ZFC since the induction axiom does not hold for all sets.

Proof. The version of induction given by corollary 1 is used.

Let x and y be objects. For $z \subseteq y$ the statement $\Phi(z)$ is $(\exists w)((a, b) \in w \Leftrightarrow (a \in x \wedge b \in z)$. For $\Phi(\emptyset)$, the statement holds with $w = x$. Let y be arbitrary.

Claim: $(\exists w)((a, y) \in w \Leftrightarrow a \in x)$

By induction on elements of x. For $s \subseteq x$ the statement $\Psi(s)$ is that there is a pair with one element in s and the other is y i.e., $(\forall z \in s)\exists(z, y)$. Then $\Psi(\emptyset)$ holds. Assume $\Psi(t)$ and let $x' \in x$. The pair (x', y) is clearly defined. As $t \cup \{x'\} \subseteq x$, there holds $(\forall z \in t \cup \{x'\})\exists(z, y)$. By induction, $(\forall t \subseteq x)\Psi(t)$, hence $\Psi(x)$ therefore $\{(a, y) \mid a \in x\}$ exists. This ends the proof of the claim.

Now assume $\Phi(z)$ i.e., $(\exists w_z)((a, b) \in w_z \Leftrightarrow (a \in x \wedge b \in z)$. As for arbitrary y it has been shown that $w_y = \{(a, y) \mid a \in x\}$ exists, set $w = w_z \cup w_y$. Then $(a, b) \in w \Leftrightarrow (a \in x \wedge b \in z) \vee (a \in x \wedge b = y)$, hence $(a, b) \in w \Leftrightarrow (a \in x \wedge b \in z \cup \{y\})$ which shows $\Phi(z \cup \{y\})$. By induction $(\forall z \subseteq y)\Phi(x)$, hence $\Phi(y)$. □

Definition 3. *A* **c***-formula* $\Phi(x, z)$ *is a unary functional relation from object* u *to object* v *if*

$$(\forall x \in u)(\forall z \in v)(\forall z' \in v)[\Phi(x, z) \wedge \Phi(x, z') \Rightarrow z = z' \ ].$$

Thus for any x, there is a unique z satisfying $\Phi(x, z)$. It is noted $z = f(x)$ and the functional relation can be represented by the object containing all pairs $f = \{(x, f(x)) \mid x \in u \wedge f(x) \in v\}$.

The object z of all f-images is the image object which is included in the range – which may not necessarily be an object. This is also denoted $f : x \to z$.

The range of a function, if it is not an object, is characterised by the property that all outputs satisfy.

A function $f : u \to v$ is a subobject of the cartesian product:

$$f = \{(x, y) \in u \times v \mid y = f(x)\}$$

Proposition 9.

$$(\forall x)(\exists z)(\forall u)\ [u \in z \Leftrightarrow (\exists w \in x)(u \in w)]$$

Proof. Let $\Phi(x)$ be $(\exists z)(\forall u)(u \in z \Leftrightarrow (\exists w \in x)(u \in w))$. Assume $\Phi(x)$ and let y be arbitrary. Let z satisfy $(\forall u)(u \in z \Leftrightarrow (\exists w \in x)(u \in w))$. Then $(\forall u)(u \in z \cup \{y\} \Leftrightarrow (\exists w \in x \cup \{y\})(u \in w))$, hence $\Phi(x \cup \{y\})$ which implies by induction $(\forall x)\Phi(x)$. □

The elements of the elements of a given object x form an object called the union of x denoted $\bigcup x$.

A functional relation can be any type of construction on objects. Replacement ensures that for a given input, the image objects exist. The following theorem shows that induction can be used to prove the existence of objects satisfying an inductive relation.

Let $F(x, \varphi)$ stand for: "there exists a functional relation φ whose domain is $x \cup \{x\}$." Formally this can be written

$$(\exists \varphi)(\forall u)(\exists v)[(u, v) \in \varphi \Leftrightarrow u \in x \cup \{x\} \wedge ((\forall v')(u, v') \in \varphi \Rightarrow v = v')].$$

Since the domain of φ is assumed to be an object, φ itself is a subobject of a cartesian product and is thus an object, hence quantification over such functional relations is possible.

Let $C(x, \varphi)$ be a contextual statement about x and a functional relation φ with domain $x \cup \{x\}$. Then $F(x, \varphi) \wedge C(x, \varphi)$ means that there is a functional relation with domain $x \cup \{x\}$ satisfying a property defined by C.

Theorem 3 (Definition by induction on objects)**.** *Let a be an object and $\Phi(x)$ be the statement $(\exists! \varphi)[(\varphi(\emptyset) = a) \wedge F(x, \varphi) \wedge C(x, \varphi)]$.*

If $\Phi(\emptyset)$ and $\Phi(x) \wedge \Phi(y) \Rightarrow \Phi(x \cup \{y\})$, then there is a unique functional relation f such that $f(\emptyset) = a$ and $(\forall x)[F(x, f) \wedge C(x, f)]$.

Proof. Since $\Phi(x) \wedge \Phi(y) \Rightarrow \Phi(x \cup \{y\})$, by induction, there holds $(\forall x)\Phi(x)$ i.e., $(\forall x)(\exists! \varphi)$ $[(\varphi(\emptyset) = a) \wedge F(x, \varphi) \wedge C(x, \varphi)]$.

For a given x, denote the unique function satisfying $C(x, \varphi)$ by φ_x.

Set $f(\emptyset) = a$ and $f(x) = \varphi_x(x)$.

Since $(\forall x)(\varphi_x(\emptyset) = a)$, $f(\emptyset)$ is uniquely defined. Similarly $\varphi_x(x)$ is uniquely defined, hence $f(x)$ is unique. Therefore $(f(\emptyset) = a) \wedge (\forall x)F(x, f) \wedge C(x, f)$. □

A transitive object y is an object such that $z \in y \wedge u \in z \Rightarrow u \in y$. The transitive closure of x is a transitive object containing x which is contained in any transitive object containing x i.e., the transitive closure is a transitive set containing x which is minimal for inclusion: it contains x, the elements of x, the elements of the elements of x (as above), the elements of these, and so on.

The object $\{\emptyset, \{\{\emptyset\}\}\}$ has a transitive closure which is

$$\{\{\emptyset, \{\{\emptyset\}\}\}, \emptyset, \{\{\emptyset\}\}, \{\emptyset\}\}.$$

Proposition 10 (Existence of a transitive closure)**.**

$$(\forall x)[(\exists! y)(x \in y \wedge (\forall u)(\forall v)(u \in y \wedge v \in u \Rightarrow v \in y) \\ \wedge [(\forall w)(x \in y \wedge (\forall u)(\forall v)(u \in w \wedge v \in u \Rightarrow v \in w) \Rightarrow y \subseteq w]$$

Proof. By definition by induction (theorem 3)

The property $C(x, \varphi_x)$ defining the function is

$$\varphi_x(x) = \{x\} \cup \left(\bigcup_{t \in x} \varphi_x(t) \right)$$

Let $u \in \varphi_x(x)$ and $v \in u$. If $u = x$ then $v \in x$. Since $\varphi_x(v) = \{v\} \cup \left(\bigcup_{t' \in v} \varphi_x(t') \right) \subset \varphi_x(x)$, there holds $v \in \varphi_x(x)$. Otherwise, $u \in \{u\} \cup \left(\bigcup_{t' \in u} \varphi_x(t') \right)$ for some $u \in \varphi_x(x)$, then $v \in u \Rightarrow \{v\} \subseteq \varphi_x(u)$ as above, hence $v \in \varphi_x(x)$.

For $x = \emptyset$ the statement holds for $\varphi_\emptyset(\emptyset) = \{\emptyset\}$. Uniqueness and minimality are clear.

Assume $\varphi_x(x)$ and $\varphi_y(y)$ exist. Set $w = \{x \cup \{y\}\} \cup \varphi_x(x) \cup \varphi_y(y)$.

Assume $u \in w$ and $v \in u$. If $u \in \varphi_x(x)$ then $v \in \varphi_x(x)$, hence $v \in w$. Similarly for $u \in \varphi_y(y)$. If $u = x \cup \{y\}$ then $u \in w$ by construction. Then either $v \in x$ which has already been discussed, or $v = y$ and since $y \in \varphi_y(y)$, there holds $v \in w$. Uniqueness and minimality follow from the construction. Hence there is a function $\varphi_{x \cup \{y\}}$ satisfying C. By induction, for each x there is such a function, hence by theorem 3, $(\exists! f)[f(\emptyset) = \{\emptyset\} \wedge F(x, f) \wedge C(x, f)]$. □

The unique image of x by f is the transitive closure of x denoted by $\mathrm{Tc}(x)$.

Now that some classical properties of collections have been established, links with levels are made.

Theorem 4 (Closure)**.** *Let $(\exists x)\Phi(x)$ be a* **c***-formula and $a_1, \ldots, ak$ are all the free variables of Φ.*

$$(\exists x)\Phi(x) \Rightarrow (\exists^{a_1, \ldots, a_k} x)\Phi(x)$$

Proof. Assume $(\exists x)\Phi(x)$. Then there is a v such that $x \sqsubseteq v$, hence

$$(\exists^{a_1, \ldots, a_k, v} x)\Phi(x).$$

This is a contextual formula since the list of parameters is unchanged and the list of level variables has an extra object. Transfer yields $(\exists^{a_1, \ldots, a_k, v} x)\Phi(x) \Leftrightarrow (\exists^{a_1, \ldots, a_k} x)\Phi(x)$. □

If an object is uniquely defined by a property, then this object appears at the level of at least one of the parameters.

If f is a functional relation, then for a given x, $f(x)$ has x as a parameter. Let a be such that $f, x \sqsubseteq a$. Then, by closure, $(\exists y)(f(x) = y) \Rightarrow (\exists^a y)(f(x) = y)$. Since this y is unique, then $f(x) \sqsubseteq a$.

Lemma 1.

$$\mathcal{P}(x) \sqsubseteq x$$

Proof. The object $\mathcal{P}(x)$ containing all parts of a given object x (proposition 7) is defined with x as only parameter. Hence, by closure, $(\exists z)(\forall u)(u \in z \Leftrightarrow u \subseteq x) \Rightarrow (\exists^x z)(\forall u)(u \in z \Leftrightarrow u \subseteq x)$ therefore $\mathcal{P}(x) \sqsubseteq x$. □

Theorem 5 (Element theorem)**.** *Let S and a be objects.*

$$(\forall a)(\forall S)\ [(S \sqsubseteq a)\ \Rightarrow (\forall x)\ (x \in S \Rightarrow (x \sqsubseteq a))]$$

Proof. This theorem is a consequence of refinement (axiom 6). If $S = \emptyset$ the theorem is trivially true. Otherwise, first assume $S \sqsubseteq a$. Then

$$(\exists^a u)(\forall y)(\exists x \in u)[y \in S\ \Rightarrow\ y = x]$$

(take $u = S$). The formula "$y \in S \Rightarrow y = x$" does not use level quantifiers, so it is a **c**-formula. Hence by refinement (contrapositive) (axiom 6) applied to the **c**-formula "$y \in S\ \Rightarrow\ y = x$", there holds $(\forall y)(\exists^a x)[y \in S \Rightarrow y = x]$ which implies $(\forall y)(y \in S \Rightarrow y \sqsubseteq a)$. □

As for any S there holds $S \sqsubseteq S$ and the element theorem is true for any a such that $S \sqsubseteq a$, then

$$(\forall x)(x \in S \Rightarrow (x \sqsubseteq S)).$$

Proposition 11 (Inclusion principle)**.**

$$(\forall x)(\forall y)(x \subseteq y \Rightarrow x \sqsubseteq y)$$

Proof. By lemma 1: $\mathcal{P}(y) \sqsubseteq y$. By the element theorem, $x \in \mathcal{P}(y) \Rightarrow x \sqsubseteq \mathcal{P}(y)$. By transitivity $x \sqsubseteq y$. □

Theorem 6 (Element theorem: converse)**.** *Let S and a be objects.*

$$(\forall a)(\forall S)\ [(S \sqsubseteq a)\ \Leftarrow (\forall x)\ (x \in S \Rightarrow (x \sqsubseteq a))]$$

Proof. Conversely: assume $(\forall x)(x \in S \Rightarrow x \sqsubseteq a)$. Then $\forall y\ \exists^a x\ (y \in S \Rightarrow x = y)$. By refinement $\exists^a u\ \forall y\ \exists x \in u\ (y \in S \Rightarrow x = y)$ which implies $S \subseteq u$, hence $S \sqsubseteq u$ by inclusion principle. As $u \sqsubseteq a$ the conclusion is that $S \sqsubseteq a$ by transitivity. □

In the following, the theorem of elements will be referred to whether for its direct form or the converse.

Theorem 7 (Comparability of levels)**.**

$$(\forall x)(\forall y)(x \sqsubseteq y \vee y \sqsubseteq x)$$

Proof. For any arbitrary x and y: Let $S = \{x, y\}$. By the element theorem: $x \sqsubseteq S \wedge y \sqsubseteq S$ and by closure: $(\exists^{x,y} S)(\forall u)(u \in S \Leftrightarrow u = x \vee u = y)$. Hence $S \sqsubseteq x \vee S \sqsubseteq y$ (see notation remarks page 5).

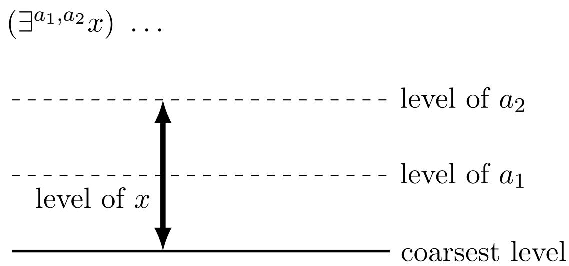
Assume $S \sqsubseteq x$. As $y \sqsubseteq S$, then by transitivity, $y \sqsubseteq x$. Otherwise $S \sqsubseteq y$ and in that case the conclusion is that $x \sqsubseteq y$. Hence $x \sqsubseteq y \vee y \sqsubseteq x$. ☐

In (F)RIST, idealisation has a restriction to finite sets. In TO there is no such restriction on objects. It follows that the inclusion property, which is true only for finite sets in (F)RIST, is true, by refinement, for all objects in TO.

Consequences of comparability

As mentioned before, $x \sqsubseteq a_1, \ldots, a_k$ stands for $x \sqsubseteq a_1 \vee \cdots \vee x \sqsubseteq a_k$

With comparability, $\neg(x \sqsubseteq y) \Rightarrow y \sqsubset x$, hence $x \sqsubseteq a_1, \ldots, a_k$ can also be understood as saying that x is not finer than the finest of $a_1, \ldots, a_k$. A mental representation can be provided by the following drawing:



Since all objects are comparable, in a contextual formula there is a coarsest level at which all of its free variables appear. This is called **the context level** of the formula, or the observation level of the formula. Any finer level is **a** context level.

Proposition 12 (Existence of a coarsest level of observation)**.**

$$(\forall x)\ (\emptyset \sqsubseteq x)$$

Proof. For any x there holds $\emptyset \subseteq x$. By inclusion principle (proposition 11) $\emptyset \subseteq x \Rightarrow \emptyset \sqsubseteq x$. ☐

The discreteness axiom states that there is a next finer level and an immediate previous level. If y is of the immediate coarser level than x, then the notation $y \sqsubset_- x$ is used and if y is of the immediate finer level, the notation is $x \sqsubset_+ y$. The immediate finer level and the immediate coarser level are unique as shown below.

Proposition 13.

$$(\forall x)(\forall y)(\forall u)(y \sqsubset_- x) \wedge (u \sqsubset_- x) \Rightarrow y \sqsupseteq\sqsubseteq u$$

and

$$(\forall x)(\forall y)(\forall u)(x \sqsubset_+ y) \wedge (x \sqsubset_+ u) \Rightarrow y \sqsupseteq\sqsubseteq u$$

Proof. For the next finer level. Assume $x \sqsubset_+ y$. Then $x \sqsubset y$ and if $x \sqsubset u$ then, by axiom 4.2, $u \sqsubseteq y$. Hence if one assumes also $x \sqsubset_+ u$, there holds $y \sqsubseteq u$. The same argument starting with u leads to $u \sqsubseteq y$. Essentially, the same argument holds for the immediate coarser level. □

An important consequence of refinement is the existence of objects of a finer level than any fixed level. In particular, not everything is of coarsest level.

Theorem 8.

$$(\forall a)(\exists y)\ (a \sqsubset y)$$

Proof. Fix a. By foundation, since $x \notin x$, then $u \in x \Rightarrow u \notin u$ and there holds $x \in u \Rightarrow x \neq u$, hence

$$(\forall^a u)(\exists y)(\forall x \in u)(y \neq x)$$

(take $y = u$). The formula "$y \neq x$" is contextual since it contains no level quantifier. Therefore by refinement (axiom 6):

$$(\exists y)(\forall^a x)(y \neq x).$$

so $y \not\sqsubseteq a$ and by comparability $a \sqsubset y$. □

This also implies that there is no "ultimate" level: there is no vantage point from which any object would be observable.

The refinement axiom shows that there are objects which are not of the coarsest level (theorem 8) but also that all the elements of an object are as visible as the object.

A singleton has same visibility as its element:

$$x \sqsupseteq\sqsubseteq \{x\}$$

$x \sqsubseteq \{x\}$ is an immediate consequence of the element theorem. It is also a consequence of closure. For the converse, consider $(\exists A)(\forall u)(u \in A \Rightarrow u = x)$. The only parameter of this statement is x, hence the level of this statement is given by x and by closure $(\exists^x A)(\forall u)(u \in A \Rightarrow u = x)$ therefore $\{x\} \sqsubseteq x$.

The following theorem generalises the fact that there is an element which has exactly the same visibility as the object. This element is called a witness of the level of the object, or witness for short. Note that it is not unique: it is even possible that all elements of an object have exact same visibility.

Corollary 3 (Existence of a witness)**.**

$$(\forall S)[S \neq \emptyset \Rightarrow (\exists x \in S)(S \sqsupseteq\sqsubseteq x)]$$

Proof. The theorem of elements shows that for all $x \in S$, there holds $x \sqsubseteq S$. The existence of a witness requires therefore to show that there is an element w such that $S \sqsubseteq w$.

By contradiction on the theorem of elements. Assume there is an object S such that $(\forall x \in S)(x \sqsubset S)$. (Contradiction is possible since there is comparability.) Then for a of immediate coarser level i.e., $a \sqsubset_{-} S$, there holds $(\forall x \in S)(x \sqsubseteq a)$. By the element theorem, this implies $S \sqsubseteq a$ which in turn implies $S \sqsubseteq a \sqsubset S$, hence $S \sqsubset S$: a contradiction. Therefore there is a w in S such that $S \sqsubseteq w$. As the element theorem states that $w \in S \Rightarrow w \sqsubseteq S$, there holds $w \sqsupseteq\sqsubseteq S$. □

The axioms of TO are based on very modest metaphysical assumptions. The idea behind this being that they should be acceptable to most. The question arose whether there were any grounds for assuming that any two objects should be comparable; this however is now shown to be a theorem. Another question was whether levels are dense or discrete. Zlatoś observed that since there seems to be no grounds for deciding the structure of levels it might be better to avoid specifying what this structure is – in a way similar to RIST. However, a strong point in favour of discreteness is the witness theorem. When working in (F)RIST it seems clear that a witness exists in all cases that can be checked. Yet the proof that such a witness exists in the case of density (one of the possibilities for FRIST) is extremely indirect: it uses a conservative extension of FRIST called GRIST and since it is true in the extension, it is true in FRIST, yet no proof has been found in FRIST itself. In TO the immediate link between discreteness and the existence of a witness can be considered reasonable grounds to accept the metaphysical assumption behind the discreteness axiom.

Proposition 14 (Pair theorem)**.**

$$(\forall x)(\forall y)\ (x \sqsubseteq (x,y)\ \textit{and}\ \ y \sqsubseteq (x,y))$$

Proof. By comparability (theorem 7) $x \sqsubseteq y$ or $y \sqsubseteq x$. By the element theorem $\{x\} \sqsubseteq (x,y)$ and $x \sqsubseteq \{x\}$ therefore $x \sqsubseteq (x,y)$. Similarly $y \sqsubseteq \{x,y\}$ and $\{x,y\} \sqsubseteq (x,y)$, hence $y \sqsubseteq (x,y)$. □

Proposition 15. *The extension of all x's such that*

$$(x \sqsubseteq a)\ ,\ (a \sqsubseteq x)\ ,\ (x \sqsubset a)\quad \textit{or}\quad (a \sqsubset x)$$

do not form objects.

Formulae such as $(x \sqsubseteq a)$ are not **c**-formulae. Nonetheless, it is worth questioning whether they can define objects (perhaps by other means).

Proof.

- Assume $x \in E \Leftrightarrow (x \sqsubseteq a)$.

  By the element theorem this implies $E \sqsubseteq a$, hence $E \in E$ which contradicts foundation.

- Assume $x \in F \Leftrightarrow (a \sqsubseteq x)$.

  By the element theorem $x \sqsubseteq F$, hence $a \sqsubseteq x \sqsubseteq F$ and therefore $F \in F$. Again a contradiction.

- Assume $x \in G \Leftrightarrow (x \sqsubset a)$. G has a witness x such that $G \sqsupseteq\sqsubseteq x$. Then $G \sqsupseteq\sqsubseteq x \sqsubset a$, hence $G \sqsubset a$ and $G \in G$.

- Assume $x \in H \Leftrightarrow (a \sqsubset x)$. H has a witness x, which implies $a \sqsubset x \sqsupseteq\sqsubseteq H$, hence $H \in H$.

□

All objects of a given observation level do not form an object yet there are objects containing all objects of a given level (and necessarily some others of finer levels).

Theorem 9.

$$(\forall a)(\exists y)(\forall^a x)(x \in y)$$

Proof. Let a be arbitrary. $(\forall^a u)(\exists y)(\forall x \in u)[x \in y]$. Take $u = y$. Hence, by refinement, $(\exists y)(\forall^a x)[x \in y]$. □

Note that proposition 15 and theorem 9 together imply that such an object y is strictly finer than a. When considering objects as visible as a it is possible to consider such an object y as a referential containing all objects as visible as a and yet some.

> It is also true in (F)RIST that whatever the level, there is a "finite" set containing all elements of that level, and more.

With respect to an ordering, if an object contains arbitrarily large members of level a, then it overflows into a finer level of observation.

Theorem 10 (Overflow)**.** *Let $\prec$ be a total (strict) **c**-ordering of object u.*

$$(\forall^a x \in u)(\exists^a y \in u)(x \prec y) \Rightarrow (\exists y \in u)\ (a \sqsubseteq y)$$

Proof. By refinement. It needs to be shown that the following holds:

$$\forall^a v\ \exists y\ \forall x \in v\ [y \in u \land (x \in u \Rightarrow x \prec y)].$$

Take $w = \{x \in v \mid x \in u\}$. Since w has a maximal element, consider $\max(w)$. By hypothesis, $\exists^a y \ \max(w) \prec y$. But then $\exists y \ \max(w) \prec y$ and refinement applies. Therefore

$$\exists y \ \forall^a x \ (y \in u \wedge (x \in u \Rightarrow x \prec y)).$$

y is strictly greater than all a-observable members of u, which implies $a \sqsubset y$. □

This of course implies $a \sqsubset u$ (by the element theorem or the existence of a witness).

> This theorem, with theorem 9, can be seen as corresponding in flavour to Vopěnka's prolongation axiom [Vopěnka, 1979]. He explains it by the following metaphor: "Imagine that we find ourselves on a long straight road with large stones set at regular distances. The stones are numbered by natural numbers. [...] The prolongation axiom assures us [...] that beyond the horizon, there is a stone S such that the stones between the 0-th stone and S form [an object]."
>
> Even though numbers have not yet been defined for TO, this can be understood as claiming that if a collection of elements seems to reach arbitrarily close to the horizon, then it overflows beyond the horizon and there is an object which also overlaps the horizon containing maybe not all such elements, but at least all those that are observable – plus an overlap.

Proposition 16 (a-density)**.** *Let $\prec$ be a linear* ***c****-ordering of object u.*

$$(\forall^a x \in u)(\forall^a y \in u)(x \prec y \Rightarrow (\exists^a z \in u)(x \prec z \prec y) \Rightarrow [\exists z \in u \ (a \sqsubset z)]$$

If between any two a-observable members of an object there is another a-observable member, then there is also a member less visible in between them.

Proof. For any fixed $x \in u$ apply overflow to elements y such that $x \prec \ y$. □

As above, this implies $a \sqsubset u$.

If in the formula used in the refinement axiom y satisfies a property which is always true, as $y = y$ for instance, then

$$(\forall^a u)(\exists y)(\forall x \in u)[(\Phi(x) \wedge (y = y)] \Leftrightarrow (\exists y)(\forall^a x)[\Phi(x) \wedge (y = y)]$$

the condition on y can thus be omitted, hence

Proposition 17.

$$(\forall^a u)(\forall x \in u) \ \Phi(x) \Leftrightarrow (\forall^a x) \ \Phi(x)]$$

If a property is satisfied by any element of any object of a given visibility then the property is satisfied by all objects of that visibility even though their collection does not form an object.

Chapter 3

Ordinals and Cardinals

> Some major properties of objects are drawn from the classical definition of ordinals : there are no limit ordinals, all objects can be linearly ordered, there are choice functions and all objects are "finite" in the sense defined by Russel (there is no one-to-one correspondence between an object and any of its strict subobjects). Even though it is shown that there are no infinite ordinals nor limit ordinals it is proven that there are "ultralarge" (finite) ordinals.

The classical definition of ordinals and the corresponding sequence are used:

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots$$

with the successor relation $\mathcal{S}(\alpha) = \alpha \cup \{\alpha\}$ (written $\alpha + 1$) and the corresponding order relation $(\alpha < \beta) \Leftrightarrow (\alpha \in \beta)$. If α and β are ordinals, then $\alpha \subset \beta \Leftrightarrow \alpha \in \beta$. To express the fact that x is an ordinal, the notation $\mathrm{Ord}(x)$ is used. Remark that $\mathrm{Ord}(x)$ is an $\in$-formula (see e.g. Krivine [Krivine, 1972], page 53), hence it is a **c**-formula.

> The collection of all ordinals cannot be a set. This is a well known fact of ZFC set theory known as the Buralli-Forti paradox. The same holds in TO.

By proposition 11 it is immediate that $\alpha \sqsubseteq (\alpha + 1)$. The following also holds:

Proposition 18. *Let α be an ordinal. Then $(\alpha + 1) \sqsupseteq\!\sqsubseteq \alpha$.*

Proof. $\alpha + 1 = \alpha \cup \{\alpha\}$. If $x \in \alpha \cup \{\alpha\}$ either $x \in \alpha$ or $x = \alpha$. In both cases, $x \sqsubseteq \alpha$, hence $(\alpha + 1) \sqsubseteq \alpha$. □

This implies that there is no greatest ordinal of a given visibility.

Proposition 19. *Ordinals of a given level of visibility do not form an object.*

Proof. A collection of ordinals of level β is totally ordered, hence if it were an object, by theorem 2 it would have a maximal element, say α. But $(\alpha + 1) \sqsubseteq \alpha$ so α is not maximal: a contradiction. □

Another way to prove the theorem is to observe that $u = \{x|\,\mathrm{Ord}(x) \wedge (x \sqsubseteq \alpha)\}$ would be an ordinal and the element theorem would imply that $u \sqsubseteq \alpha$, hence $u \in u$, also a contradiction.

Nonetheless, by adding the condition that y and x be ordinals, theorem 9 shows that there is an ordinal y such that for any ordinal x, there holds $(x \sqsubseteq \alpha) \Rightarrow x \in y$. Necessarily $\alpha \sqsubset y$.

Definition 4. *An ordinal N is α-large if*

$$(\forall x)\ [\mathrm{Ord}(x) \wedge (x \sqsubseteq \alpha) \Rightarrow x < N]$$

Relative to a given observation level, such an ordinal may be referred to as **ultralarge**.

Proposition 20 (Existence of ultralarge ordinals)**.** *Let α, β and γ be ordinals.*

$$(\forall \alpha)(\exists \beta)(\forall \gamma)[(\gamma \sqsubseteq \alpha) \Rightarrow \gamma < \beta]$$

Proof. Claim:

$$(\forall^{\alpha} u)\ (\exists \beta)\ (\forall \gamma \in u)\ [\mathrm{Ord}(\beta) \wedge (\mathrm{Ord}(\gamma) \Rightarrow \gamma < \beta)].$$

Let $v = \{x \in u \mid \mathrm{Ord}(x)\}$ which is a **c**-formula. This object has a maximal element m. Take $\beta = m + 1$. This shows that the claim holds.

Hence, by refinement,

$$(\exists \beta)\ (\forall^{\alpha} \gamma)\ [\mathrm{Ord}(\beta) \wedge (\mathrm{Ord}(\gamma) \Rightarrow \gamma < \beta)].$$

□

This implies $\alpha \sqsubset \beta$. As an immediate consequence, if α and β are ordinals, $\alpha \sqsubset \beta$ implies that β is α-large.

The existence of ultralarge ordinals shows that relative to some observation level, there are ordinals which are too large to be observable. The following proposition asserts that relative to any level, there *is* an observable ordinal.

Proposition 21.

$$(\forall x)(\exists^{x} \alpha)\,\mathrm{Ord}(\alpha)$$

Proof. By induction. Let $\Phi(x)$ be the statement $(\exists^{x} \alpha)\,\mathrm{Ord}(\alpha)$. It has one free variable (x) which appears as level variable, hence the formula is contextual. $\Phi(\emptyset)$ is clear. Assume $\Phi(x)$ and $\Phi(y)$ i.e., there are α and β such that $\alpha \sqsubseteq x$ and $\beta \sqsubseteq y$. If $\alpha \sqsubseteq \beta$ then $\beta \sqsubseteq x \cup \{y\}$ otherwise $\alpha \sqsubseteq x \cup \{y\}$. In both cases one can conclude that $(\exists^{x \cup \{y\}} \gamma)\,\mathrm{Ord}(\gamma)$, hence $\Phi(x \cup \{y\})$ and the statement holds for all x. □

In set theory, a set u with the property that $\emptyset \in u$ and $x \in u \Rightarrow x \cup \{x\} \in u$ is called an inductive set (see e.g. Hrbacek-Jech's textbook "Introduction to Set Theory" [Hrbacek and Jech, 1999]). The next theorem shows that there is no inductive set in TO.

Theorem 11.

$$\neg(\exists u)[(\emptyset \in u) \wedge (\forall x)(x \in u \Rightarrow x \cup \{x\} \in u)]$$

Proof. By contradiction. Assume such an object u exists. It is ordered by $x \prec y \Leftrightarrow x \in y$. But then it has no maximal element, since if m is this maximal element then $m \prec m \cup \{m\} \in u$ which contradicts theorem 2. Hence such an object does not exist. □

There is no "limit" ordinal in the sense it has in ZFC.

In fact, every ordinal has a predecessor.

Corollary 4. *Let α be an ordinal.*

$$(\forall\beta)[\mathrm{Ord}(\beta) \wedge (\beta \neq \emptyset)] \;\Rightarrow (\exists u)(\mathrm{Ord}(u) \wedge (\beta = u \cup \{u\})$$

Proof. Every ordinal β has a greatest element $u < \beta$. It is also an ordinal. Then $u+1 \geq \beta$ since otherwise it would not be the greatest. $u+1$ cannot be strictly greater than β since that would imply $u = \beta$. Hence $\beta = u + 1$. □

Every nonzero ordinal has a predecessor. It is immediate – by the same argument as above – that except for the coarsest level, there is no least ordinal of a given visibility.

A consequence of corollary 4 is induction on ordinals.

Corollary 5 (Induction on ordinals)**.** *Let Φ be a* **c***-formula and α an ordinal.*

$$[\Phi(\emptyset) \wedge [(\Phi(\alpha) \Rightarrow \Phi(\alpha+1)]] \Rightarrow (\forall\alpha)\Phi(\alpha) \qquad (*)$$

Proof. By contradiction. Suppose that (*) holds but that there is a β such that $\Phi(\beta)$ does not hold. Then $u = \{\alpha \in \beta \mid \neg\Phi(\alpha)\}$ is an object since Φ is a **c**-formula. Then let γ be the least element of u. As γ is an ordinal (and is not $\emptyset$) it has a predecessor, and as (*) holds, $\Phi(\gamma)$ also holds: a contradiction. □

This induction is, of course, the transposition of classical induction.

The notation $f : x \xrightarrow{1-1} y$ stands for "f is a one-to-one correspondence from x to y". It can be formally written

$$(f \subset x \times y) \wedge [(\forall u \in x)\ (\exists! v \in y)((u,v) \in f) \wedge (\forall v' \in y)(\exists u' \in x)((u',v') \in f) \wedge ((u,v) \in f \wedge (u,v') \in f \Rightarrow v = v')].$$

It is thus defined by a **c**-formula.

Proposition 22. $(\forall x)(\exists\alpha)(\exists f)(\mathrm{Ord}(\alpha) \wedge f : x \xrightarrow{1-1} \alpha)$

Proof. By proposition 1 (induction). Let $\Phi(x)$ be the **c**-formula: $(\exists\alpha)(\exists f)(\mathrm{Ord}(\alpha) \wedge f : x \xrightarrow{1-1} \alpha)$. Trivially, $\Phi(\emptyset)$. Assume $\Phi(x)$ and let y be arbitrary (wlog assume $y \notin x$). Let f be such that $f : x \xrightarrow{1-1} \alpha$. Then set

$$\widehat{f} : u \mapsto \begin{cases} f(u) & \text{if } u \in x \\ \alpha + 1 & \text{if } u = y \end{cases}$$

Then $\widehat{f} : x \cup \{y\} \xrightarrow{1-1} \alpha + 1$. Therefore $\Phi(x \cup \{y\})$, hence by induction the statement is true for any object. □

Corollary 6 (Existence of a linear ordering)**.** *Every object can be linearly ordered.*

Proof. Let x be an object and α an ordinal such that $f : x \xrightarrow{1-1} \alpha$. Then x is linearly ordered by (for u and v in x) $u \preceq v \Leftrightarrow f(u) \leq f(v)$. □

A choice function $g : x \to x$ is a function such that $\forall u \subset x$ with $u \neq \emptyset$, there holds $g(u) \in u$.

Corollary 7. *For any object, there exists a choice function.*

Proof. Consider a linear ordering of x. Then $(\forall u \subset x)(g(u) = \min(u))$ defines a choice function. □

Proposition 23. *Let α and β be ordinals. If there exists a one-to-one correspondence $f : \alpha \to \beta$, then $\alpha = \beta$.*

Proof. By induction on ordinals (corollary 5). It is clear that $f : \emptyset \xrightarrow{1-1} \beta$ implies $\beta = \emptyset$. Assume $f : \alpha \xrightarrow{1-1} \delta \Rightarrow \alpha = \delta$ and that there is a function $f : \alpha + 1 \xrightarrow{1-1} \beta$. First, if $f(\alpha) \notin \alpha$, then $f \restriction \alpha \to \gamma \subset \beta$ implies $\gamma = \alpha$, hence $f(\alpha) = \alpha$ and $\beta = \alpha + 1$. Second, if $f(\alpha) = \gamma \in \alpha$, then set

$$g : x \mapsto \begin{cases} f(x) & \text{if } x \in \alpha \text{ and } x \neq f^{-1}(\alpha) \\ \gamma & \text{if } x = f^{-1}(\alpha) \\ \alpha & \text{if } x = \alpha \end{cases}$$

then $g : \alpha + 1 \xrightarrow{1-1} \beta$ is a mapping as in the first situation, from which it follows that $\alpha + 1 = \beta$. By induction on ordinals, it is true for all ordinals. □

The major difference with ZFC is, of course, that since in TO all ordinals have a predecessor, this proof extends to *all* ordinals whereas in ZFC it is true only for finite ordinals.

Proposition 24. *Let x be an object and α and β ordinals. If there exist one-to-one correspondences $f : x \to \alpha$ and $g : x \to \beta$, then $\alpha = \beta$.*

There is a unique ordinal in one-to-one correspondence with a given object.

Proof. Assume there are ordinals α and β and one-to-one correspondences f and g such that $f : x \to \alpha$ and $g : x \to \beta$, then $f \circ g^{-1} : \beta \to \alpha$ is a one-to-one correspondence, hence $\alpha = \beta$. □

Theorem 12. *There is no one-to-one correspondence between an object and one of its proper sub-objects.*

Proof. Assume there is a one-to-one function $f : x \xrightarrow{1-1} w \subsetneq x$. Then there are unique ordinals α and β such that $g : x \xrightarrow{1-1} \alpha$ and $h : w \xrightarrow{1-1} \beta$.

But then $\alpha \xrightarrow{1-1} x \xrightarrow{1-1} w \xrightarrow{1-1} \beta$ and by proposition 23 $\alpha = \beta$ but then $w = h^{-1}(\beta) = h^{-1}(\alpha) = g^{-1}(\alpha) = x$. □

A proof using the induction theorem is also possible.

> Theorem 12 states that all objects are in fact "finite" in the sense given by Russel [Russel, 1903]. Since the transitive closure of an object is also an object (proposition 10), hence Russel-finite, it is also possible to conclude that no object can be "infinitely" deep. All descent in elements of elements, elements of these, etc., ends by the atomic object.

Definition 5. *The cardinal of an object x, noted $|x|$ is the unique ordinal which can be in one-to-one correspondence with the object x.*

Proposition 25. *Every ordinal is a cardinal.*

Proof. $\emptyset$ is a cardinal. Assume there is an ordinal which is not a cardinal. Then there is a least such ordinal. It has a predecessor α which is a cardinal. Consider any object x which has that cardinal and add any element y to the object extending the mapping by $\alpha \mapsto y$, thus $x \cup \{y\}$ has cardinal $\alpha + 1$: a contradiction. Hence all ordinals are cardinals. □

As mentioned before, ordinals do not altogether form an object, hence cardinals do not altogether form an object.

Proposition 26. *Let x be an object.*

$$|x| \sqsubseteq x$$

Proof. By closure. □

Note that $x \sqsubseteq |x|$ is not necessarily true. If $x = \{1, N\}$ then $|x| = 2$ even if N is 2-large.

A consequence is that if $|x|$ is a-large for some a then $a \sqsubset x$ and, by the witness theorem, x contains at least one member which is less visible than a

Definition 3 is extended to cases where the objects satisfying the input condition are too numerous to be collected into a set.

Definition 6. *A* **c***-formula* $\Phi(x, z)$ *is a one argument functional relation if*

$$(\forall x)(\forall z)(\forall z')[(\Phi(x, z) \wedge \Phi(x, z') \Rightarrow z = z')\ ].$$

For a functional relation, $y = f(x)$ is a notation for $\Phi(x, y)$. In this case it is interpreted as meaning that for any given x, there is a corresponding y.

> Following [Wildberger, 1996] it is not necessary to be able to contemplate the extension of all possible inputs to define a function. The domain, here, is a condition which the inputs must satisfy, such as "x is an ordinal".

Proposition 27. *Let* $R(x, y)$ *be a well* **c***-ordering such that all* x *satisfy a condition* $D(x)$ *but the extension of all such* x *do not define an object. Then there is a functional relation* J *such that* $D(x) \Rightarrow (\exists y)(\mathrm{Ord}(y) \wedge J(x, y)$ *and* $x \prec x' \wedge J(x, y) \wedge J(x', y') \Rightarrow y < y'$*.*

Proof. Let $\alpha = J(x)$ be the functional relation given by "$\mathrm{Ord}(\alpha)$ and α is isomorphic to the segment $S_x(R)$." This is a functional relation since for any x such that $D(x)$, there is a unique ordinal isomorphic to $S_x(R)$. It is an isomorphism which respects the R order on D and the natural order on ordinals: if $x \prec y$ then $S_x(R)$ is an initial segment of $S_y(R)$. Since there is an isomorphism from $S_x(R)$ to $J(x)$, then $J(x) < J(y)$ and since J is a one-to-one correspondence, J^{-1} is also a one-to-one correspondence. The domain of J^{-1} is not an object since the domain of J is not an object. Hence the domain of J^{-1} is an initial segment of ordinals which is not an object. By previous theorems, an initial segment of ordinals which has a greatest element would be an object, hence the domain of J^{-1} has no greatest element.

This means that any ordinal is the image of some x such that $D(x)$. □

> In ZFC there is no collection of all ordinals. The phrase "all ordinals" should therefore not be used any more than the phrase "all sets". Yet the set-theoretical belief that all extensions exist lead to abuses in applying the idea that any property can be used to define sets, such as Krivine's [Krivine, 1972] conclusion to the theorem corresponding to the one above: "c'est donc On tout entier" (where On stands for ordinals).

Even though for any ordinal there is an object which has same observability (in particular: the ordinal itself) and for any object there is an ordinal which has same visibility, it is not possible to *number* levels by indexing them on ordinals.

The statement $(\exists k_0)(\exists k_1)(k_0 \sqsubset_+ k_1)$ is a contextual formula. It is immediate that $(\exists k_0)(\exists k_1)(\exists k_2)(k_0 \sqsubset_+ k_1 \sqsubset_+ k_2)$ is also a contextual formula, and so forth. It states that there are two (three) consecutive levels. This statement is true for $n = 3$ consecutive levels. Assuming the statement to be true for n consecutive levels, it is clear that it is true for $n + 1$ consecutive levels. Hence it is true for all n, including ultralarge n.

For every n, there is an nth level. The converse would be that for every object, there is an n such that this object is of the nth level. The statement that for any x, there is an n such that x has the exact same visibility as the nth level would require to define some function which assigns to each object the number of its level. If such a function existed the inverse image of a given ordinal would be an object; but the inverse image of an ordinal would also be a level which cannot be an object: a contradiction.

It is possible to write about objects of the coarsest level, of the next finer level and so on for "naive" counting numbers. It is the generalisation to assuming that "$(\forall x)(\exists n)$ x is in the nth level" which fails.

3.1 Addition

When considering operation on ordinals, the notation $0 = \emptyset$, $1 = \{\emptyset\}$, etc, is used.

Addition is defined in the usual way: If a and b are two ordinals, then two new objects of ordered pairs are formed, $A = \{(x, 0) \mid x \in a\}$ and $B = \{(y, 1) \mid y \in b\}$, then $a + b = |A \cup B|$. Clearly $(a + b) \sqsupseteq\sqsubseteq a, b$.

The following hold.

Let a, b and c be ordinals.

(1) $(a \sqsubseteq c) \wedge (b \sqsubseteq c) \;\Leftrightarrow\; (a + b \sqsubseteq c)$.

(2) $(c \sqsubset a) \;\Rightarrow\; (c \sqsubset a + b)$.

(3) $(a + b \sqsubseteq c) \Rightarrow (a \sqsubseteq c) \;\wedge (b \sqsubseteq c)$.

(4) $(c \sqsubset a + b) \;\Rightarrow\; (c \sqsubset a) \vee (c \sqsubset b)$.

3.2 Subtraction

For every $n > m$, there exists a unique k such that $m + k = n$. Such a k is denoted by $n - m$.

3.3 Multiplication

Multiplication is given by the cardinal of the cartesian product: For x and y, consider $X \times Y = \{(u, v) \mid u \in x, v \in y\}$, then $x \cdot y = |X \times Y|$.

By applying the element theorem, it is immediate that $x \cdot y \sqsupseteq\sqsubseteq x, y$.

3.4 Exponentiation

A function $f : a \to b$ is an object containing ordered pairs (x, y) with $x \in a$ and $y \in b$ such that for each x, there is a unique pair (x, y).

The exponentiation a^b is defined as the cardinal number of the object whose elements are all functions $f : a \to b$ (it is easy to verify that such an object actually exists). Closure yields $b^a \sqsupseteq\sqsubseteq a, b$.

Chapter 4

Incomplete Objects

Certain properties $\varphi(x)$ defined by statements cannot be used to form objects by writing $\{x \mid \varphi(x)\}$ because the resulting collection would be too large. Objects may be extremely big, they may contain an ultralarge number of items, but some statements describe potentially infinitely many objects. This is what is intuitively meant by "too large". An example is that objects strictly less visible than the coarsest level cannot be all collected. Another (more classical) example is that ordinals cannot be collected, neither can objects be collected into an object. Some $\in, \sqsubseteq$-statements do not define objects yet they may nonetheless define a somewhat fuzzy subcollection of a well-defined object. Consider an ordinal α not of the coarsest level. The ordinals of coarsest level all belong to α yet they do not form an object (theorem 19). If the collection of all objects x such that $x \in A \wedge \Phi(x)$ does not form an object, then it will be said to form an **incomplete object**. An incomplete object does not necessarily satisfy the axioms or theorems which are statements about objects. For incomplete objects a different font is used whenever possible.

> These incomplete objects are somewhat similar to the semisets of Vopěnka [Andreev and Gordon, 2006, Vopěnka, 1979] and bare similarities also to fuzzy sets as given by Zadeh [Zadeh, 1965]. They are also similar to external sets of (F)RIST or IST.

Since they are not objects, the braces "{" and "}" will not be used. The collection (in the intuitive sense) of the x's of an incomplete object will be denoted with special delimiters " $\wr$ ".

Definition 7 (Incomplete Object)**.** *Let Φ be a $\in$-$\sqsubseteq$-formula and A an object.*

If $\neg(\exists x)(\forall u)(u \in A \Leftrightarrow \Phi(x))$ then $\mathsf{x} = \wr\, u \in A \mid \Phi(u)\, \wr$ and x is an incomplete object.

Paraphrasing [Koudjeti and van den Berg, 1995], an incomplete object is a collection of elements of an object when this collection disobeys at least one theorem about objects.

Examples:

$$\mathsf{a} = \wr\, x \mid x \sqsubseteq \alpha\, \wr, \ \mathsf{b} = \wr\, x \mid x \sqsupseteq\!\sqsubseteq \alpha\, \wr \text{ and } \mathsf{c} = \wr\, x \mid x \sqsubset \alpha\, \wr$$

are incomplete objects, since by theorem 9 they are contained in an object of a level finer than that of α, yet by theorem 4, they are not objects. The other properties mentioned in theorem 4, such as $\alpha \sqsubseteq x$, do not form incomplete objects as they cannot be contained in an object (they are the complement with respect to "everything").

Similarly, $\bigcirc^\alpha = \wr\ x \mid \text{Ord}(x) \wedge (x \sqsubseteq \alpha)\ \wr$ is an incomplete object. Clearly it has no maximal member.

The term **incomplete** comes from the fact that, being part of an object, it is possible to complete them by adding members in such a way that the extended collection becomes an object. The completion is not, in general, unique.

4.1 Levels of incomplete objects

Following the principle given by the element theorem, if all members of an incomplete object are of a given observation level, the level of that incomplete object is defined to be the level of its members. Hence $\bigcirc^\alpha \sqsubseteq \alpha$.

If $\bigcirc^\alpha_0 = \wr\ \text{Ord}(x) \mid x \sqsupseteq\sqsubseteq \alpha\ \wr$ then $\bigcirc^\alpha_0 \sqsupseteq\sqsubseteq \alpha$.

Every level is contained in an object of immediate finer level (theorem 36, below) hence levels are incomplete objects. In this sense, a reification of levels as "things" is acceptable. The writing $\mathbf{V}(0)$ (resp. $\mathbf{V}(x)$) may therefore be used to denote the incomplete object which is the coarsest level (resp. the coarsest level containing x). The next finer level after $\mathbf{V}(x)$ will be denoted by $\mathbf{V}_+(x)$ and the immediate coarser level before will, of course, be denoted by $\mathbf{V}_-(x)$ (if x is not of the coarsest level). In the case of a context level, $\mathbf{V}$ with no decorations is used.

It is not always possible to define the level of an incomplete object. For $0 \sqsubset N$, consider $A = \{a_n \mid n \leq N\}$ such that $a_n \sqsubset a_{n+1}$ and all a_n are ordinals. There is no coarsest level which contains A.

An incomplete object may have a one-to-one correspondence with a subpart of itself. Consider the one-to-one mapping $f : \bigcirc^\alpha \to \bigcirc^\alpha \setminus \{\emptyset\}$ by setting $f(x) = x \cup \{x\}$.

Proposition 28. *Let A be an object. $B \subset A$ is an object if and only if B has a cardinality.*

Proof. As A is an object, it has a cardinality (by theorem 24). Similarly for B, if it is an object.

For the converse, assume B has a cardinality. Let $\beta = |B|$. The fact that B has a cardinality means that there is an ordering f which is a one-to-one correspondence from B to β. Then $B = \{x \in A \mid (\exists\gamma \leq \beta)(x = f^{-1}(\gamma))\}$ is given by a $\mathbf{c}$-formula, hence B is an object (by specification). $\square$

4.2 Coarsest level

Let A be an object, and P a $\mathbf{c}$-formula. By specification, $\{x \in A \mid P(x)\}$ determines an object, hence incomplete objects may be formed only if $\in, \sqsubseteq$-statements which are not contextual are used. If an object A is of the coarsest level, referring to the $. \sqsubseteq .$ predicate will not enable any distinction between members of A, hence an object of coarsest level cannot contain an incomplete object.

More specifically, if $A \sqsubseteq \emptyset$ and Φ^u is a statement with level variable u, then $x \in A \wedge \Phi^u(x)$ is a **c**-statement, hence $\{x \in A \mid \Phi^u(x)\}$ is an object.

In particular, any collection of ordinals less than a given ordinal of coarsest level, forms an object.

> In AST, classes whose proper subclasses are all proper subsets are *defined* as finite [Vopěnka, 1979] hence are sets. In the theory of objects, this characterisation cannot be used to determine objects of the coarsest level yet, as shown above, it is true that objects of the coarsest level have no subcollections which are incomplete objects. Conversely, consider $A = \{a, b\}$ where a and b are not of coarsest level, any subcollection of A is an object. The transposition of Vopěnka's claim would be that if any object x is such that any subcollection $\{y \in x \mid \Phi(y)\}$ is an object – for any $\in, \sqsubseteq$-formula – then the cardinal of x is of the coarsest level.

Let $B \subset A$ stand for $x \in B \Rightarrow x \in A$ without implying that A or B be proper objects.

Theorem 13. *Let A be an object. Then every subcollection of objects of A forms an object if and only if the cardinal of A is of the coarsest level.*

Proof. Assume $|A| = \alpha \sqsubseteq \emptyset$ then there is a one-to-one correspondence $g : A \to \alpha$. As $\alpha \sqsubseteq \emptyset$ and $B \subset A$, then g maps B to a subcollection of α. Since α is of coarsest level, then the subcollection is an object, hence has a cardinality. Since g is one-to-one, this implies that B has a cardinal, hence (by proposition 28) B is an object.

Conversely, assume there is a $B \subset A$ which is not an object. Since A is an object, it has a cardinality α and the one-t-one mapping $g : A \to \alpha$ restricted to B defines a subcollection of α which is not an object. This implies that $\emptyset \sqsubset \alpha$. □

4.3 Incomplete objects and "Infinite Sets"

One might interpret that the proof that all objects satisfy classical definitions of finiteness is equivalent to the claim of a negation of infinity. This is not so. If one desires to believe in the existence of actually infinite quantities, in a way similar to the axioms of classes of the Alternative Set Theory, an extension of objects could be considered, only another name should be used to denote these. This justifies the claim of infinity-agnosticism.

Note that in ZF without the infinity axiom, there remain ambiguous situations. An infinite set could, in principle, be constructed if one admits unending processes. But this may be considered to beg the question: what is an unending process if not an infinite process? Are these unending processes accepted? This is a question answered not by the axioms of set theory but rather by what type of formula is accepted. This situation forces one to choose, in ZF, either between an axiom of infinity or its negation. In AST and in TO the induction axiom schema allows to extend the proof of "finiteness" of small objects to all objects without legislating on the length of admissible formulae. ZF without the assertion of the existence of infinity and without its negation is not powerful enough to perform such proofs. But this is not surprising since it is usually understood that ZFC is a theory about infinite sets, hence removing its founding axiom destroys most of the construction.

Considering blurry horizons of incomplete objects it is clear that this will conflict with the fundamental philosophy of ZFC that every collection should be sharp and precisely delineated. As ZFC and TO are both attempts at modelling and generalising the intuitive concept of number (among other things), it may be asked whether ordinals of the coarsest level of TO are the usual finite ordinals of ZFC. On the one hand, if this is so, then the ordinals of TO are a much larger collection, since there are ordinals of finer levels – though ordinals of finer levels (ultralarge) are not similar to limit ordinals. On the other hand, if one considers that all ordinals of TO are the finite ordinals of ZFC (with an extra distinction possibility), then ordinals of TO are a much smaller collection (even though they cannot be collected...) than the collection of ordinals (which does not exist either...) of ZFC, since in ZFC there are limit ordinals. In fact both views are equally acceptable and the permanent horizon shifting which takes place in TO indicates that a definite answer is not even desirable. It is certainly possible to find a model within ZFC for either of these interpretations, but as Vopěnka already noted for AST [Benešová et al., 1989], it is not possible to find a model in which the shifting from one view to another is possible.

(1) In ZFC and (F)RIST: Every infinite set has a one-to-one correspondence with a proper subset. Finite sets do not.

    This is the condition of finiteness given by Russel [Russel, 1903]. Theorem 12 shows that no object can be in one-to-one correspondence with a proper sub-object, hence all objects satisfy the condition for being finite in this sense.

    Incomplete objects can have a relation defined by a one-to-one rule where some elements are not reached. Incomplete objects may have properties assigned to infinite sets in ZFC.

(2) In ZFC and (F)RIST: An infinite ordered set does not necessarily have minimal or maximal elements. Finite ordered sets have maximal and minimal elements.

    It has been shown that all objects can be ordered and all have maximal and minimal members. Objects fit this criterion of finiteness.

    Incomplete objects, such as levels, may satisfy the definition of infinite sets.

(3) In ZFC and (F)RIST: Limit ordinals exist: they are ordinals which do not have a predecessor or some of their members do not have a predecessor.

    All ordinals in the theory of objects have predecessors: this is theorem 4. Even incomplete objects cannot produce the equivalent of a specific limit ordinal having no predecessor.

    Limit ordinals can intuitively be understood as "beyond the horizon" as can be the case for ordinals of finer levels. Still, there is a fundamental difference: even ultralarge ordinals (greater than some "infinite" incomplete objects) are finite in the sense of ZFC.

(4) In AST: A set which contains no proper sub-semisets is finite. Transposing the concept of semiset to incomplete set, theorem 13 states that an object which contains no incomplete object has a cardinal of the coarsest level. Hence objects which have a cardinal of a finer level could be considered infinite in the sense of AST.

Incomplete objects satisfy many properties of infinite sets of ZFC. The name "incomplete" carries the flavour of "infinite" in the etymological sense of *un-finite* or *un-finished*. Objects of the coarsest level agree with all classical definitions of finiteness.

If one decides to ignore classical characterisations of infinity, other conclusions can be met. Ordinals of the coarsest level can be huge. $10^{10^{10}}!$ would need around 10 million pages to be written, yet it is still of coarsest level. An ordinal of finer level is thus beyond bounds that can be expressed by writable numbers i.e., very very big. Hence one could assume that numbers of finer levels are either (1) finite numbers which have a flavour of infinity, or conversely, (2) infinite numbers which have the flavour of finiteness. Philosophically, there is nothing wrong in considering the second interpretation but it clashes with historical mathematical characterisations. In particular, it does not seem pertinent to try to establish a link with the concepts of denumerability and non-denumerability.

Language (especially mathematics) has the purpose of communicating with as little ambiguity as possible, hence it is reasonable to adopt the point of view that ordinals of finer levels are finite – with a flavour of infinity.

Potential infinity is less ambiguous. If the words "infinity" or "infinitely" are assumed to refer to their potentiality, then it is possible to say that there are infinitely many ordinals in the sense that there are more ordinals than observable. By transfer, there are more ordinals than at any observation level.

A rather curious conclusion arises from considering the classical characterisations of finite and infinite. Infinite collections are proper sub-parts of finite collections. Infinite can be smaller than finite.

"Infinity" is not intrinsically a question of size but rather a matter of structure.

Chapter 5

Integers and rationals

Rationals introduce the concept of ultrasmall number. Classical nonstandard results about the level of operations on integers and rationals are shown to hold in TO. The Ackermann coding is defined in TO, which shows a one-to-one correspondence between objects and ordinals. This correspondence respects ordering according to levels.

5.1 Similitude and Indiscernibility

Classically, numbers (integers, rationals, reals, etc.) are constructed using equivalence classes and saying that a given number is in fact a tag for a class and that two numbers are equal if they are (maybe different) tags of the same class. But the only way to check that they belong to the same class is to check that they satisfy a certain equivalence relation. Hence the claim that the fundamental concept is that of equivalence, not that of class. The equivalence class constructions of ZFC use the equal sign to mean "belongs to the same equivalence class as".

Similitude and indiscernibility are in some sense richer concepts than absolute equality. In a given context, different objects are considered as representing the same concept (similitude) or sufficiently similar (indiscernibility). This way of considering similitude and indiscernibility has been used by many authors already quoted, including [Andreev and Gordon, 2006]. This concept is related to the universe of the discourse. Two objects are similar (equivalent) if, within the context of the discourse, no difference between them can possibly be exhibited.

In the following, the concept of similitude will be used in places where, classically, equivalence classes would be introduced.

5.2 Integers

Integers are defined classically as ordered pairs of (finite) ordinals, the canonical form being $(n, 0)$ or $(0, n)$.

The notation $\mathbb{N}(x)$ stands for "$(\exists n)(\mathrm{Ord}(n) \wedge x = (n, 0))$" and $\mathbb{Z}(x)$ means that x is of the form (a, b) with $a = 0$ or $b = 0$. Then $\mathbb{N}(x) \Rightarrow \mathbb{Z}(x)$.

5.2.1 Addition:

Addition is defined as an extension of ordinal addition.

$$x + y = \begin{cases} (n+m, 0) & \text{if } x \text{ is of the form } (n,0) \text{ and } y \text{ is of the form } (m,0) \\ (0, n+m) & \text{if } x \text{ is of the form } (0,n) \text{ and } y \text{ is of the form } (0,m) \\ (n-m, 0) & \text{if } x \text{ is of the form } (n,0) \text{ and } y \text{ is of the form } (0,m) \\ & \text{and } n > m \\ (0, m-n) & \text{if } x \text{ is of the form } (n,0) \text{ and } y \text{ is of the form } (0,m) \\ & \text{and } m > n \end{cases}$$

The pair $(n, 0)$ is written as n or $+n$ and $(0, n)$ as $-n$.

5.2.2 Multiplication:

$$x \cdot y = \begin{cases} (n \cdot m, 0) & \text{if } x \text{ is of the form } (n,0) \text{ and } y \text{ is of the form } (m,0) \\ (0, n \cdot m) & \text{if } x \text{ is of the form } (n,0) \text{ and } y \text{ is of the form } (0,m) \\ 0, n \cdot m) & \text{if } x \text{ is of the form } (0,n) \text{ and } y \text{ is of the form } (m,0) \\ (n \cdot m, 0) & \text{if } x \text{ is of the form } (0,n) \text{ and } y \text{ is of the form } (0,m) \end{cases}$$

The distributive law of multiplication over addition is extended from ordinal addition and multiplication.

As there exist α-large ordinals (for any α), by the identification $n \hookrightarrow (n, 0)$, there exist α-large integers, and the definition of ultralargeness is extended to negative values by stating that there exist α-large negative integers $(0, n)$. Hence a number is α-large if it is larger in absolute value than any positive α-observable integer.

5.3 Rationals

A rational number is an ordered pair of integers (a, b) with $b > 0$. $\mathbb{Q}(x)$ indicates that x satisfies the condition of being a rational number.

Similitude between rationals (a, b) and (c, d) is defined by saying that $a \cdot d = b \cdot c$. Writing (a, b) as $\frac{a}{b}$ and (c, d) as $\frac{c}{d}$, similitude is noted $\frac{a}{b} = \frac{c}{d}$. Thus 3/2 and 6/4 are two different representations of the same quantity. [3] All other operations are defined in the usual way.

[3] The word "quantity" is used here in the informal way of everyday language, as defined by Webster: The attribute of being so much, and not more or less; the property of being measurable, or capable of increase and decrease, multiplication and division; greatness; and more concretely, that which answers the question "How much?"

An integer n is identified as a rational by the pair $(n, 1)$. Going all the way back to cardinals, the fraction $-3/2$ is thus $((0,3),(2,0))$.

Then, all the way back through the definitions:
$0 \equiv \emptyset$, $2 \equiv \{\emptyset, \{\emptyset\}\}$ and $3 \equiv \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, hence
$(0,3) = \{0, \{0,3\}\} \equiv \{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}\}$
$(2,0) = \{2, \{0,2\}\} \equiv \{\{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset, \{\emptyset\}\}\}\}$ and
$-\frac{3}{2} = ((0,3),(2,0)) = \{(0,3), \{(0,3),(2,0)\}\} =$

$$\{\{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}\}, \{\{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}\}, \{\{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset, \{\emptyset\}\}\}\}\}\}.$$

We believe very strongly that no one has ever, ever, considered that this is really what they think of when they summon $-3/2$. Axioms do not "create" or "construct" numbers such as $-3/2$ nor do they define what they "really are": rather they provide a coding, often much more complicated than the coded concept, which allows for logical proofs and internal consistency.

Proposition 29.
$$\mathbb{Q}(x) \wedge \mathbb{Q}(y) \wedge (y \neq 0) \Rightarrow \frac{x}{y} \sqsubseteq x, y$$

Proof. As rationals are pairs, the pair theorem (page 21) provides the required property. □

Definition 8. *A rational number $\frac{c}{b}$ is a-small if $\frac{c}{b} \neq 0$ and*
$$(\forall^a x > 0)\ \mathbb{Q}(x) \Rightarrow \left|\frac{c}{b}\right| < x$$

Theorem 14. *Relative to level a, there exist a-small rational numbers.*

Proof. Let N be an a-large whole number. Then $\frac{1}{x} > \frac{1}{N}$. Hence $\frac{1}{N}$ is a-small. □

If x is a-small or zero, then the notation
$$x \simeq_a 0$$
stands for "x is a-indiscernible from zero", and similarly x is a-indiscernible from y, written
$$x \simeq_a y$$
if $x - y$ is a-small or $x = y$. If $x \simeq_a y$ it may also be said that x and y are a-close.

Proposition 30. *If $x \sqsubseteq a$ and $y \sqsubseteq a$ and $x \simeq_a y$ then $x = y$.*

Proof. By closure, $x - y \sqsubseteq a$ and $x \simeq_a y \Rightarrow x - y \simeq_a 0$. As $x - y$ cannot be a-small, there holds $x - y = 0$. □

Definition 9 (Interval condition)**.** *If $a < x < b$ the notation $]a, x, b[$ is used and similarly for $[a, x, b]$, $]a, x, b]$ and $[a, x, b[$ for the open and closed conditions. To say that x is in an interval, it is always in the sense of being in between the bounds.*

Proposition 31. *An interval of rational numbers which is neither a singleton nor empty is not an object.*

Proof. By theorem 16, an interval of rationals – having the density property – would have no level of observation as it would indefinitely overflow into finer levels. □

Subscript Convention and Context Dependent Writing

When an observation level is clearly defined, the subscripts in $\simeq_a$ will not be written. Ultralarge (resp. ultrasmall) will stand for a-large (resp. a-small), where the level of a is the observation level. This observation level is the **context level** of a statement.

With this convention on levels, classical rules of analysis with levels (as proven in [Hrbacek et al., 2010b]) can be written in the form:

Proposition 32. *Relative to a context level, if $x \simeq a$ and $y \simeq b$, $\varepsilon \simeq 0$ and $\delta \simeq 0$ and let a and b be not ultralarge.*

Then

(1) $\varepsilon + \delta \simeq 0$

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

(3) $a \cdot \varepsilon \simeq 0$

(4) $a \cdot b \simeq x \cdot y$

(5) (In addition, if $b \not\simeq 0$) $\frac{a}{b} \simeq \frac{x}{y}$

(6) $a \cdot b \simeq 0 \Rightarrow a \simeq 0 \vee b \simeq 0$

Proof. (Observable, ultralarge and ultrasmall refer to the context level.)

For (6): wlog assume $a > 0$ and $b > 0$. If a and b are not ultrasmall, then $a > a_0$ and $b > b_0$ for some non zero observable a_0 and b_0. But then $a \cdot b > a_0 \cdot b_0$ which is observable by closure (and non zero), hence $a \cdot b \not\simeq 0$. By item (3), one of a or b must be ultrasmall. □

The advantage of this notation is that "illegal" formulae become very difficult to write by accident, as shown in [O'Donovan, 2009].

5.3.1 The Ackermann ordering

Let $\mathrm{ac}(x)$ be the Ackermann number of object x:

$$\mathrm{ac} : x \mapsto \begin{cases} 0 & \text{if } x = \emptyset \\ \displaystyle\sum_{y \in x} 2^{\mathrm{ac}(y)} & \text{otherwise} \end{cases}$$

For "small" objects, the Ackermann number can be computed by hand even though they become very quickly absurdly high since, for instance, the pair $x = \{3, \{3, 2\}\}$ has $\mathrm{ac}(x) = 2^{11} + 2^{2056}$.

Proposition 33. *The Ackermann function is a one-to-one correspondence between objects and ordinals.*

Proof. By induction on ordinals: for every ordinal there is an Ackermann-type function which maps it to a unique object.

Let $F(\alpha) =$

$$\max\{\beta \leq \alpha \mid (\forall \gamma \leq \beta)(\exists! x)(\exists! g_x : \{x\} \to \gamma + 1)(g_x(x) = \sum_{u \in x} 2^{g_x(u)} = \gamma)\} \quad (*)$$

By direct computation, it is clear that $F(0) = 0$, $F(1) = 1$, $F(2) = 2$ and so on. It needs to be shown that for any α, $F(\alpha) = \alpha$.

By contradiction: Assume there is an ordinal γ such that $F(\gamma) \neq \gamma$ i.e. there is an ordinal not greater than γ, yet greater than $\emptyset$, for which there is no Ackermann type function. Then for any $\beta > \gamma$, there holds $F(\beta) \neq \beta$. Let κ be the least ordinal such that $F(\kappa) \neq \kappa$. (The statement given by (*) is a **c**-formula, hence if there is an ordinal for which $F(\beta) \neq \beta$ then the collection of all ordinals β less than γ such that $F(\beta) \neq \beta$ is an object, and therefore has a least member.)

Let $C_{\kappa-1} = \{x \mid g_x(x) \leq \kappa - 1\}$. There holds $\kappa = \sum_{n=0}^{\kappa-1} a_n \cdot 2^n$ where a_n is zero or 1. This sum is unique. Since all n in the sum satisfy $n \leq \kappa - 1$, for each of these there is a unique object $x \in C_{\kappa-1}$ such that $g_x(x) = n$.

For $\sum_{n=0}^{\kappa-1} a_n 2^n = \kappa$ set $x_\kappa = \{x \in C_{\kappa-1} \mid (\exists i)(a_i = 1) \wedge (i = g_x(x))\}$. Then $g_{x_\kappa}(x_\kappa) = \kappa$: a contradiction. Hence $(\forall \alpha)(F(\alpha) = \alpha)$.

For any ordinal α, there is a unique object and a unique function as defined above such that $g_x(x) = \alpha$.

For the converse: By definition by induction (theorem 3). For every object, there is a unique ordinal and a unique function with the required property.

The statement $\Phi(x)$ is

$$\exists! \varphi : x \cup \{x\} \to \text{Ord} \text{ such that } \varphi(x) = \sum_{t \in x} 2^{\varphi(t)}$$

The statement is obviously true for $\emptyset$.

Assume $\Phi(x)$ and $\Phi(y)$ i.e., $\exists \varphi_x : x \cup \{x\} \to \text{Ord}$ and $\varphi_y : y \cup \{y\} \to \text{Ord}$. If $y \in x$ then $x \cup \{y\} = x$ and there is nothing to prove. Otherwise define φ by

$$\varphi : u \mapsto \begin{cases} \varphi_x(u) & \text{if } u \in x \cup \{x\} \\ \varphi_y(u) & \text{if } u \in y \setminus x \\ \varphi_y(y) & \text{if } u = y \\ \varphi_x(x) + 2^{\varphi_y(y)} & \text{if } u = x \cup \{y\} \end{cases}$$

there holds $\varphi : x \cup \{y\} \cup \{x \cup \{y\}\} \to \text{Ord}$ and $\Phi(x \cup \{y\})$ holds, hence by theorem 3, there is a unique function f such that for any x, there holds $f : x \mapsto \sum_{t \in x} 2^{f(t)}$. □

The one-to-one mapping from objects to ordinals defined above is denoted by $\text{ac}(x) = \sum_{u \in x} 2^{\text{ac}(u)}$ and is called the Ackermann function.

In ZFC it is an isomorphism between "hereditarily finite" sets and whole numbers.

Proposition 34.

$$(\forall x)\ \ \mathrm{ac}(x) \sqsupseteq\sqsubseteq x$$

Proof. The definition of the Ackermann function does not refer to levels and its parameters are of coarsest level, hence $\mathrm{ac} \sqsubseteq 0$. It is a one-to-one correspondence of coarsest level: closure applied both ways yields the result. □

The Ackermann numbering provides an ordering of objects which respects the ordering of levels in the following sense:

$$(\forall x)(\forall y)\ (x \sqsubset y) \Rightarrow \mathrm{ac}(x) < \mathrm{ac}(y)$$

This is straightforward from the fact that $\mathrm{Ord}(x)$, $\mathrm{Ord}(y)$ and $x \sqsubset y$ implies $x < y$.

An immediate consequence is

Proposition 35.

$$(\forall x)(\exists \alpha)(\mathrm{Ord}(\alpha) \wedge (x \sqsupseteq\sqsubseteq \alpha))$$

For any level, there is an ordinal which has exactly the same observability.

The result of theorem 9 can be made more specific:

Proposition 36.

$$(\forall \alpha)(\exists y)(\alpha \sqsubset_+ y)(\forall^\alpha x)(x \in y)$$

There is an object of next finer level containing all members of a given level.

Proof. Fix α and let $\alpha \sqsubset_+ x$. Let $y = \{u \mid \mathrm{ac}(u) \leq \mathrm{ac}(x)\}$. By closure $y \sqsupseteq\sqsubseteq x$, hence $\alpha \sqsubset_+ y$ and $u \sqsubseteq \alpha \Rightarrow u \in y$. □

Chapter 6

Numeric grains: a first exploration

Following Sochor [Benešová et al., 1989], reals can be considered to be approximations to rationals. Consider rationals whose squares are ultraclose to 2 and then allow our acuteness of vision to decrease. The individuality of rationals sink below the discernibility horizon and the blurry haziness which remains is called "a real number equal to $\sqrt{2}$". A similar concept is now developed but since it has properties which are not necessarily attributed to real numbers, the name "numeric grain" is used.

Let $\mathsf{V}(x)$ denote the level of x. As mentioned page 34 this is an incomplete object, but can also be used to indicate the observability of x: the notation $x \in \mathsf{V}(a)$ stands for "x is a-observable". The next finer level is $\mathsf{V}_+(a)$ and the inclusion symbol may be used : $\mathsf{V}(a) \subset \mathsf{V}_+(a)$ in the sense that any a-observable number is also observable at next finer level. The following construction extends some ideas of Davis [Davis, 1977] on equivalence classes given by the $\simeq$ relation.

Let V stand for the coarsest level and V_+ for the next finer level. The $\simeq$ relation (relative to V) is clearly an equivalence relation on V_+. For a given rational x,

$$\big\} \, u \in \mathsf{V}_+ \mid u \simeq x \, \big\}$$

is an incomplete object which is itself included in the incomplete object V_+. (It is an incomplete object since, for instance, there is no greatest u such that $u \simeq x$ yet it is included in an object.)

The collection of all such incomplete objects is given by

$$\mathsf{V}_{+}/_{\simeq}.$$

This is not an object nor even an incomplete object. Recall that incomplete objects are sub-collections of objects, hence the members of incomplete objects are objects. Here, the collection of numeric grains should be understood as a syntactic object.

If an equivalence class contains a rational number which is V-large, then all of its members are V-large.

The same holds if instead of the coarsest level and the next finer level, any arbitrary level and its next finer level are considered.

A rational number is a-limited if it is not a-large.

Let $\mathbf{V}_+^{\lim}(a)$ denote rationals of $\mathbf{V}_+(a)$ which are a-limited.

Definition 10 (Numeric grains of level a)**.** *An incomplete object of the form*

$$\wr\, u \in \mathbf{V}_+^{\lim} \mid u \simeq x \,\wr$$

is called a numeric grain of level a*.*

Such a collection is denoted by $\wr\, x \,\wr$. If $u \in \wr\, x \,\wr$ then $x \simeq u \Rightarrow x \in \wr\, u \,\wr$ and clearly $\wr\, u \,\wr = \wr\, x \,\wr$.

If x is a rational and $x \in \wr\, u \,\wr$, then x is said to *represent* the numeric grain $\wr\, u \,\wr$.

The absolute value of a numeric grain is obtained by considering the absolute value of its representatives. If a numeric grain is given by $\wr\, x \mid x \simeq_a u \,\wr$ then its absolute value is $\wr\, x \mid x \simeq_a |u| \,\wr$, denoted by $|\wr\, x \,\wr|$ and $|\wr\, x \,\wr| = \wr\, |x| \,\wr$.

Definition 11 (u-Rational numeric grains)**.** *If a numeric grain of level* u *contains a rational* x *of level* u*, then this grain is said to express the rational number* x*.*

It will be called a u*-rational numeric grain and* x *is its canonical representative.*

A u-rational numeric grain contains no other rationals of level u since by proposition 30 no other rational of level u can be ultraclose to x.

If a numeric grain of level u contains no rational of level u, then it is an u-irrational numeric grain.

Proposition 37. *For numeric grains* x *and* y *(of level* a*).* $(\forall t \in \wr\, x \,\wr)(\forall s \in \wr\, x \,\wr)(\forall u \in \wr\, y \,\wr)(\forall v \in \wr\, y \,\wr)$ *(for rational* t, s, u *and* v*)*

(1) $t + u \simeq s + v$

(2) $t \cdot u \simeq s \cdot v$

(3) If $v \not\simeq 0$, $t/u \simeq s/v$

Proof. The context level is given by a. Immediate by proposition 32. □

This shows that it is possible to define operations between numeric grains by considering operations between representatives. The following operations are thus well defined.

If $\wr\, x \,\wr$ is a numeric grain of level a and $0 \in \wr\, x \,\wr$, then $\wr\, x \,\wr = \wr\, 0 \,\wr$.

Definition 12. *Let* $\wr\, x \,\wr$ *and* $\wr\, y \,\wr$ *be of level* a

(1) $\wr\, x \,\wr + \wr\, y \,\wr = \wr\, x + y \,\wr$

(2) $\wr\, x \,\wr \cdot \wr\, y \,\wr = \wr\, x \cdot y \,\wr$

(3) If $u \in \wr\, y\, \wr \Rightarrow u \not\simeq 0$ *then* $\wr\, x\, \wr / \wr\, y\, \wr = \wr\, x/y\, \wr$

If necessary, the reference to observability can be indexed i.e., $\wr\, x\, \wr_a$ is a numeric grain of level a.

A numeric grain of level a can be expressed at a finer level (say, b) by

$$\wr\, u \in \mathbf{V}_+^{\lim}(b) \mid u \simeq_a x\, \wr$$

and denoted by $\wr\, x\, \wr_a^b$.

With this notation, $\wr\, x\, \wr_a^a$ is the same as $\wr\, x\, \wr_a$. If $b \sqsubset a$ then the numeric grain $\wr\, x\, \wr_a^b$ is reduced to a rational singleton.

Let $\wr\, x\, \wr_a^a$ be a numeric grain of level a, (where x is some rational number of $\mathbf{V}_+(a)$), then $\wr\, x\, \wr$ is *expressed* at level b with $a \sqsubset b$ by considering $\wr\, u \in \mathbf{V}_+(b) \mid u \simeq_a x\, \wr$. Let $a \sqsubset b$, and $\wr\, x\, \wr_a^a$ and $\wr\, y\, \wr_a^b$. If there is a $u \in \wr\, x\, \wr$ such that $u \in \wr\, y\, \wr$, then $\wr\, x\, \wr \subset \wr\, y\, \wr$ and they represent the same quantity since they are characterised by the same indiscernibility relation.

A numeric grain is a-large if it is greater in absolute value than any numeric grain of level a.

> Calculus with numeric grains is thus a form of interval calculus as described in [Koudjeti and van den Berg, 1995] yet it differs from neutrices and external numbers (from the same book) in many ways. The neutrix part of an external number is, in general, an infinite external set. External numbers are designed to allow a calculus of magnitudes hence, say, in a computation, an appreciable number will be replaced by the external set of all appreciable numbers. Yet similarities appear.

6.1 Ordering numeric grains

For numeric grains of a given level, say a, the ordering is straightforward.

As for equivalence classes in general, one has $\wr\, x\, \wr_a \cap \wr\, y\, \wr_a = \emptyset$ or $\wr\, x\, \wr_a = \wr\, y\, \wr_a$.

If $\wr\, x\, \wr \neq \wr\, y\, \wr$ and $(\exists t \in \wr\, x\, \wr)(\exists u \in \wr\, y\, \wr)(t < u)$, then $(\forall t \in \wr\, x\, \wr)(\forall u \in \wr\, y\, \wr)(t < u)$. Similarly for $t > u$. This justifies the following definition:

Definition 13 (Order on numeric grains)**.** *If* $\wr\, x\, \wr_a \cap \wr\, y\, \wr_a = \emptyset$ *then*

$$\wr\, x\, \wr_a < \wr\, y\, \wr_a \Leftrightarrow x < y$$

For numeric grains of different levels, say $a \sqsubset b$, it is only possible to show that for any $\wr\, x\, \wr_a^b$ and $\wr\, y\, \wr_b^b$, either $\wr\, x\, \wr_a^b < \wr\, y\, \wr_b^b$ or $\wr\, x\, \wr_a^b > \wr\, y\, \wr_b^b$ or, relative to the coarsest of the two levels, $\wr\, x\, \wr_a^b \simeq_a \wr\, y\, \wr_b^b$ in the case that $\wr\, y\, \wr_b^b \subset \wr\, x\, \wr_a^b$.

This last case comes from the fact that a numeric grain of finer level will be either strictly included in a numeric grain of coarser level or disconnected.

Definition 14 (Shadow)**.**

Let $0 \sqsubseteq a \sqsubset b \sqsubseteq c$. *If* $\rbrace\, y \,\lbrace_b^c$ *and* $\rbrace\, x \,\lbrace_a^c$ *are such that* $\rbrace\, y \,\lbrace_b^c \subset \rbrace\, x \,\lbrace_a^c$, *then* $\rbrace\, x \,\lbrace_a^c$ *is said to be an a-shadow of* $\rbrace\, y \,\lbrace_b^c$.

The shadow of a numeric grain is the incomplete object obtained by considering a coarser equivalence relation.

Proposition 38. *If a numeric grain is not a-large, then its a-shadow exists and is unique.*

Proof. By proposition 32, the shadow is unique if it exists.

Existence is given by observing that if $a \sqsubset b$ then $u \simeq_b x \Rightarrow u \simeq_a x$. □

The unique a-shadow of $\rbrace\, x \,\lbrace$ is denoted by $\mathsf{sh}_a \rbrace\, x \,\lbrace$.

At this point, numeric grains differ from real numbers. In (F)RIST a real number of a given level is also a real number of all finer levels. Here, a numeric grain is not "a point" which can be included in finer levels, but a blurry incomplete object which is bigger than – in the sense that it overlaps – numeric grains of finer levels. Hence, in general, if two numeric grains are ultraclose it cannot be said which is greatest.

The shadow is not a dimensionless point ultraclose to a number of finer level by being to one of its sides, but a grain which *contains* the numeric grain of a finer level.

With numeric grains, the inclusion is in fact in opposite direction than in (F)RIST.

Numeric grains of a given level have representatives in finer levels. These representatives are unique. They express the same equivalence relation only on different levels of rationals. They also have shadows in coarser levels (if they are not already of coarsest level) which are unique. These express coarser equivalence relations.

Proposition 39. *For* $\mathsf{sh}_a \rbrace\, x \,\lbrace$ *and* $\mathsf{sh}_a \rbrace\, y \,\lbrace$.

(1) $\mathsf{sh}_a \rbrace\, x \,\lbrace + \mathsf{sh}_a \rbrace\, y \,\lbrace = \mathsf{sh}_a(\rbrace\, x \,\lbrace + \rbrace\, y \,\lbrace)$

(2) $\mathsf{sh}_a \rbrace\, x \,\lbrace \cdot \mathsf{sh}_a \rbrace\, y \,\lbrace = \mathsf{sh}_a(\rbrace\, x \,\lbrace \cdot \rbrace\, y \,\lbrace)$

(3) If $0 \notin \rbrace\, y \,\lbrace$, *then* $\mathsf{sh}_a \rbrace\, x \,\lbrace / \mathsf{sh}_a \rbrace\, y \,\lbrace = \mathsf{sh}_a(\rbrace\, x \,\lbrace / \rbrace\, y \,\lbrace)$

Proof. By proposition 32. □

Proposition 40. *For shadows* $\mathsf{sh}_a \rbrace\, x \,\lbrace$ *and* $\mathsf{sh}_b \rbrace\, y \,\lbrace$ *for* $a \sqsubset b$.

(1) $\boldsymbol{sh}_a \wr x \wr + \boldsymbol{sh}_b \wr y \wr = \boldsymbol{sh}_a(\wr x \wr + \wr y \wr)$

(2) $\boldsymbol{sh}_a \wr x \wr \cdot \boldsymbol{sh}_b \wr y \wr = \boldsymbol{sh}_a(\wr x \wr \cdot \wr y \wr)$

(3) If $0 \notin \wr y \wr$, *then* $\boldsymbol{sh}_a \wr x \wr / \boldsymbol{sh}_b \wr y \wr = \boldsymbol{sh}_a(\wr x \wr / \wr y \wr)$

Proof. By proposition 32 and the fact that if $a \sqsubset b$, $x \simeq_b y \Rightarrow x \simeq_a y$. □

The metaphor here is that if two levels are used, the precision of the finer one is lost in the fuzziness of the coarsest.

> This is contrary to what one would expect of closure in (F)RIST in which the sum of two rational numbers is observable at the level of the finest, not necessarily at the level of the coarsest.

6.2 Groups, Rings and Fields

As mentioned several times before: some properties cannot be used to define objects because they characterise potentially infinitely many objects. In these cases, the properties cannot even define incomplete objects. It is nonetheless possible to consider objects satisfying a given property without considering the extension of all such objects.

> In ZFC one often proves that a property holds for *all* whole numbers by first assuming that it holds for *any* whole number. Here, we remain content with the observation that a given property will hold for *any* whole number (or integer or rational).

The classical definition of groups, rings and fields are extended:

Let $G(x)$ be a formula about x. We write $G[x, y, z]$ as short for $G(x) \wedge G(y] \wedge G[z]$

The pair $(G, *)$ defines a group if

(1) $G(x, y) \Rightarrow G(x * y)$

(2) $G(x, y, z) \Rightarrow x * (y * z) = (x * y) * z$

(3) $G(x) \Rightarrow (\exists e)G(e) \wedge (x * e = e * x = x)$

(4) $G(x) \Rightarrow (\exists y)G(y) \wedge (x * y = y * x = e)$

If the condition G defines an object, the classical definition holds.

Similarly, rings and fields are defined for objects satisfying the usual conditions.

> This usage follows common practice in nonstandard analysis as, for instance, in [Koudjeti and van den Berg, 1995], where neutrices are subgroups which are in fact external subsets.

Proposition 41. *Numeric grains of a given level form a field.*

Proof. For addition and multiplication defined as above (definition 12 or proposition 40), there is an identity for addition and an identity for multiplication):

Let x, y and u be rationals:

Relative to some context level,

$$(\forall x)(\forall y)(\forall u)(x \simeq y \wedge u \simeq 0 \Rightarrow x + u \simeq x \wedge y + u \simeq y).$$

It is thus legitimate, here, to quantify over numeric grains:

$$(\forall \wr\, x \,\wr)(\wr\, x \,\wr + \wr\, 0 \,\wr = \wr\, x \,\wr)$$

Similarly

$$(\forall \wr\, x \,\wr)(\wr\, x \,\wr \cdot \wr\, 1 \,\wr = \wr\, x \,\wr)$$

$\wr\, 1 \,\wr$ is multiplicative identity. $\wr\, 0 \,\wr$ is additive identity.

The numeric grain $\wr\, 0 \,\wr$ is unique and will be called the zero of the field. Similarly, $\wr\, 1 \,\wr$ is the unit and is called "one".

A numeric grain $\wr\, x \,\wr$ has an opposite with respect to addition which is $\wr\, -x \,\wr$. It is trivial that it contains all opposites of members of $\wr\, x \,\wr$ since $x \simeq y \Rightarrow -x \simeq -y$. The opposite is denoted $-\wr\, x \,\wr$.

A nonzero numeric grain $\wr\, x \,\wr$ has an inverse $\wr\, 1/x \,\wr$ and one also has that the inverse numeric grain contains all inverses of $\wr\, x \,\wr$. The inverse is denoted $1/\wr\, x \,\wr$.

By closure applied to rationals; addition, multiplication and division of numeric grains of a given level yield numeric grains of that level.

All other properties of fields are immediate consequences of proposition 32. ☐

Note that zero has no inverse as expected. Even though nonzero members of $\wr\, 0 \,\wr$ have inverses (they are rationals), $x \simeq 0 \wedge y \simeq 0 \Rightarrow x \simeq y$ but this does not imply that $1/x \simeq 1/y$. (For ultralarge whole number N consider $x = 1/N$ and $y = 1/N^2$. Thus the inverses of all ultrasmall rationals of $\mathbf{V}_+(a)$ are not necessarily a-close.)

Definition 15. *A a-cut $\in$-property $P \sqsubseteq a$ is such that $P(x) \wedge \neg P(y) \Rightarrow x < y$ and P has rational parameters only.*

Example:

$$(x < 0) \vee (x^2 \leq 2)$$

Theorem 15. *Let $P \sqsubseteq a$ be an a-cut property which holds for rational b and does not hold for some c. Then there is a $\wr\, d \,\wr_a$ such that $P(x) \Rightarrow x < \wr\, d \,\wr$ and $y > \wr\, d \,\wr \Rightarrow \neg P(y)$.*

Proof. By closure, there is an a-observable c such that $P(c)$ does not hold and an a-observable b such that $P(b)$ holds.

Take a-large whole number N in $\mathbf{V}_+(a)$. Set $b_0 = b$ and $b_k = b + k \cdot \frac{c-b}{N}$ for $0 \leq k \leq N$. All b_i's are in $\mathbf{V}_+(a)$. The collection $B = \{b_k \mid 0 \leq k \leq N\}$ is an object, hence there is a smallest $b_k \in B$ which is an upper bound. It is a representative of a numeric grain $\wr\, b_k \wr_a$. Clearly no numeric grain of level a can be a least upper bound. □

This means that property P has a least upper bound $\wr\, d \wr$, even if the extension of all numbers satisfying P do not form an object.

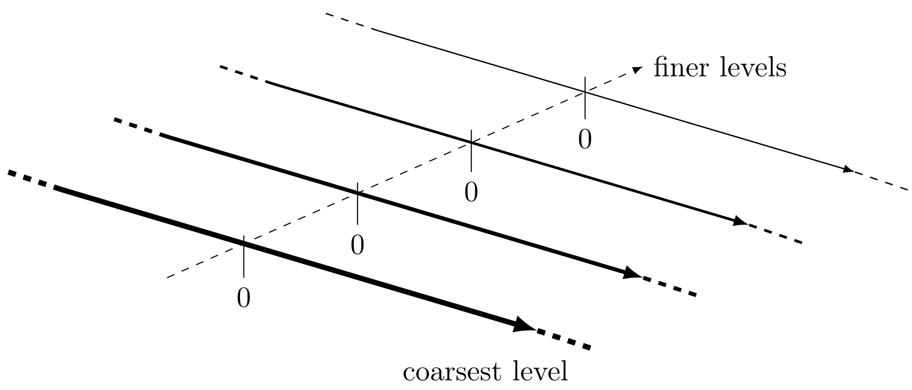
This means that numeric grains of level a seem complete.

Example

Let $P(x)$ be the property $x > 0 \wedge x^2 < 2$. This property satisfies $P \sqsubseteq 0$. For any $x > 2$ the property is false, hence it has an upper bound. $P(1)$ holds. The numeric grain $\wr\, x \wr$ of all positive rationals whose squares are ultraclose to 2 is a least upper bound of coarsest level. And $\wr\, x \wr^2 = \wr\, 2 \wr$.

> An interesting philosophical observation is that here the paradox of "real numbers" being points with no dimension and yet covering the whole line does not hold: numeric grains are close in some aspects to real numbers, but are not dimensionless points. Set on a geometric line, they have spatial extension. Hence a line is "covered" by numeric grains of a given level. It is also covered by numeric grains of finer levels.

A mental representation can be provided by the following drawing. Each line is a finer (complete) description of a given interval. Yet no line describes an interval completely.



6.3 An extension

Consider a deleted numeric grain [4] $\rbrace\, 0 \,\lbrace \setminus \{0\}$ – in the sense that it would contain all inverses of ultrasmall rationals – noted $\rbrace\, 0 \,\lbrace^*$ (or $\rbrace\, 0 \,\lbrace^*_a$ to indicate that it is the additive identity for level a.) Using the ∞ symbol to mean "beyond the horizon of the level", $\rbrace\, 1 \,\lbrace / \rbrace\, 0 \,\lbrace^*_a = \rbrace\, \infty \,\lbrace_a$ denotes the collection of all a-large rationals of $\mathbf{V}_+(a)$.

It will be called the inverse of zero but not "infinity", for obvious reasons.

The philosophical reason to consider that the entirety of what is beyond the a-horizon is a single grain is that the level a is the level of discernibility. One knows that the world continues beyond the horizon but, seen from level a, nothing more can be distinguished. Seen from the earth, a distant galaxy is but a point.

If context level is clear, $\rbrace\, \infty \,\lbrace$ with no index may be used.

$\rbrace\, 0 \,\lbrace^*$ is also an additive identity – even though the 0 symbol represents a rational which is *not* in the grain.

Note that $\rbrace\, \infty \,\lbrace$ contains both positive and negative ultralarge rationals. Writing $\rbrace\, 0_+ \,\lbrace^*$ (resp. $\rbrace\, 0_- \,\lbrace^*$) for all positive (resp. negative) ultrasmall rationals, then $\rbrace\, +\infty \,\lbrace = \rbrace\, 1 \,\lbrace / \rbrace\, 0_+ \,\lbrace^*$ (resp. $\rbrace\, -\infty \,\lbrace = \rbrace\, 1 \,\lbrace / \rbrace\, 0_- \,\lbrace^*$).

6.3.1 Operations in the extension.

Relative to a given level, let $\rbrace\, u \,\lbrace$ be nonzero and non inverse of zero.

The following properties are immediate consequences of proposition 32.

With respect to a given level:

- $\rbrace\, u \,\lbrace + \rbrace\, 0 \,\lbrace^* = \rbrace\, u \,\lbrace$
- $\rbrace\, u \,\lbrace \cdot \rbrace\, 0 \,\lbrace^* = \rbrace\, 0 \,\lbrace^*$
- $\rbrace\, u \,\lbrace + \rbrace\, \infty \,\lbrace = \rbrace\, \infty \,\lbrace$
- $\rbrace\, u \,\lbrace \cdot \rbrace\, \infty \,\lbrace = \rbrace\, \infty \,\lbrace$
- $\dfrac{\rbrace\, u \,\lbrace}{\rbrace\, 0 \,\lbrace^*} = \rbrace\, \infty \,\lbrace$
- $\dfrac{\rbrace\, u \,\lbrace}{\rbrace\, \infty \,\lbrace^*} = \rbrace\, 0 \,\lbrace^*$
- $\dfrac{\rbrace\, 0 \,\lbrace^*}{\rbrace\, u \,\lbrace} = \rbrace\, 0 \,\lbrace^*$
- $\dfrac{\rbrace\, \infty \,\lbrace}{\rbrace\, u \,\lbrace} = \rbrace\, \infty \,\lbrace$
- $\rbrace\, 0 \,\lbrace^* \cdot \rbrace\, 0 \,\lbrace^* = \rbrace\, 0 \,\lbrace^*$

The following results show that these "numbers" require special attention.

With respect to a given level $\mathbf{V}$:

[4] extending the concept of deleted interval.

- $\wr 0 \wr^* + \wr 0 \wr^* = \wr 0 \wr$ (since for ultrasmall rational x, both x and $-x$ are in $\wr 0 \wr^*$) whereas $\wr 2 \wr \cdot \wr 0 \wr^* = \wr 0 \wr^*$, hence $\wr 2 \wr \cdot \wr 0 \wr^* \neq \wr 0 \wr^* + \wr 0 \wr^*$.
- $\wr 0_+ \wr^* + \wr 0_+ \wr^* = \wr 0_+ \wr^*$
- $\wr 0_+ \wr^* + \wr 0_- \wr^* = \wr 0 \wr$
- Even though $\wr 1 \wr / \wr 0 \wr^* = \wr \infty \wr$ and $\wr 1 \wr / \wr \infty \wr = \wr 0 \wr^*$ it is not true that $\wr 0 \wr^* \cdot \wr \infty \wr = \wr 1 \wr$. Consider ultralarge whole numbers N and $2N$ in $\wr \infty \wr$, then $1/N$ and $1/(2N)$ are both in $\wr 0 \wr^*$. Then 1 and 2 are in the product. The members of the product are not even all limited: consider $1/N^2 \cdot N = 1/N$. In fact $\wr 0 \wr^* \cdot \wr \infty \wr = \mathbf{V}_+ \setminus \{0\}$.
- $\wr +\infty \wr + \wr +\infty \wr = \wr +\infty \wr$.
- $\wr \infty \wr + \wr \infty \wr = \mathbf{V}_+$.
- $\wr +\infty \wr + \wr -\infty \wr = \mathbf{V}_+$.
- $\dfrac{\wr u \wr}{\wr 0 \wr^*} \cdot \wr 0 \wr^* \neq \wr u \wr$.

The fact that $\wr 0 \wr^* \cdot \wr \infty \wr = \mathbf{V}_+ \setminus \{0\}$ conveys the classical intuition that "tending to zero times tending to infinity can be almost anything".

6.4 Formula-defined numeric grains and properties

If an equation given by a **c**-formula, such as $x^2 = 2 \wedge x > 0$, has no solution in the rationals, then the equal sign can be replaced by $\simeq_a$ for some a. If $x^2 \simeq_a 2$ and $y^2 \simeq_a 2$ (and x and y are both positive rationals of $\mathbf{V}_+(a)$), then $x^2 - y^2 = (x+y)(x-y) \simeq_a 0$. A simple approximation shows that $1 < x < 2$ and $1 < y < 2$, hence $x+y$ is neither a-large nor a-small, hence (by proposition 32) $x - y$ must be a-small. Hence $\wr x \wr = \wr y \wr$.

Let $sqrt_a(2)$ denote this grain. It is said to **express** $\sqrt{2}$ at the level of a. It is an a-expression of $\sqrt{2}$ in the sense that it is an a-expression of a numeric grain whose square is equal to $\wr 2 \wr$.

Such formulae may be considered at any finer level $a \sqsubset b$ by

$$u_b = \wr x \in \mathbf{V}_+(b) \mid x^2 \simeq_b 2 \wr.$$

This numeric grain is the b-expression of $\sqrt{2}$.

Numeric grains of level a can be expressed at any finer order. These will be said to express the same **quantity** at different levels.

Considering numeric grains as a-expressions of some property, and considering only those numeric grains for which there is a known defining property, an inclusion into finer levels which preserves total ordering may be defined .

A numeric grain defined by a **c**-formula is stated as being a **c**-defined numeric grain.

The ordering of **c***-defined numeric grains requires that they all be expressed at the same level.*

Definition 16. *If two* **c***-defined numeric grains are equal at all levels, then they are equal.*

If $\wr\, x \,\wr$ is the a-expression of some property of level a, its shadow in coarser levels are simply coarser expressions of the same value (provided the numeric grain is not ultralarge in the coarser levels).

Numeric grains need to be explored in more detail to understand to what extent analysis can be performed using these incomplete objects and what other properties they might have. This is an opening for further research.

Chapter 7

Consistency issues

When studying a new axiomatics, the question arises as to whether there is a model for the new theory. Describing a model of TO in classical model theory (within ZFC) raises some philosophical difficulties. To a certain extent, the question is not even relevant. As Nelson puts it [Nelson, 2006]: "A semantic proof of consistency, by appeal to a model, simply replaces the question of the consistency of a simpler theory, such as P, by the question of the consistency of the far more complicated set theory in which the notion of a model can be expressed." A "finitary" proof of the consistency of a theory would be much more preferable than an "infinitary" proof grounded on the consistency of ZFC. It would be contradictory to the philosophy of the theory of objects to attempt to show anything relative to its consistency using infinite sets and non constructive methods of ultrapowers. The use of nonconstructive methods in building models of nonstandard analysis seems unavoidable as it appears to be the only way to describe external sets such as collections of real numbers bounded above having no least upper bound (which is the case, for instance, for whole numbers of coarsest level).

The fact that the theory of objects is relatively consistent with (F)RIST should therefore not be understood as a proof about the consistency of TO. The embedding of TO in FRIST shows the important fact that the "truths" of TO can be interpreted as truths of (F)RIST and of ZFC.

7.1 Relative Consistency

Working in this chapter within (F)RIST, hence within an extension of ZFC, the concept of finite and infinite will be used and must be understood in the classical sense.

In [Hrbacek, 2004] Hrbacek shows that FRIST is a conservative extension of ZFC. Péraire does the same for RIST in [Péraire, 1992]. As a reminder, (F)RIST is obtained by taking all classical axioms of ZFC and also extra axioms: Idealisation, Standardisation and Transfer – and the axioms defining levels. The theory of objects can be embedded in (F)RIST in the sense that theorems of TO can be interpreted as theorems of FRIST and therefore also interpreted as theorems of ZFC.

The proof that FRIST is a conservative extension of ZFC uses a succession of iterations of ultrapowers. The limit ultrapower construction starts by fixing an ultrafilter over a chosen set and choosing a linearly ordered set $\langle \Lambda, \leq \rangle$. This set will in fact be used to index levels. Proposition 3.6 of [Hrbacek, 2004] stipulates that if $\langle \Lambda, \leq \rangle = \langle \omega, \leq \rangle$, then the model $\mathbb{O}^*\Lambda$ constructed starting with Λ is isomorphic to $^*\omega$, hence discrete. Later in the paper, Proposition 3.8 of [Hrbacek, 2004] specifies that if Λ is densely ordered, then the corresponding model has a dense ordering of levels. [5] Proposition 4.7 of [Hrbacek, 2004] (here, metatheorem 1) states that both choices are possible.

Metatheorem 1. *[Hrbacek, 2004] FRIST is a conservative extension of ZFC for any choice of* Λ.

(Incidentally, if $\Lambda = \{1\}$ this produces a model of IST).

The notation FRIST(D) stands for FRIST with the axiom of discreteness instead of the axiom of density and modelled by $\Lambda = \omega$.

In [Péraire, 1992], Péraire proved that RIST is a conservative extension of ZFC. The crucial point in the construction of his model is that levels are indexed by p and that the range of p is $\mathbb{N}$ (hence discrete). The model provided for RIST is a discrete model, thus proving that a version of RIST with discreteness is a conservative extension. This fact had simply not been made explicit. The notation RIST(D) stands for RIST with the additional axiom of discreteness of levels. Then:

Metatheorem 2. *[Péraire, 1992] RIST(D) is a conservative extension of ZFC.*

7.2 An embedding

Theorem 12 shows that objects are finite in the sense defined in ZFC and (F)RIST. But objects are not only finite, they are hereditarily finite.

Recursively: The empty set is hereditarily finite. A set is hereditarily finite if it is finite and all of its elements are hereditarily finite.

Equivalently, a set is hereditarily finite if its transitive closure is finite. Since the transitive closure of an object is an object, it is finite in the sense defined in ZFC, hence objects are hereditarily finite.

In ZFC, hence in FRIST, there is a set of all hereditarily finite sets denoted by V_ω.

The interpretation of TO in FRIST(D) will thus require that all sets be restricted to hereditarily finite sets, denoted by $h\text{-}fin$ and that quantifiers also restrict to *h-fin* in analogy with its use in IST and RIST: $(\exists^{h\text{-}fin} x)P(x)$ stands for $(\exists x)(x \in V_\omega \wedge P(x)$ and $(\forall^{h\text{-}fin} x)P(x)$ stands for $(\forall x)(x \in V_\omega \Rightarrow P(x))$.

The $\in$ predicate of TO is interpreted as the $\in$ predicate of FRIST(D) and the $\sqsubseteq$ predicate of TO is interpreted as the $\sqsubseteq$ predicate of FRIST(D). An object is interpreted as a hereditarily finite set.

[5] It is an uncomfortable coincidence that discrete and density share the same initial. In the present work, (D) is never used for density but only for discreteness.

Metatheorem 3. *Contextual formulae of TO are interpreted as internal formulae of FRIST(D) and of RIST(D).*

Proof. For FRIST(D): Since V_ω is definable in ZFC, it is a set of the coarsest level in FRIST. The definition of restricted formula of FRIST is given page 80, as given in [Hrbacek et al., 2010b]. Since all objects of TO are hereditarily finite, any formula of the form $(\exists x)P(x)$ (resp. $(\forall x)P(x)$) can be replaced in the interpretation by $[(\exists x)(x \in V_\omega \wedge P(x))]$ (resp. $(\forall x)(x \in V_\omega \Rightarrow P(x))$). Similarly, the reference to V_ω can be added to any list of level parameters since referring also to the coarsest level does not change the validity. Hence any contextual formula is interpreted as a restricted formula of FRIST. Theorem 16, page 81, shows that formulae of this form are indeed internal formulae of FRIST.

Internal formulae of FRIST which do not refer to density are internal formulae of FRIST(D) since any choice of Λ in metatheorem 1 is acceptable and the proof uses neither discreteness nor density.

For RIST(D) the proof is essentially the same.

Standardisation is used in the proof that contextual statements (whose elements are bound by belonging to a referential set) are equivalent to internal proofs with no reference to levels. It is the use of standardisation which requires bounding the variables. The standard set V_ω is used in the same way.

□

Metatheorem 4. *The axioms of TO are interpreted as theorems of RIST(D) and of FRIST(D).*

(F)RIST(D) denotes both FRIST(D) and RIST(D) when no major difference occurs.

Proof.

- The empty object axiom (axiom 1), is interpreted as:

$$(\exists^{h\text{-}fin} x)(\forall^{h\text{-}fin} y)(y \notin x).$$

  In (F)RIST(D) this is satisfied uniquely by $x = \emptyset$. Axiom 1 of TO is interpreted as a theorem of (F)RIST(D).

- Extensionality (axiom 2) is interpreted as:

$$(\forall^{h\text{-}fin} x)(\forall^{h\text{-}fin} y)(\forall^{h\text{-}fin} z)\ (z \in x \Leftrightarrow z \in y) \Leftrightarrow (x = y).$$

  This restriction defines equality between hereditarily finite sets. Since extensionality is an axiom applying to all sets of (F)RIST(D) without restriction, it is true in particular for hereditarily finite sets. Axiom 2 of TO is interpreted as a theorem of (F)RIST(D).

- The successor object (axiom 3) is interpreted as:

$$(\forall^{h\text{-}fin} x)(\forall^{h\text{-}fin} y)(\exists^{h\text{-}fin} z)(\forall^{h\text{-}fin} u)\ [u \in z \Leftrightarrow u \in x \vee u = y\ ].$$

In (F)RIST(D), if x and y are hereditarily finite sets, then $\{y\}$ exists by the powerset axiom and $x \cup \{y\}$ exists by the union axiom and both $\{y\}$ and $x \cup \{y\} = z$ are hereditarily finite. Hence the interpretation of axiom 3 is a theorem in (F)RIST(D).

- Axiom 4.1

$$(\forall^{h\text{-}fin} x)(\forall^{h\text{-}fin} y)(\forall^{h\text{-}fin} z)\ [\ (x \sqsubseteq y) \wedge (y \sqsubseteq z) \Rightarrow (x \sqsubseteq z)\ ]$$

is a restriction of an axiom of (F)RIST(D) which applies to all sets, hence also in the case of the restriction. Their interpretations are theorems of (F)RIST(D).

- Axiom 4.2 is interpreted as:

$$\begin{aligned}(\forall^{h\text{-}fin} x)(\forall^{h\text{-}fin} y)[(x \sqsubset y) \\ \Rightarrow [(\exists^{h\text{-}fin} z)(x \sqsubset z \sqsubseteq y) \wedge \neg(\exists^{h\text{-}fin} u)(x \sqsubset u \sqsubset z) \\ \wedge (\exists^{h\text{-}fin} w)(x \sqsubseteq w \sqsubset y) \wedge \neg(\exists^{h\text{-}fin} v)(w \sqsubset v \sqsubset y)]].\end{aligned}$$

Without the reference to the restriction to V_ω, this would be the axiom of discreteness of (F)RIST(D). Finite ordinals are hereditarily finite sets and for any level in (F)RIST(D), there is an ordinal which appears at that level. Proposition 35 given here also applies in (F)RIST(D). Hence for any level, there is a hereditarily finite object appearing at that level. Hence in particular, for any level at which an *h-fin* object appears, there is a next finer (resp. immediate previous) level at which *h-fin*-objects appear. Thus discreteness of levels applies to hereditarily finite sets of (F)RIST(D).

The interpretation of axiom 4 is a theorem of (F)RIST(D).

- If Φ is a **c**-formula of TO, then $\widehat{\Phi}$ is the corresponding *h-fin*-formula of (F)RIST(D) (a formula in which all unbounded variables of Φ are bound by belonging to V_ω.)

The interpretation of the induction (axiom 5) is given by the following:

$$(\forall^{h\text{-}fin} x)(\forall^{h\text{-}fin} y)[\widehat{\Phi}(\emptyset) \wedge \Big(\widehat{\Phi}(x) \wedge \widehat{\Phi}(y) \Rightarrow \widehat{\Phi}(x \cup \{y\}\Big))] \Rightarrow (\forall^{h\text{-}fin} x)\widehat{\Phi}(x).$$

Let V_ω be ordered by the Ackermann coding in (F)RIST(D). Assume $\widehat{\Phi}(\emptyset)$ and that $\widehat{\Phi}(x) \wedge \widehat{\Phi}(y) \Rightarrow \widehat{\Phi}(x \cup \{y\})$. By contradiction: Assume there is a z such that $\neg\widehat{\Phi}(z)$. Since V_ω is well ordered by the Ackermann coding, let m be the least set such that $\neg\widehat{\Phi}(m)$. $\mathrm{ac}(m) = \sum_{i \in \mathbb{N}} a_i 2^i$ where each a_i is either 0 or 1. This can also be written as $\mathrm{ac}(m) = \sum_{i=0}^{k} a_i 2^i$ for $k = \max\{i \mid a_i \neq 0\}$, then $\mathrm{ac}(m) = \sum_{i=0}^{k-1} a_i 2^i + 2^k$. But $k = \mathrm{ac}(y)$ for some set y and $\sum_{i=0}^{k-1} a_i 2^i = \mathrm{ac}(x)$ for some x. Then $\sum_{i=0}^{k} a_i 2^i = \mathrm{ac}(x \cup \{y\})$. Clearly, $\mathrm{ac}(y) < \mathrm{ac}(m)$ and $\mathrm{ac}(x) < \mathrm{ac}(m)$, hence $\widehat{\Phi}(x)$ and $\widehat{\Phi}(y)$. But this proves $\widehat{\Phi}(x \cup \{y\})$ and since $\mathrm{ac}(x \cup \{y\}) = \mathrm{ac}(m)$, there holds $m = x \cup \{y\}$ which implies $\widehat{\Phi}(m)$. A contradiction.

Induction is interpreted as a theorem in (F)RIST(D).

- The refinement axiom (6) is interpreted as:

  Let $\widehat{\Phi}$ be an *h-fin*-formula, with free variables x and y and possibly other free variables $x_1, \ldots x_n$

  For all v such that $u \sqsubset v$ and $x_1, \ldots x_n \sqsubset v$

$$(\forall^{u,h\text{-}fin} a)(\exists^{h\text{-}fin} y)(\forall x \in a)\widehat{\Phi}^v(x, y) \Leftrightarrow (\exists^{h\text{-}fin} y)(\forall^{u,h\text{-}fin} x)\widehat{\Phi}^v(x, y).$$

  In FRIST(D), idealisation is given by the following: Given $\mathbf{U} \subset \mathbf{V}$, A, $B \in \mathbf{U}$ and $x_1, \ldots, x_k \in \mathbf{V}$:

$$(\forall a \in \mathcal{P}^{fin} A \cap \mathbf{U})(\exists y \in B)(\forall x \in a)\mathcal{P}(x, y, x_1, \ldots, x_k; \mathbf{V}) \Leftrightarrow (\exists y \in B)(\forall x \in A \cap \mathbf{U})\mathcal{P}(x, y, x_1, \ldots, x_k; \mathbf{V}).$$

  Since $x \in a$ and a is required to be hereditarily finite, x is necessarily hereditarily finite.

  Replacing the restriction to finite sets by a restriction to hereditarily finite sets does not change the validity of the formula. The major difference here is that y is bounded by belonging to some set.

  Refinement can be rewritten as

$$(\forall^u a \in V_\omega)(\exists y \in V_\omega)(\forall x \in a)\widehat{\Phi}^v(x, y) \Leftrightarrow (\exists y \in V_\omega)(\forall^u x \in V_\omega)\widehat{\Phi}^v(x, y)$$

  which transposes to:

$$(\forall a \in V_\omega \cap \mathbf{U})(\exists y \in V_\omega \cap \mathbf{V})(\forall x \in a)\widehat{\Phi}(x, y, x_1, \ldots, x_k; \mathbf{V}) \Leftrightarrow (\exists y \in V_\omega \cap \mathbf{V})(\forall x \in V_\omega \cap \mathbf{U})\widehat{\Phi}(x, y, x_1, \ldots, x_k; \mathbf{V})$$

  which is a particular case of FRIST(D) idealisation.

  Hence the interpretation of refinement is a theorem in FRIST(D).

  Furthermore:

  If u is hereditarily finite and $u \subseteq V_\omega$, then $u \in V_\omega$. Conversely, if $u \in V_\omega$ then u is hereditarily finite and $u \subset V_\omega$, hence

$$(\forall^u a \subseteq V_\omega)(\exists y \in V_\omega)(\forall x \in a)\widehat{\Phi}^v(x, y) \Leftrightarrow (\forall^u a \in V_\omega)(\exists y \in V_\omega)(\forall x \in a)\widehat{\Phi}^v(x, y) \Leftrightarrow (\exists y \in V_\omega)(\forall^u x \in V_\omega)\widehat{\Phi}^v(x, y),$$

  hence

$$(\forall^u a \subseteq V_\omega)(\exists y \in V_\omega)(\forall x \in a)\widehat{\Phi}^v(x, y) \Leftrightarrow (\exists y \in V_\omega)(\forall^u x \in V_\omega)\widehat{\Phi}^v(x, y).$$

  Which is the formulation of idealisation in RIST(D) restricted to hereditarily finite sets. Hence refinement is a theorem of RIST(D).

- For the transfer principle: Let $\widehat{\Phi}$ be an *h-fin*-formula such that $\widehat{\Phi} \sqsubseteq \alpha \sqsubseteq \beta$. Then

$$\widehat{\Phi}^{\alpha} \Leftrightarrow \widehat{\Phi}^{\beta}$$

  As it makes no specific reference to the size of collections, its restriction to hereditarily finite sets is immediate.

  It remains to be checked that the passage from Φ to Φ^v in TO is equivalent to the passage from $\widehat{\Phi}$ to $\widehat{\Phi}^v$ in FRIST(D). (The interpretation of Φ^v is the same in TO as in RIST(D).)

  In TO, Φ^v is obtained from Φ by adding v to the list of level variables in every occurrence of level quantifiers.

  In FRIST(D), Φ^v is obtained from Φ by replacing each occurrence of $\sqsubseteq$ by $\sqsubseteq_v$. This predicate is defined by $x \sqsubseteq_v y \equiv (x \sqsubseteq v \wedge y \sqsubseteq v) \vee (x \sqsubseteq y)$.

  If $\exists^{a_1,\dots,a_k} x$ (which stands for $x \sqsubseteq a_1 \vee \cdots \vee x \sqsubseteq a_k$) appears in Φ, then in TO $\exists^{a_1,\dots,a_k,v} x$ will appear in Φ^v. But then this stands for $x \sqsubseteq a_1 \vee \cdots \vee x \sqsubseteq a_k \vee x \sqsubseteq v$ which is $x \sqsubseteq_v a_1 \vee \cdots \vee x \sqsubseteq_v a_k$. Hence the relativisation from Φ to Φ^v in TO is interpreted as the relativisation from $\widehat{\Phi}$ to $\widehat{\Phi}^v$ in FRIST(D). The interpretation of Φ^v is the same in TO as in RIST(D).

  The interpretation of the transfer axiom of TO is a theorem of (F)RIST(D).

  This completes the proof that the axioms of TO are interpreted as theorems of FRIST(D) and RIST(D).

□

Writing ZF-I for ZFC without the axiom of choice and with the negation of infinity as axiom:

Metatheorem 5. *True statements of ZF-I are interpreted as true statements of TO.*

Proof. It is immediate that in ZF-I only hereditarily finite sets can be constructed. Axioms of ZF-I can be interpreted in a straightforward way as theorems of TO. This shows that ZF-I can be embedded in TO and that truths of ZF-I are truths of TO i.e., ZF-I is a restriction of TO. □

7.3 Completeness

Gödel's incompleteness result [Gödel, 1992] holds in the theory of objects as its proof never refers to infinite sets and is in fact predicative in nature. The theory of objects is sophisticated enough to spell out completely the celebrated formula $17\,\mathrm{Gen}\,r$. Indeed, Gödel's proof is about hereditarily finite sets.

Appendix A

Analysis with ultrasmall numbers

This article was co-authored with Karel Hrbacek and Olivier Lessmann and was published in the Monthly Magazine of the Mathematical Association of America in November 2010. It is the first article that presents FRIST in a mathematical magazine which has a non specialised readership. The intention here is both mathematical – to show that it is correct – and pedagogical – to show that it is actually usable at a rather basic level. It is referred to as [Hrbacek et al., 2010a] in the bibliography.

Abstract : *We develop a context-based theory of ultrasmall (infinitesimal) and ultralarge real numbers from a few simple principles, and present some examples of their use in analysis. In this theory, perhaps for the first time, definitions and arguments involving infinitesimals can be presented in a style that is both as informal and as rigorous as is customary in standard textbooks of real analysis.*

Introduction.

Systematic use of infinitesimals in mathematics originated with Leibniz in the 17th century. Infinitesimals served as a mainstream tool of calculus for the next 150 years; unfortunately, mathematicians of the period never succeeded in formulating unambiguous rules for working with them, and the idea was gradually abandoned in favor of the now-standard ε–δ method of Weierstrass.

A rigorous theory of infinitesimals, known as nonstandard analysis, was developed by Robinson [Robinson, 1961] in the 1960s. The usual framework for nonstandard analysis is based on a suitable non-Archimedean extension ${}^*\mathbb{R}$ of the field of real numbers $\mathbb{R}$, often constructed as an ultrapower of $\mathbb{R}$. Such an approach is satisfactory to research mathematicians, and a number of important results proved using nonstandard methods testify to the power of these ideas. The advent of nonstandard analysis also raised hopes that the teaching of calculus at the elementary level could be made easier by replacing the ε–δ arguments with simpler, physically intuitive, yet rigorous reasoning about infinitesimals. Attempts to do so using nonstandard analysis show that the concept of infinitesimal is easy for students to grasp; however, technical details inherent in the development of analysis in the Robinsonian framework (see, for example, [Keisler, 2000, Stroyan, 1997]) present some serious pedagogical difficulties (see [O'Donovan, 2007, Hrbacek, 2007] for a discussion of these matters).

In this paper we propose instead an axiomatic approach motivated by physics. The literature of physics is replete with references to quantities at different *scales*: large scale versus small scale, macroscopic, microscopic, and atomic scales, etc. For example, quantities at the macroscopic scale are those observable with an unaided eye. Sums and products of macroscopic quantities are macroscopic. Compared to macroscopic quantities, such as the diameter of a soccer ball, the quantities observed at the atomic scale (diameter of an atom) or cosmic scale (diameter of a galaxy) are "infinitesimal" and "infinitely large," respectively. We stress that "infinitesimals" in this conception are just ordinary real numbers that are "very small" compared to those observable at the macroscopic scale. They have nothing to do with infinity. To avoid misunderstandings, we abandon the historic terminology and talk about "ultrasmall" and "ultralarge" numbers, rather than "infinitesimal" and "infinitely large" ones.

Another important point suggested by the above discussion is that "ultrasmall" is a *relative* concept. The diameter of a bacterium is ultrasmall relative to macroscopic quantities, but at the microscopic scale it is significant. The diameter of an atom, in its turn, can be regarded as ultrasmall relative to the microscopic scale, etc.

Scales of magnitude play an important role in the thinking of physicists, but to a mathematician the concept seems incoherent. This becomes evident if we formulate some of the requisite properties of macroscopic quantities as axioms:

(a) 1 is macroscopic.

(b) If n is macroscopic, then $n + 1$ is macroscopic.

(c) Not all natural numbers are macroscopic.

The principle of mathematical induction fails for the property of being macroscopic! This is an example of the ancient *sorites paradox* attributed to Eubulides of Miletus (from Greek *soros* for heap). Usually it is formulated as follows: One grain of sand is not a heap of sand. If a number of grains of sand does not make a heap, then adding one more grain still does not make a heap. Yet heaps of sand do exist.

There are many examples of "soritic properties" for which mathematical induction does not hold ("number of grains in a heap," "number that can be written down with pencil and paper in decimal notation," "macroscopic number,"...), but mathematicians traditionally take no account of them in their theories, with the excuse that such properties are vague. We present here a mathematically rigorous theory in which a soritic property is put to constructive use. This theory — we call it *relative analysis* (RA) — axiomatizes the traditional mathematical concepts such as numbers, sets, and membership, but, in addition, it includes a new primitive binary predicate $\sqsubseteq$. The intuitive meaning of the statement "$x \sqsubseteq y$" is "*x is observable at every scale where y is observable*"; rigorously, it is given by the axioms of RA. In this richer language we can express distinctions ignored by traditional mathematics, such as the idea that one real number may be ultrasmall or ultralarge relative to another. The notions of ultrasmall and ultralarge can then be used to develop calculus in the style of Leibniz.

Formally, relative analysis is an *extension* of **ZFC** (Zermelo-Fraenkel set theory with the axiom of choice, the theory that is thought to codify current mathematical practice). Just as in the case of **ZFC**, all objects of RA are sets. Natural numbers and real numbers are defined in the usual way. RA does not change any of the properties of numbers and sets that are expressible in the traditional mathematical language of **ZFC**. Concepts such as "finite," "infinite," "well ordered," and so on have the usual meaning. All traditional theorems and proofs remain valid. For example, every nonempty set of natural numbers still has a least element, the field $\mathbb{R}$ is Archimedean (for every $x \in \mathbb{R}$ there is $n \in \mathbb{N}$ such that $x < n$), and it is the unique complete (every nonempty bounded set has a supremum) ordered field, up to an isomorphism. Because of this, we can work in RA as if it described the mathematical universe with which we are familiar.

Relative analysis extends the description of the mathematical universe given by **ZFC** by taking into account also the intuitive notion, often used by physicists, that some quantities are negligible compared to others. The axioms about $\sqsubseteq$ portray an idealized version of this notion, just as the axioms of Euclidean geometry portray an idealization of our experience with, or intuition about, straight line segments. We stress that it is an idealization that is suitable for pure mathematics, but abstracts from many specific details of its intuitive motivation. For example, while 10^{100} is surely not macroscopic in the sense discussed above, in RA it is observable at every scale.

The presence of $\sqsubseteq$ in the language of relative analysis enables us to describe properties of natural numbers for which mathematical induction fails, for example $n \sqsubseteq x$ (n is observable whenever x is observable, for a fixed x). This is a vestige of the intuitive origin of $\sqsubseteq$ as a soritic property. It simply means that in RA there is no set S such that, for all $n \in \mathbb{N}$, $n \in S$ if and only if $n \sqsubseteq x$. The existence of such properties may be difficult to reconcile with the intuitive interpretation of natural numbers, real numbers, and sets of RA as the usual, familiar ones. Yet we strongly urge the reader to engage in such an interpretation, if only at the level of fiction, because it is very helpful when practicing mathematics in RA. A brief discussion of alternatives can be found at the end of Section starting page 65.

Soritic collections like $\{n \in \mathbb{N} : n \sqsubseteq x\}$ are not entities of our theory: relative analysis does not postulate their existence. Nevertheless, it turns out that the axioms of RA are simplified and made more intuitive when they are stated in terms of certain collections we call *levels*, rather than directly in terms of $\sqsubseteq$, as is the case in the extant literature. In this paper, we write $x \in \mathbf{V}(y)$ in place of $x \sqsubseteq y$, and read it as "*x appears at the level of y.*" Levels are not sets, so they are entities outside the traditional mathematical universe, but the use of levels is just a convenient way of speaking that can easily be eliminated in favor of $\sqsubseteq$. There is a precedent: mathematicians realized long ago that certain useful collections cannot be sets — the collection of all sets, the collection of all groups, various functors used in algebraic topology, etc. Yet mathematicians freely work with such collections (*proper classes*) because they simplify statements and proofs of important results. As in the case of levels, the use of proper classes can in principle be eliminated.

These are not entirely new ideas. The earliest mathematical theory that takes soritic concepts seriously seems to be the *theory of semisets* of Vopěnka, intended as an axiomatization of the method of forcing [Vopěnka and Hájek, 1972]. In the mid-1970s, several axiomatic nonstandard set theories were proposed with the objective of making nonstandard analysis more accessible (see the first author's papers [Hrbacek, 1978, Hrbacek, 1979], Nelson [Nelson, 1977], and Vopěnka [Vopěnka, 1979]); also the monograph of Kanovei and Reeken [Kanovei and Reeken, 2004]. In particular, our intuitive interpretation of relative analysis is similar to that proposed for Nelson's **IST** [Nelson, 1977, Diener and Diener, 1995]; we note especially that "x is standard" of **IST** is a soritic property in our sense. Relative analysis goes beyond **IST** in that it extends the mathematical language by a *binary* $x \sqsubseteq y$ (which can also be read as "*x is standard relative to y*"), rather than by a *unary* predicate "*x is standard*," as do **IST** and other nonstandard set theories. This move is suggested by the physicists' idea of scales as described above, and it is essential if the theory is to work at the elementary level. In **IST**, infinitesimals can be used to define calculus concepts, such as $f'(a)$, but only for *standard* functions f at *standard* points a. It is of course imperative to have a definition of $f'(a)$ that works for *all* real numbers and functions, whether standard or not. But if a is nonstandard, no "infinitesimals relative to a" are available in **IST**, and one has to fall back on the ε–δ method (see [Hrbacek, 2007, O'Donovan, 2007] for a detailed discussion).

Relative analysis is a *conservative* extension of **ZFC**. This means that every traditional mathematical statement provable in RA is provable in **ZFC**; the point is of course that the proof in RA may be simpler, shorter, and more intuitive. In particular, RA is safe: any contradiction discovered in it would give rise to a contradiction in **ZFC**.

A set theory with a binary relative standardness predicate was first developed by Péraire in [Péraire, 1992] under the acronym **RIST** (*Relative Internal Set Theory*); a different approach to relative analysis is presented in Gordon's [Gordon, 1997]. The first author further extended **RIST**, and showed that the resulting theory **FRIST** is a conservative extension of **ZFC**, in [Hrbacek, 2004, Hrbacek, 2009]. This paper is based on a fragment of **FRIST** [Hrbacek, 2010].

The plan of the rest of the paper is as follows. In Section starting page 65 we state the axioms of relative analysis and give definitions of its key concepts. This approach can be used to introduce nonstandard ideas into elementary calculus. It has been successfully tested in several high school calculus classes in Geneva in the spring of 2009. The presentation for beginning students would focus on axioms I – V; students seem to accept VI and VII implicitly, and are not immediately concerned about sophisticated issues like induction. More details on how relative analysis can be employed in the high school environment can be found in the third author's [O'Donovan, 2009]. Section starting page 65 concludes with some further discussion of foundational issues raised by relative analysis.

Section starting page 70 is a sampler of definitions and theorems from calculus of one real variable. Our main goal is to demonstrate that, once the principles of relative analysis have been assimilated, it becomes possible to do mathematics with ultrasmall numbers in a style that is just as informal and natural as the treatments found in traditional textbooks, but with important advantages. Use of ultrasmall numbers disposes with the ε–δ machinery that students find notoriously difficult, and with the associated bookkeeping. The proofs become simpler and more focused on the "combinatorial" heart of arguments. Fundamental results, such as the extreme value theorem, can be fully proved from the axioms immediately, without the need to master difficult notions of supremum or compactness. As a result, calculus can be presented as *mathematics* — with proofs — even at a student level where vague arguments about "approaching" have become the norm. Derivatives and definite integrals can be developed before limits, and independently of each other. The relative framework allows arguments involving two or more levels; this simplifies proofs about double limits, even compared to the Robinsonian framework. A rigorous theory of ultrasmall and ultralarge numbers also enables the construction of entirely new models of mathematical and physical phenomena, although we do not pursue this direction here. A full, systematic development of relative analysis is in preparation (a first version is [Hrbacek et al., 2010b]).

Principles of relative analysis.

We stress again that relative analysis *does not change* the definitions of any mathematical concepts or invalidate any mathematical theorems; this is captured formally by postulating the axioms of **ZFC**. Relative analysis merely *adds* something to mathematics — namely, the concept of level — and is thus able to make distinctions that cannot be expressed in the language of traditional mathematics. The levels of relative analysis should be thought of as an *idealized* version of the "scales" discussed in the introduction. Given a list of real numbers $a_1, \ldots, a_k$, we visualize the level determined by $a_1, \ldots, a_k$ as containing all real numbers obtained from $a_1, \ldots, a_k$ by the operations of addition, multiplication — indeed, to obtain a smoothly functioning theory, we generalize and require at once that levels be closed under *every* operation defined in traditional mathematics. It is convenient to assign levels also to sets, functions, operations, and in fact to all mathematical objects, by the simple idea that if an object is defined from parameters $a_1, \ldots, a_k$, then it appears at the level of $a_1, \ldots, a_k$. For example, the quadratic function f, defined for all x by $f(x) = ax^2 + bx + c$, appears at the level where the parameters a, b, c appear.

We now proceed to rigorously state the basic principles of relative analysis.[6] We add to the mathematical language a new sort of variables, denoted by $\mathbf{V}$ with various decorations, and assumed to range over *levels*. (Other variables, such as $x, y, A, \ldots$, range over sets, unless the context indicates otherwise.) The notion of level is an additional primitive concept, whose meaning is given implicitly by the axioms. Levels are not sets, and are allowed only on the right side of $\in$; to stress this, we read "$x \in \mathbf{V}$" as "*x appears at the level* $\mathbf{V}$," and we read $\mathbf{V}_1 \subseteq \mathbf{V}_2$ as "$\mathbf{V}_1$ is *coarser* than $\mathbf{V}_2$" or "$\mathbf{V}_2$ is *finer* than $\mathbf{V}_1$."

Axiom I. *For every $x_1, \ldots, x_k$ there is a level $\mathbf{V}(x_1, \ldots, x_k)$ such that $x_1, \ldots, x_k \in \mathbf{V}(x_1, \ldots, x_k)$ and,*

for all levels $\mathbf{V}$, $x_1, \ldots, x_k \in \mathbf{V}$ implies $\mathbf{V}(x_1, \ldots, x_k) \subseteq \mathbf{V}$.

This axiom postulates that for all $x_1, \ldots, x_k$ there is a coarsest level where $x_1, \ldots, x_k$ appear; we denote it $\mathbf{V}(x_1, \ldots, x_k)$ and call it *the level of* $x_1, \ldots, x_k$.

Axiom II. *For every $\mathbf{V}_1$ and $\mathbf{V}_2$ either $\mathbf{V}_1 \subseteq \mathbf{V}_2$ or $\mathbf{V}_2 \subseteq \mathbf{V}_1$.*

[6] For a description of how these axioms can be presented at the high school level, see [15].

Levels enable us to define ultrasmall and ultralarge numbers.

Definition A.1. *Given a level* $\mathbf{V}$*:*

(1) A real number ε *is* ultrasmall *relative to* $\mathbf{V}$ *if* $|\varepsilon| < r$ *for all* $r > 0$*,* $r \in \mathbf{V}$*.*

(2) A real number x *is* ultralarge *relative to* $\mathbf{V}$ *if* $|x| > r$ *for all* $r > 0$*,* $r \in \mathbf{V}$*; it is* limited *relative to* $\mathbf{V}$ *if it is not ultralarge relative to* $\mathbf{V}$*.*

(3) Real numbers a *and* b *are* ultraclose *relative to* $\mathbf{V}$*, written* $a \simeq_{\mathbf{V}} b$*, if* $a - b$ *is ultrasmall relative to* $\mathbf{V}$*.*

Note that 0 is the only number ultrasmall relative to $\mathbf{V}$ that appears at $\mathbf{V}$, and no numbers ultralarge relative to $\mathbf{V}$ appear at $\mathbf{V}$. A number x is limited relative to $\mathbf{V}$ if and only if there are $r, s \in \mathbf{V}$ such that $r < x < s$.

Axiom III. *For every level* $\mathbf{V}$ *there exist nonzero real numbers ultrasmall relative to* $\mathbf{V}$*.*

Axiom IV (Neighbor Principle). *For every real number* x *limited relative to* $\mathbf{V}$ *there is a real number* $r \in \mathbf{V}$ *such that* $x \simeq_{\mathbf{V}} r$*.*

Intuitively, the Neighbor Principle asserts that every real number has a "best approximation" at every given "scale." It is a version of completeness of $\mathbb{R}$.

Axiom V (Closure Principle).

A number, function, operation, or set that is uniquely defined (in traditional mathematics, that is, without any mention of levels) from parameters that appear at level $\mathbf{V}$ *does itself appear at level* $\mathbf{V}$*.*

Thus the numbers $1, 2, 196883, 4/5, \sqrt{2}, \pi$, the functions and operations $+$, $\times$, sin, ln, and the sets $\varnothing, [0,1], \mathbb{N}, \mathbb{R}$, which can be defined outright (without any parameters), appear at every level. If $a, b \in \mathbb{R}$ appear at $\mathbf{V}$, then $a + 3$, $\sin a$, and $[a, b]$ appear at $\mathbf{V}$. We stress that the set $\mathbb{R}$ of real numbers appears at every level $\mathbf{V}$, but, by Axiom **III**, it has elements that do not appear at $\mathbf{V}$. In relative analysis, every infinite set has elements that do not appear at the level of that set. This is in agreement with a dictum of Hilbert, "*We know sets before we know* [all of] *their elements.*"

The reader may perhaps wish for a specific example of an ultrasmall or ultralarge number. We cannot give one, but this is precisely the point: numbers ultralarge relative to some level are so large that they cannot be uniquely specified by the means available at that level.

We now prove some basic properties of these new concepts. Throughout the paper we adopt the convention that the mention of $\mathbf{V}$ can be omitted if $\mathbf{V}$ is given, or understood from the context.

Proposition A.1. *Relative to a given level* $\mathbf{V}$*:*

(1) If x and y are limited real numbers, then $x \pm y$ and $x \cdot y$ are limited.

(2) If δ and ε are ultrasmall, then $\delta \pm \varepsilon$ is ultrasmall.

(3) If ε is ultrasmall and x is limited, then $\varepsilon \cdot x$ is ultrasmall.

(4) If $x \neq 0$, then x is ultralarge if and only if $1/x$ is ultrasmall.

(5) $a \simeq a$; if $a \simeq b$, then $b \simeq a$; and if $a \simeq b$ and $b \simeq c$, then $a \simeq c$.

(6) If $x \simeq a$ and $y \simeq b$, then $x \pm y \simeq a \pm b$.

(7) If $x \simeq a$, $y \simeq b$, and a and b are limited, then $x \cdot y \simeq a \cdot b$.

(8) For $a, b \in \mathbf{V}$, $a \simeq b$ if and only if $a = b$.

Proof. As an example, we prove (1), (3), and (7).

(1) If x and y are limited relative to $\mathbf{V}$, then $|x| \leq r$ and $|y| \leq s$, for some $r, s \in \mathbf{V}$. It follows that $|x \pm y| \leq r + s$ and $|x \cdot y| \leq r \cdot s$, where $r + s$, $r \cdot s \in \mathbf{V}$ by the Closure Principle. Hence $x \pm y$ and $x \cdot y$ are limited relative to $\mathbf{V}$.

(3) Let $|x| \leq r_0$, where $r_0 > 0$, $r_0 \in \mathbf{V}$. For every $r > 0$, $r \in \mathbf{V}$, we have $r/r_0 > 0$ and $r/r_0 \in \mathbf{V}$, by the Closure Principle. Hence $|\varepsilon| < r/r_0$, and $|\varepsilon \cdot x| < (r/r_0) \cdot r_0 = r$. This shows that $\varepsilon \cdot x$ is ultrasmall.

(7) We have $x = a + \delta$ and $y = b + \varepsilon$, for some ultrasmall δ and ε. Hence $x \cdot y = a \cdot b + \varepsilon \cdot a + \delta \cdot b + \delta \cdot \varepsilon \simeq a \cdot b$, as $\varepsilon \cdot a$, $\delta \cdot b$, and $\delta \cdot \varepsilon$ are ultrasmall by (3), and a sum of ultrasmall numbers is ultrasmall by (2). □

By Proposition A.1, the number r in the Neighbor Principle is uniquely determined; we call it the $\mathbf{V}$-*neighbor* of x and denote it $\mathbf{n}_{\mathbf{V}}(x)$.

Ultrasmall numbers can be used to define and calculate derivatives.

Convention. *In definitions where no level is specified, the context level is always the level of the parameters of the property or operation being defined.*

Definition A.2. *Let f be a function whose domain is an open interval about a.*

*The **derivative** of f at a is a real number L at the context level such that*

$$\frac{f(a+dx) - f(a)}{dx} \simeq L \qquad \text{for all ultrasmall } dx \neq 0.$$

By the convention, the context level is $\mathbf{V}(f, a)$. By Closure, the domain of f appears at the context level and hence f is defined for all $a + dx$ where dx is ultrasmall.

The derivative of f at a is "the best approximation, at the observation level" to the average rate of change of f in an ultrasmall interval about a. If the number L exists, it is uniquely determined, by Proposition A.1 (8); we denote it $f'(a)$.

Example Let $f(x) = x^2$ and $a \in \mathbb{R}$. Compute $f'(a)$.

We work relative to $\mathbf{V}(f,a)$ (actually $\mathbf{V}(f,a) = \mathbf{V}(a)$, because the function f appears at every level). For $dx \neq 0$ ultrasmall,

$$\frac{f(a+dx)-f(a)}{dx} = \frac{(a+dx)^2 - a^2}{dx} = 2a + dx \simeq 2a.$$

We observe that $2a \in \mathbf{V}(f,a)$; thus $f'(a) = 2a$.

We note an important fact: the calculation of $f'(a)$ in this example gives the same result if instead of $\mathbf{V}(f,a)$ we work relative to any level $\mathbf{V}$, as long as f and a appear at $\mathbf{V}$ (as long as $\mathbf{V}$ is an "observation level" for the objects we are studying). Relative analysis makes this remark into a fundamental general principle.

Axiom VI (Stability Principle).
Let $\mathcal{P}(x_1, \ldots, x_k; \mathbf{V})$ be a statement about $\mathbf{V}$.
If $x_1, \ldots, x_k \in \mathbf{V}_1$ and $x_1, \ldots, x_k \in \mathbf{V}_2$, then

$$\mathcal{P}(x_1, \ldots, x_k; \mathbf{V}_1) \iff \mathcal{P}(x_1, \ldots, x_k; \mathbf{V}_2).$$

A *statement about* $\mathbf{V}$ is any statement (well-formed formula of the extended language) with free variables among $x_1, \ldots, x_k$ (also referred to as *parameters*) and $\mathbf{V}$. Quantifiers over levels are allowed in the form $(\forall\, \mathbf{V}' \supseteq \mathbf{V})$ and $(\exists\, \mathbf{V}' \supseteq \mathbf{V})$, although in applications of Stability in analysis, quantifiers over levels usually do not occur explicitly. We call the level $\mathbf{V}(x_1, \ldots, x_k)$ of the parameters of a statement $\mathcal{P}(x_1, \ldots, x_k; \mathbf{V})$ *the context level* for that statement.

The Stability Principle asserts that if a statement is true about its context level, then it remains true about every finer level. Hence in the definition of the derivative, any level where f and a appear can be used in place of $\mathbf{V}(f,a)$. This is a general fact; the level does not matter, as long as it is sufficiently fine (at least as fine as the context level).

An immediate consequence of Stability is the following generalization of the Closure Principle, known in nonstandard analysis as the *Transfer Principle.*

Closure Principle. *Let $\mathcal{P}(x, x_1, \ldots, x_k)$ be a statement that does not mention levels (formally, a statement in the language of* **ZFC***).*

If $x_1, \ldots, x_k \in \mathbf{V}$ and $(\exists x)\mathcal{P}(x, x_1, \ldots, x_k)$, then $(\exists x \in \mathbf{V})\mathcal{P}(x, x_1, \ldots, x_k)$.

In the contrapositive form:

If $x_1, \ldots, x_k \in \mathbf{V}$ and $(\forall x \in \mathbf{V})\mathcal{P}(x, x_1, \ldots, x_k)$, then $(\forall x)\mathcal{P}(x, x_1, \ldots, x_k)$.

Proof. Let $x_1, \ldots, x_k \in \mathbf{V}$ and an arbitrary x be such that $\mathcal{P}(x, x_1, \ldots, x_k)$ holds. By Axiom **I**, there is some level $\mathbf{V}'$ such that $x, x_1, \ldots, x_k \in \mathbf{V}'$. The statement $(\exists x \in \mathbf{V}')\mathcal{P}(x, x_1, \ldots, x_k)$ about $\mathbf{V}'$ is true; hence, by Stability, $(\exists x \in \mathbf{V})\mathcal{P}(x, x_1, \ldots, x_k)$ is also true. □

An important consequence of the Closure Principle is: if $F(x_1, \ldots, x_k)$ is any operation defined in conventional mathematics (i.e., **ZFC**), then for any level V, $x_1, \ldots, x_k \in \mathsf{V}$ implies that $F(x_1, \ldots, x_k) \in \mathsf{V}$. Taking F to be the addition operation on natural numbers, the Closure Principle implies in particular that, for any level V:

$$1 \in \mathsf{V} \text{ and } (n \in \mathsf{V} \Rightarrow n + 1 \in \mathsf{V}) \text{ for all } n \in \mathbb{N}.$$

Hence not all statements in our extended language define sets or are subject to the principle of mathematical induction; "$n \in \mathbb{N} \wedge n \in \mathsf{V}$" is a simple counterexample. Fortunately, statements that behave in the conventional way are very easy to recognize by inspection.

A statement $\mathcal{Q}(x_1, \ldots, x_k)$ is *internal* if either it does not mention levels, or all levels mentioned in it are at least as fine as its context level $\mathsf{V}(x_1, \ldots, x_k)$. Technically, $\mathcal{Q}(x_1, \ldots, x_k)$ is internal if all quantifiers over levels that occur in it are of the form ($\exists \mathsf{V}$ such that $x_1, \ldots, x_k \in \mathsf{V}$) or ($\forall \mathsf{V}$ such that $x_1, \ldots, x_k \in \mathsf{V}$) (and there are no free variables ranging over levels).

Axiom VII (Definition Principle).
If $\mathcal{Q}(x, x_1, \ldots, x_k)$ is internal and A is a set, then there is a set B such that

$$(\forall x)(x \in B \iff x \in A \wedge \mathcal{Q}(x, x_1, \ldots, x_k)).$$

By Closure Principle, easily generalized to all internal statements $\mathcal{P}$, if $A, x_1, \ldots, x_k \in \mathsf{V}$, then also $B \in \mathsf{V}$.

Corollary. *The principle of mathematical induction holds for internal statements.* Indeed, if $\mathcal{Q}(x)$ is internal, then $\{n \in \mathbb{N} : \mathcal{Q}(n)\}$ is a set.

We conclude this section with a few comments on some questions that mathematicians often raise at the first encounter with relative analysis.

(1) *Do levels "really" exist?*

In the view of some mathematicians, *infinite sets* do not "really" exist; they are mere figments useful for proving theorems about entities that do exist, such as natural numbers. Levels serve exactly the same purpose. However, many mathematicians do seem to ascribe some sort of existence to sets. Even proper classes, such as the class of all sets, initially intended as just a convenient way of speaking about extensions of statements (formulas), tend to be endowed by practicing set theorists with the same degree of reality as sets. We can testify from our own experience that the reality of levels grows on one with exposure to their usefulness.

(2) *Should not the sets at the coarsest level be identified with the "real" sets?*

The level $\mathsf{V}(\cdot)$ (empty list) is the coarsest of all levels. By Closure, every mathematical object that is uniquely described in traditional mathematics appears at this level. An interpretation of nonstandard set theory in which the sets at the coarsest level are regarded as the usual, "standard," sets is philosophically coherent (see [Hrbacek, 1979]), and trained mathematicians often prefer it. But every infinite set at the coarsest level has elements at all finer levels. This interpretation endows $\mathbb{N}$ with a multitude of ideal, "hypernatural" elements; hence $\mathbb{N}$ is not the *usual* set of natural numbers! It suggests that the collection $^{\circ}\mathbb{N}$ containing all natural numbers at the coarsest level *and nothing else* should be viewed as the usual set of natural numbers. But $^{\circ}\mathbb{N}$ is not an entity of relative analysis! While it is possible to introduce such objects into the theory, doing so would saddle it with pedagogical problems similar to those that arise in the Robinsonian approach.

(3) *Are not* all *natural numbers uniquely definable, hence at the coarsest level?*

It is of course true that every natural number n can be uniquely *represented* in, say, decimal notation as a finite sequence $i_0 \dots i_m$ of digits. There is no contradiction here: if $n \notin \mathbf{V}$, then $m \notin \mathbf{V}$. Unique definability in the Closure Principle means something else: the existence of a statement $\mathcal{P}(n)$ such that **ZFC** proves that there is a unique n for which $\mathcal{P}(n)$ holds. It is *not* a theorem of **ZFC** that "every natural number is uniquely definable" in this sense; indeed, the statement in quotes confuses mathematics and metamathematics, and cannot even be expressed in **ZFC**. Whether the statement has some truth value, and what it might be, seems to be a function of one's foundational views. We sympathize with a famous dictum of Reeb, "*Les entiers naïfs ne remplissent pas* $\mathbb{N}$."[7]

Applications of relative analysis.

In this part of the paper we present some selected applications of the relative framework to analysis.

Definition A.3. *Let f be a function defined on an interval I and $a \in I$. We say that f is* continuous *at a if $f(x) \simeq f(a)$ for all $x \simeq a$, $x \in I$.*

The definition is intuitive: continuity means that an ultrasmall change of the argument results in an ultrasmall change of the value of the function.

We recall our convention from Section starting page 65: in a definition where no level is specified, the context level is always the level of the parameters of the property or operation being defined. Relative concepts, such as "ultrasmall," "ultralarge," $\simeq$, and $\mathbf{n}$, are to be taken relative to this context level.

The context level for the above definition is thus $\mathbf{V} = \mathbf{V}(f, a, I)$ and $\simeq$ means $\simeq_{\mathbf{V}}$. We note that actually $\mathbf{V} = \mathbf{V}(f, a)$, because the domain I of the function f is uniquely determined by f, and therefore I appears at the level of f. Having said that, we immediately point out that such care in specifying the context level is not necessary. *Any* level where all the parameters of the property or operation being defined appear can be used equally well, by Stability. Therefore, the above definition can equivalently be restated as: f is *continuous* at a if $f(x) \simeq_{\mathbf{V}} f(a)$ for all $x \simeq_{\mathbf{V}} a$, $x \in I$, where $f, a \in \mathbf{V}$.

[7] The naive integers do not fill up $\mathbb{N}$.

The statement that defines continuity of f at a is obviously internal.[8] Concepts defined by statements that are equivalent to internal statements are called *internal concepts*. Definitions that use the above convention automatically define internal concepts.

At the beginning of a proof of any theorem, we fix a context level for the proof, by some formulation such as "We work relative to the level $\mathbf{V}(a, b, \ldots, z)$." Usually, $a, b, \ldots, z$ are the entities the theorem is about. By Stability, any finer level could be used equally well, and we often say more loosely "Let $\mathbf{V}$ be a level where $a, b, \ldots, z$ appear," and refer to this $\mathbf{V}$ subsequently as "the context level." All relative concepts mentioned in the course of the proof are to be taken relative to this context level, unless explicitly stated otherwise. Of course, objects introduced by means of such relative concepts do not have to *appear* at the context level.

Theorem A.1. *If f and g are functions defined on I and continuous at a, then $f \cdot g$ is continuous at a.*

Proof. We work relative to $\mathbf{V} = \mathbf{V}(f, g, a)$. By Closure, also $f \cdot g \in \mathbf{V}$. If $x \simeq a$, then $f(x) \simeq f(a)$ because $f, a \in \mathbf{V}$ and f is continuous at a, and $g(x) \simeq g(a)$ because $g, a \in \mathbf{V}$ and g is continuous at a. Again by Closure, $f(a), g(a) \in \mathbf{V}$; hence they are not ultralarge, and by Proposition A.1 (7) we have $f(x) \cdot g(x) \simeq f(a) \cdot g(a)$. This proves that $f \cdot g$ is continuous at a. □

Theorem A.2. (Extreme Value) *Let f be a function defined on $[a, b]$ and continuous at every $x \in [a, b]$. Then f attains its maximum and minimum on $[a, b]$.*

Proof. We consider the case of the maximum. Work relative to a level where f, a, b appear. Let N be an ultralarge positive integer, $\Delta = (b - a)/N$, and $x_i = a + i \cdot \Delta$, for $i = 0, \ldots, N$. The set $\{f(x_0), \ldots, f(x_N)\}$ is finite (albeit ultralarge), so it has a greatest element, say $f(x_j)$. Let $c = \mathbf{n}(x_j)$ (it exists because x_j is limited); clearly $c \in [a, b]$. By continuity of f at c, we have $f(x_j) \simeq f(c)$, and by Closure, $f(c)$ appears at the context level.

Let $x \in [a, b]$ appear at the context level. There is i such that $x_i \leq x < x_{i+1}$. Hence $x_i \simeq x$ and $f(x_i) \simeq f(x)$, because f is continuous at x and x appears at the context level. By definition of x_i and c we have

$$f(x) \simeq f(x_i) \leq f(x_j) \simeq f(c).$$

As $f(x)$ and $f(c)$ appear at the context level, this implies that $f(x) \leq f(c)$. We have proved that $f(c)$ is the maximum value of $f(x)$ for all x at the context level. By Stability, the same is true about every finer level; hence $f(c)$ is the maximum value of f. □

Definition A.4. *Let f be a function defined on an interval I. We say that f is* uniformly continuous *if $f(x) \simeq f(y)$ for all $x \simeq y$, $x, y \in I$.*

[8] Technically, it can be written as: f is continuous at a if and only if for *some* $\mathbf{V}$ such that $f, a \in \mathbf{V}$, $f(x) \simeq_{\mathbf{V}} f(a)$ for all $x \simeq_{\mathbf{V}} a$, $x \in I$, or, equivalently, if and only if for *every* $\mathbf{V}$ such that $f, a \in \mathbf{V}$, $f(x) \simeq_{\mathbf{V}} f(a)$ for all $x \simeq_{\mathbf{V}} a$, $x \in I$.

Uniform continuity is an internal concept; the symbol $\simeq$ is relative to $\mathbf{V}(f)$. Explicitly, f is uniformly continuous if $f(x) \simeq_{\mathbf{V}(f)} f(y)$ whenever $x \simeq_{\mathbf{V}(f)} y$. By Stability, $\mathbf{V}(f)$ can be replaced by any $\mathbf{V}$ with $f \in \mathbf{V}$. Compare this with the following.

Definition A.5. *Let f be a function defined on an interval I. We say that f is* continuous *if f is continuous at x for all $x \in I$.*

This definition does not explicitly mention levels, but levels are of course used in the definition of continuity of f at x. When it is spelled out in full, it asserts that f is continuous if $f(x) \simeq_{\mathbf{V}(f,x)} f(y)$ whenever $x \simeq_{\mathbf{V}(f,x)} y$. The level $\mathbf{V}(f,x)$ here varies with x, but it is always at least as fine as the context level $\mathbf{V}(f)$; hence continuity on an interval is an internal concept. It is easy to see in general that statements that do not mention levels explicitly, but refer to them indirectly via previously defined internal concepts, are equivalent to internal statements. The Stability Principle remains valid for statements about $\mathbf{V}$ that refer to previously defined internal concepts.

It follows immediately from Stability that uniform continuity implies continuity. Indeed, given $x \in I$, if $f(x) \simeq f(y)$ for all $y \in I$ such that $x \simeq y$ is true relative to $\mathbf{V}(f)$, then the same is true relative to $\mathbf{V}(f,x)$. The converse holds on closed intervals.

Theorem A.3. *If f is defined and continuous on $[a,b]$, then f is uniformly continuous.*

Proof. We work relative to a level where f (hence also a and b, the uniquely determined left and right endpoints of the domain of f) appears. Let $x, y \in [a,b]$ with $x \simeq y$. Let $c = \mathsf{n}(x)$ (x is limited); then $c \in [a,b]$. But $x \simeq c$ and also $y \simeq c$ (since $x \simeq y$). Since f is continuous at c and c appears at the context level, we have $f(x) \simeq f(c)$ and $f(y) \simeq f(c)$. This implies that $f(x) \simeq f(y)$. □

Definition A.6. *Let f be a function defined on a neighborhood of a, except possibly at a. We say that a real number L is a* limit *of f at a if L appears at the context level and $f(x) \simeq L$ for all $x \simeq a$, $x \neq a$. If the number L exists, it is uniquely determined; we denote it* $\lim_{x\to a} f(x)$.

The context level here is of course $\mathbf{V}(f,a)$.[9] We prove that any finer level can be used instead.

Proof. Assume that $f, a, L \in \mathbf{V} = \mathbf{V}(f,a)$ and $f, a, L' \in \mathbf{V}'$ are such that $f(x) \simeq_{\mathbf{V}} L$ for all $x \simeq_{\mathbf{V}} a$, $x \neq a$, and $f(x) \simeq_{\mathbf{V}'} L'$ for all $x \simeq_{\mathbf{V}'} a$, $x \neq a$. By Axiom **I**, there is a level $\mathbf{V}''$ such that $f, a, L, L' \in \mathbf{V}''$. By Stability, $x \simeq_{\mathbf{V}''} a$, $x \neq a$ implies $f(x) \simeq_{\mathbf{V}''} L$ and $f(x) \simeq_{\mathbf{V}''} L'$. Hence $L \simeq_{\mathbf{V}''} L'$ and, as $L, L' \in \mathbf{V}''$, we get $L = L'$. □

It is convenient to write $x \simeq +\infty$ if x is positive and ultralarge, and $x \simeq -\infty$ if x is negative and ultralarge (relative to a given context level).

[9] By Closure, f is defined on a neighborhood at the context level, and in particular, for all $x \simeq a$, except perhaps $x = a$.

Definition A.7. *Let f be a function defined on a deleted neighborhood of a. We say that $\lim_{x\to a} f(x) = +\infty$ if $f(x) \simeq +\infty$ for all $x \simeq a$, $x \neq a$. Similarly for $-\infty$ and $\lim_{x\to\pm\infty} f(x) = L$.*

The familiar properties of limits are easily proved by arguments analogous to those used for continuity. We also have (see Definition A.2)

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

The definition of derivative can be rewritten in a number of ways; one of the most useful is the following.

Theorem A.4. (Increment Equation) *Let f be a function defined on an open interval containing a. Then f is differentiable at a if and only if there is L at the context level such that, for all ultrasmall dx,*

$$f(a + dx) = f(a) + L \cdot dx + \varepsilon \cdot dx,$$

where $\varepsilon \simeq 0$. If f is differentiable at a, then $L = f'(a)$.

"Ultrasmall" and $\simeq$ are relative to the context level $\mathbf{V}(f, a)$. The increment equation shows immediately that differentiability implies continuity.

Writing, as usual, $\Delta f(a) = f(a + dx) - f(a)$ and $df(a) = f'(a) \cdot dx$, the increment equation can be restated as

$$\frac{\Delta f(a)}{dx} \simeq f'(a) = \frac{df(a)}{dx}, \qquad \text{for ultrasmall } dx \neq 0.$$

Theorem A.5. (Chain Rule) *Let f and g be functions such that g is differentiable at a and f is differentiable at $g(a)$. Then the composition $f \circ g$ is differentiable at a and*

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

Proof. Fix a level where f, g, and a appear. Let $dx \neq 0$ be ultrasmall. Since g is differentiable at a, we have $\Delta g(a) = g'(a) \cdot dx + \varepsilon \cdot dx$, for some $\varepsilon \simeq 0$, by the increment equation. Since f is differentiable at $g(a)$ and $\Delta g(a) \simeq 0$, we have

$$f(g(a) + \Delta g(a)) - f(g(a)) = f'(g(a)) \cdot \Delta g(a) + \delta \cdot \Delta g(a),$$

for some $\delta \simeq 0$, again by the increment equation. Hence

$$\begin{aligned} \frac{\Delta f(g(a))}{dx} &= \frac{f(g(a + dx)) - f(g(a))}{dx} = \frac{f(g(a) + \Delta g(a)) - f(g(a))}{dx} \\ &= f'(g(a)) \cdot \frac{\Delta g(a)}{dx} + \delta \cdot \frac{\Delta g(a)}{dx} \simeq f'(g(a)) \cdot g'(a) \end{aligned}$$

because $\Delta g(a)/dx \simeq g'(a)$ (hence is limited), so $\delta \cdot (\Delta g(a)/dx) \simeq 0$, as $\delta \simeq 0$. □

The statement defining $f'(x)$ is internal, so $f' : x \mapsto f'(x)$ is a function, and appears at every level where f does, by the Definition Principle.

Proofs involving double limits can often be simplified with the help of yet another principle about levels. We introduce it now and give two examples.

Axiom VIII (Density of Levels). If a real number $\varepsilon \neq 0$ is ultrasmall relative to $\mathbf{V}$ and $\varepsilon \in \mathbf{V}'$, then there is $\mathbf{V}^+$ and a real number $\delta \in \mathbf{V}^+$, $\delta \neq 0$, such that δ is ultrasmall relative to $\mathbf{V}$ and ε is ultrasmall relative to $\mathbf{V}^+$. (Note that $\mathbf{V} \subset \mathbf{V}^+ \subset \mathbf{V}'$.)

Example We paraphrase the historic argument due to Euler and use density of levels to show that

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!}.$$

We assume that convergence of $\lim_{n\to\infty}\sum_{k=0}^{n} 1/k!$ to some real number L has been established, say by the ratio test. We fix an arbitrary level $\mathbf{V}$ where L appears and prove that $\big(1+1/N\big)^N \simeq L$ for every ultralarge positive integer N.

By the Density of Levels, there is a level $\mathbf{V}^+$ and a positive ultralarge integer $M \in \mathbf{V}^+$ such that N is still ultralarge relative to $\mathbf{V}^+$. We use $\overset{+}{\simeq}$ to denote $\simeq$ relative to the level $\mathbf{V}^+$. By the binomial formula, we have

$$\left(1+\frac{1}{N}\right)^N = \sum_{n=0}^{N}\frac{N\cdot(N-1)\cdots(N-(n-1))}{n!}\cdot\frac{1}{N^n} =$$

$$\sum_{n=0}^{M}\frac{1\cdot(1-1/N)\cdots(1-(n-1)/N)}{n!} + \sum_{n=M+1}^{N}\frac{1\cdot(1-1/N)\cdots(1-(n-1)/N)}{n!}$$

But on the one hand,

$$\sum_{n=0}^{M}\frac{1\cdot(1-1/N)\cdots(1-(n-1)/N)}{n!} \overset{+}{\simeq} \sum_{n=0}^{M}\frac{1}{n!}.$$

To see this, note that for all $n \leq M$ we have $1 \geq 1\cdot(1-1/N)\cdots(1-(n-1)/N) \geq (1-(n-1)/N)^M \overset{+}{\simeq} 1$, because $(n-1)/N \overset{+}{\simeq} 0$ and $x \mapsto x^M$ is continuous at $x = 1$. Each term on the left side is thus $1/n! + \varepsilon_n$ for $\varepsilon_n \overset{+}{\simeq} 0$. Let $\varepsilon = \max\{|\varepsilon_0|, \ldots, |\varepsilon_M|\}$; the left sum is $\sum_{n=0}^{M} 1/n! + \sum_{n=0}^{M}\varepsilon_n$, where $|\sum_{n=0}^{M}\varepsilon_n| \leq \varepsilon\cdot(M+1) \overset{+}{\simeq} 0$.[10]

On the other hand,

$$0 \leq \sum_{n=M+1}^{N}\frac{1\cdot(1-1/N)\cdots(1-(n-1)/N)}{n!} \leq \sum_{n=M+1}^{N}\frac{1}{n!} \simeq 0,$$

since M is ultralarge and $\lim_{n\to\infty}\sum_{k=0}^{n} 1/k!$ exists.

[10] Similar arguments can be used to prove in general that cases (6) and (7) in Proposition A.1 hold for M terms and factors, as long as $M \in \mathbf{V}$.

It follows that

$$\left(1+\frac{1}{N}\right)^N \simeq \sum_{n=0}^{M} \frac{1}{n!}.$$

But M is ultralarge, so $\sum_{n=0}^{M} 1/n! \simeq L$, and we conclude that $(1+1/N)^N \simeq L$. □

Theorem A.6. (L'Hospital's Rule for ∞/∞) *Let f and g be differentiable in a deleted neighborhood of a. Suppose that $\lim_{x\to a} |f(x)| = \lim_{x\to a} |g(x)| = \infty$ and $\lim_{x\to a} f'(x)/g'(x)$ exists. Then*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Proof. We follow the argument in [Benninghofen and Richter, 1987]. And let $f, g, a \in \mathbf{V}$. Let

$$L = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

and let $x \simeq a$, $x \neq a$. Assume that $x > a$ (the case $x < a$ is similar).

It is easy to see that, by the Density of Levels, we can choose $y > a$, $y \simeq a$ relative to $\mathbf{V}$, such that $x \simeq a$ relative to $\mathbf{V}^+ \supseteq \mathbf{V}$ and $y \in \mathbf{V}^+$. We use $\overset{+}{\simeq}$ when we work relative to the finer level $\mathbf{V}^+$. We have $x \overset{+}{\simeq} a$ and necessarily $a < x < y$.

By the Cauchy mean value theorem,

$$\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(c)}{g'(c)} \quad \text{for some } c \in (x, y).$$

But $c \simeq a$, so $f'(c)/g'(c) \simeq L$. Since $\lim_{x\to a} |f(x)| = \lim_{x\to a} |g(x)| = +\infty$, and $x \overset{+}{\simeq} a$, we have $f(x), g(x) \overset{+}{\simeq} \pm\infty$. Hence $f(y)/f(x) \overset{+}{\simeq} 0$ and $g(y)/g(x) \overset{+}{\simeq} 0$. The proof is completed by observing that

$$L \simeq \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f(x)}{g(x)} \cdot \frac{1-f(y)/f(x)}{1-g(y)/g(x)} \overset{+}{\simeq} \frac{f(x)}{g(x)}$$

□

We conclude this sampler with some brief remarks about integration. At the most elementary level it suffices to integrate continuous functions. If f is continuous on $[a, b]$, we define the *definite integral of f from a to b* as follows. We fix a level where f, a, b appear, and a positive integer N ultralarge relative to this level. Let $dx = (b-a)/N$ and $x_i = a + i \cdot dx$, for $i = 0, \ldots, N$. Then

$$\int_a^b f(x) \cdot dx = \mathbf{n}\left(\sum_{i=0}^{N-1} f(x_i) \cdot dx\right).$$

It is not too hard to prove that the definition is independent of the choice of N, and to derive linearity and additivity properties of definite integrals (for continuous functions) from it.

The fundamental problem of integration is to find the original function F, given its derivative $F' = f$. This is particularly easy if the derivative is continuous.

Theorem A.7. (Uniform Increment Equation) *Let F be a function differentiable on a closed interval $[a, b]$ (one-sided derivatives, defined in the obvious way, suffice at end-points). The following conditions are equivalent.*

(1) F' is continuous on $[a, b]$.

(2) For all $x \in [a, b]$ and all $dx \simeq 0$ such that $x + dx \in [a, b]$,

$$F(x + dx) - F(x) = F'(x) \cdot dx + \varepsilon \cdot dx,$$

where $\varepsilon \simeq 0$.

We stress that dx is ultrasmall relative to the context level $\mathbf{V}(F, a, b)$, which is independent of x; compare this with the "local" increment equation, Theorem A.4. The uniform increment equation gives

$$F(x_{i+1}) - F(x_i) = F'(x_i) \cdot dx + \varepsilon_i \cdot dx = f(x_i) \cdot dx + \varepsilon_i \cdot dx$$

where $\varepsilon_i \simeq 0$, for $i = 0, \ldots, N - 1$.

By adding up these equations we obtain

$$F(b) - F(a) = \sum_{i=0}^{N-1} f(x_i) \cdot dx + \sum_{i=0}^{N-1} \varepsilon_i \cdot dx \simeq \sum_{i=0}^{N-1} f(x_i) \cdot dx.$$

Hint: let $\varepsilon = \max\{|\varepsilon_0|, \ldots, |\varepsilon_{N-1}|\}$ and observe that $\left|\sum_{i=0}^{N-1} \varepsilon_i \cdot dx\right| \leq \varepsilon \cdot (b - a)$.

Since $F(b) - F(a)$ appears at the context level, we have

$$F(b) - F(a) = \mathbf{n}\left(\sum_{i=0}^{N-1} f(x_i) \cdot dx\right) = \int_a^b f(x) \cdot dx.$$

This formula (the first fundamental theorem of calculus) solves the integration problem for $b > a$; the same formula works for $b \leq a$.

The Riemann theory of integration for bounded functions on $[a, b]$ can be obtained as a slight generalization of this approach.

Definition A.8. *A* tagged partition *of an interval $[a, b]$ is a pair $(\mathcal{P}, \mathcal{T})$ where $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ with $a = x_0 < x_1 < \cdots < x_n = b$, and $\mathcal{T} = \{t_0, \ldots, t_{n-1}\}$ with $x_i \leq t_i \leq x_{i+1}$, for $i = 0, \ldots, n - 1$. We let $dx_i = x_{i+1} - x_i$. The* Riemann sum *$\sum(f; \mathcal{P}, \mathcal{T})$ is defined as*

$$\sum(f; \mathcal{P}, \mathcal{T}) = \sum_{i=0}^{n-1} f(t_i) \cdot dx_i.$$

Relative to a context level (that is, where f, a, b appear), a tagged partition $(\mathcal{P}, \mathcal{T})$ is fine *if all dx_i are ultrasmall. A bounded function f on $[a, b]$ is* Riemann integrable *if there is $R \in \mathbb{R}$ at the context level such that $\sum(f; \mathcal{P}, \mathcal{T}) \simeq R$ for all fine tagged partitions $(\mathcal{P}, \mathcal{T})$ of $[a, b]$. If this is the case, we let $\int_a^b f(x) \cdot dx = R$. It is not hard to show that continuous functions remain integrable according to this definition.*

The above considerations explain why continuity of $f = F'$ is essential for the recovery of F via Riemann integrals: it implies the *uniform* version of the increment equation, which allows us to use fine partitions, that is, partitions where each dx_i is ultrasmall relative to the context level, independent of the tag t_i. This suggests replacing fine partitions in the definition of the Riemann integral by "superfine" partitions, where a tagged partition $(\mathcal{P}, \mathcal{T})$ is *superfine* if each dx_i is ultrasmall relative to a level that also contains t_i. It turns out that superfine partitions in this sense do not exist [Hrbacek, 2010]. One needs to use, for this purpose, a weaker notion of relative ultrasmallness, due to Benninghofen and Richter [Benninghofen and Richter, 1987] and Gordon [Gordon, 1997].

For $a \in \mathbb{R}$, we say that a real number ε is *a-ultrasmall* relative to $\mathbf{V}$ if $|\varepsilon| < f(a)$ for all positive real-valued functions $f \in \mathbf{V}$ with $a \in \operatorname{dom} f$.

In the definition of superfine partitions we require that each dx_i be t_i-ultrasmall (relative to the context level). With this modification, the idea works well; it is developed in [Benninghofen, 1984] and [Hrbacek et al., 2010b], and yields the generalized Riemann integral, also known as the Henstock-Kurzweil integral, which agrees with the Lebesgue integral on absolutely integrable functions. In this theory, every derivative is integrable and its integral is the original function, up to a constant.

Final Remark. This paper demonstrates that the fundamental concepts and theorems of analysis can be introduced and proved without any use of the ε–δ techniques that students find notoriously difficult to learn. Of course we recognize that it is crucially important that students bound for more advanced mathematics courses master these techniques. They can be introduced naturally, almost as an afterthought, when studying estimation and elementary topology of the real line. Here we show only that our definition of continuity is equivalent to the usual ε–δ one.

Let us assume that f is continuous at a according to Definition A.3. Given $\varepsilon > 0$, let $\delta > 0$ be ultrasmall relative to some level $\mathbf{V}$ where f, a, ε appear. Then $|x - a| < \delta$ implies $|f(x) - f(a)| < \varepsilon$.

Conversely, suppose that f is continuous at a according to the ε–δ definition, and let $x \simeq_{\mathbf{V}} a$ where $f, a \in \mathbf{V}$. Given $\varepsilon > 0$, $\varepsilon \in \mathbf{V}$, there is a $\delta > 0$ such that, for all z, $|z - a| < \delta$ implies $|f(z) - f(a)| < \varepsilon$. By Closure, some such δ appears at the level $\mathbf{V}$. Then $|x - a| < \delta$, and so $|f(x) - f(a)| < \varepsilon$. As this is true for every $\varepsilon \in \mathbf{V}$, we conclude that $f(x) \simeq_{\mathbf{V}} f(a)$.

Acknowledgements. We are grateful to R. Almeida, I. van den Berg, P. Ehrlich, R. Goldblatt, E. Gordon, W. Henson, P. Loeb, V. Neves, Y. Péraire, V. Prabhu, J. Roquette, S. Sanders, F. Soares, K. Stroyan, T. Todorov, P. Zlatoš, and the anonymous referees for their comments on earlier versions of this paper, some encouraging, some critical, but all helpful.

Appendix B

Internal formulae of FRIST

The following is co-authored with Karel Hrbacek and Olivier Lessmann. It is a series of extracts of a book in preparation [Hrbacek et al., 2010b] on analysis with ultrasmall numbers. These exerpts discuss the question of internality of formulae.

Definition 17. *A statement is **internal** if it is about its observation level.*

Formally, a statement is internal if it is of the form

$$\mathcal{P}(x_1, \ldots, x_k; \mathbf{V}(x_1, \ldots, x_k)).$$

*A statement is **external** if it is not internal.*

Speaking more loosely, a statement is internal if it either does not mention levels at all (i.e., it is a traditional mathematical statement), or if every level mentioned in the statement contains all the objects the statement is about.

[...]

Here are some examples.

(1) *$x \in \mathbb{N}$ is a prime number.*

The statement does not mention any levels and therefore it is internal.

(2) *y is ultrasmall relative to the level of x.*

In detail: $0 \neq |y| < r$, for every $r > 0$, $r \in \mathbf{V}(x)$.

The parameters are x and y, but the statement refers to $\mathbf{V}(x)$, not to $\mathbf{V}(x, y)$, so it is not internal.

(3) *There exists an x which is ultrasmall relative to the coarsest level $\mathbf{V}(0)$.*

The variable x is bound (by the existential quantifier "there exists"), so x is not a parameter. (One could just as well say that there exists a y which is ultrasmall relative to $\mathbf{V}(0)$, or even, that there is a number which is ultrasmall relative to $\mathbf{V}(0)$.) This statement has no parameters, so its observation level is $\mathbf{V}(0)$. The statement is internal.

(4) *Each x ultrasmall relative to the level of y is such that x^2 is ultrasmall relative to the level of y.*

Again, the variable x is bound (by the universal quantifier "for all"), and this statement has parameter y. The statement refers only to the level of y, and is therefore internal. (In fact, it is equivalent to $\lim_{x\to 0} x^2 = 0$.)

[...]

An important example: By definition, a function f is continuous on an interval I if f is continuous at a, for all $a \in I$. The defining statement does not mention levels, and hence it is internal. One might object that we defined continuity of f at a using levels, and therefore it is not a concept of traditional mathematics, at least not until our definition is proved equivalent to the traditional definition. However, when the definition of continuity on an interval is restated directly in terms of levels

f is continuous on I if and only if

$$f(x) \simeq f(y) \text{ whenever } x, y \in I,\ x \simeq y$$

holds relative to $\mathbf{V}(f, I, x)$,

it becomes apparent that it is indeed internal, because the level $\mathbf{V}(f, I, x)$ it refers to contains the parameters f and I of the definition. More formally, with the help of stability one can rephrase the definition as follows: "f is continuous on I if and only if for every level $\mathbf{V} \supseteq \mathbf{V}(f, I)$, if $x, y \in I$, $x \in \mathbf{V}$ and $x \simeq_{\mathbf{V}} y$, then $f(x) \simeq_{\mathbf{V}} f(y)$," which is clearly a statement about $\mathbf{V}(f, I)$.

This example illustrates a general phenomenon. We say that a defined *concept* is *internal* if its defining statement is internal. Continuity of a function at a point, differentiability of a function at a point, limit, and integral are some of the examples of internal concepts introduced so far.

*From now on, we allow statements about **V** and internal statements to refer to previously defined internal concepts. We postulate that the stability principle remains valid when "statements about **V**" are understood in this extended sense.*

[...]

Restricted definitions are for sets given in the form

$$\{(x, y) \mid x \in A \text{ and } P(x, y)\}$$

i.e., where the elements are bound by belonging to some predefined set.

[...]

One of the parameters of such a definition is A; the **defining statements** $\mathcal{P}(x)$ or $\mathcal{P}(x, y)$ can involve additional parameters. Note that x and y do not count as parameters of the definition, because they become bound in it; for example, "For every x, $x \in \{x \in A : \mathcal{P}(x)\}$ if and only if $x \in A$ and $\mathcal{P}(x)$."

In our richer language, even this restriction is not sufficient. It is possible to make statements about natural numbers that do not define sets! Here is the simplest example.

Example

Consider the statement "$n \in \boldsymbol{V}(0)$." There is no set S such that $n \in S$ if and only if $n \in \mathbb{N}$ and $n \in \boldsymbol{V}(0)$. In other words, $\{n \in \mathbb{N} : n \in \boldsymbol{V}(0)\}$ is not a set.

The proof is easy. Assume $S = \{n \in \mathbb{N} : n \in \boldsymbol{V}(0)\}$ is a set. Then

(1) *$0 \in S$, because $0 \in \mathbb{N}$ and $0 \in \boldsymbol{V}(0)$, and*

(2) *If $n \in S$, then $n \in \mathbb{N}$ and $n \in \boldsymbol{V}(0)$, so $n+1 \in \mathbb{N}$ and $n+1 \in \boldsymbol{V}(0)$ (the latter follows from the closure principle: $n+1$ is defined from the parameter $n \in \boldsymbol{V}(0)$); hence $n+1 \in S$.*

By the principle of mathematical induction, $\mathbb{N} = S$, i.e., all $n \in \mathbb{N}$ are in $\boldsymbol{V}(0)$, a contradiction.

Fortunately, there is a very simple rule that singles out the acceptable definitions of sets and functions.

Definition Principle

Restrictive definitions whose defining statement is internal define sets and functions. Moreover, these sets and functions are observable at the level that contains all the parameters of the definition.

Here, "sets" are the legitimate sets of the theory i.e., internal sets

[...]

The definition principle generalizes the axioms of separation and replacement. We recall that a statement $\mathcal{P}(x_1, \dots, x_k)$ is called **internal** if it is of the form

$$\mathcal{Q}(x_1, \dots, x_k; \mathbf{V}(x_1, \dots, x_k)).$$

Theorem 16. *The definition principle follows from the axioms. More formally:*

(1) *Let $\mathcal{P}(x, A, x_1, \dots, x_k)$ be an internal statement. Then*
$(\exists B \in \boldsymbol{V}(A, x_1, \dots, x_k))(\forall x)(x \in B \leftrightarrow x \in A \wedge \mathcal{P}(x, A, x_1, \dots, x_k))$.

(2) *Let $\mathcal{P}(x, y, A, B, x_1, \dots, x_k)$ be an internal statement. If*

$$(\forall x \in A)(\exists! y \in B)\mathcal{P}(x, y, A, B, x_1, \dots, x_k),$$

then there is a function $F : A \to B$, $F \in \boldsymbol{V}(A, B, x_1, \dots, x_k)$, such that

$$(\forall x \in A)\mathcal{P}(x, F(x), A, B, x_1, \dots, x_k).$$

Proof. (1) By standardization applied to $\mathbf{V} = \mathbf{V}(A, x_1, \dots, x_k)$, there is a set $B \in \mathbf{V}$ such that

$$(\forall x \in \mathbf{V})(x \in B \leftrightarrow x \in A \wedge \mathcal{Q}(x, A, x_1, \dots, x_k; \mathbf{V})).$$

By stability, for any $\mathsf{V}' \supseteq \mathsf{V}$,

$$(\forall x \in \mathsf{V}')(x \in B \leftrightarrow x \in A \wedge \mathcal{Q}(x, A, x_1, \ldots, x_k; \mathsf{V}')).$$

Let x be arbitrary and $\mathsf{V}' = \mathsf{V}(x, A, x_1, \ldots, x_k)$; the above statement gives

$$x \in B \leftrightarrow x \in A \wedge \mathcal{Q}(x, A, x_1, \ldots, x_k; \mathsf{V}'),$$

but $\mathcal{Q}(x, A, x_1, \ldots, x_k; \mathsf{V}(x, A, x_1, \ldots, x_k)) \leftrightarrow \mathcal{P}(x, A, x_1, \ldots, x_k)$ because $\mathcal{P}$ is internal.

(2) The proof is similar.

□

Appendix C

Teaching analysis with ultrasmall numbers

This article was published in the Mathematical Teaching and Research Journal, August 2009. Even though chronologically is was published before the article of appendix A, it should be considered as a sequel: a presentation of how FRIST works in the classroom. It is referred to as [O'Donovan, 2009] in the bibliography.

Abstract: *When teaching an introductory course of analysis or calculus, many colleagues resort to infinitesimals as useful metaphors to convey understanding. For some ten years or so a group of teachers in Geneva (Switzerland) has been striving to go beyond the metaphor — the holy grail being to introduce analysis in a mathematically rigorous way, yet close enough to intuitive concepts and with a lower degree of technical difficulty than classical approaches. In collaboration with professor Hrbacek (CUNY) a first version of this approach is now completed and has been used in several classes this year. In the following we present classroom material and discuss how students respond.*

After some attempts at teaching what we might call "naive nonstandard analysis" [O'Donovan and Kimber, 2006] it became obvious that there are pedagogical advantages. The cognitive improvements are considerable. Still, if it was to be more than a metaphor and truly mathematics, some foundational research had to be done. This lead to the collaboration with professor Hrbacek (CUNY) who had developed a form of nonstandard analysis which seemed best fit to meet our requirements [O'Donovan, 2007].

A preliminary remark is necessary for the mathematically trained reader. Many of us have been taught that infinitesimals do not exist, or that if they do exist they are not real numbers. Here, we will be talking about extremely small quantities, called *ultrasmall numbers*, within the real numbers. This may be felt as conflicting knowledge by some. We can testify that students, being ignorant of the subject, do not share this feeling (for this once, we will say that their ignorance is a blessing!) The existence of these ultrasmall and corresponding ultralarge real numbers are due to extra axioms which will not be discussed here. It should be enough to know that it has been proven in [Hrbacek, 2004] that these axioms add no contradiction to the universe of mathematics i.e., they are perfectly safe. [11]

This description does not aim at being complete but rather at showing why we consider that it provides an interesting option for teaching and learning analysis.

Context

In Geneva high schools (pre-university: years 12 and 13) students have had no prior teaching of calculus when they start studying analysis and they will be required to be able to prove most of the theorems they use. They know the definition of a function in terms of input set, output set and a rule. They can solve $f(x) = 0$ in simple polynomial or trigonometric cases, and by plotting selected points, they draw rough approximations of curves. The notion of slope applies only to straight lines. Classical courses in analysis have dropped teaching the full blown Weierstrassian ε-δ theory as it is considered too complicated. Instead, a hand-waving definition of the limit is given and the main properties (sum of limits is limit of sums, etc.) are given without proof. The motivation for our research is to increase meaning and yet bring back some mathematical rigor in the classroom.

New concepts

First exercises are meant to make students suggest answers following their intuition and are given with no prior explanation.

If h is a positive value which is extremely small (even smaller than that!), what can you say about the size of h^2, $2h$ and $-h$? What can you say about $2 + h$ and $2h$? What can you say about $\frac{1}{h}$?

If N is a positive huge number (really very huge!), what can you say about N^2, $2N$ and $-N$? What can you say about $N + 2$ and $N - 2$? What can you say about $\frac{1}{N}$? What can you say about $\frac{N}{2}$?

Students respond well to these exercises and it is clear that the concept of "infinitesimal" is already present in their mind. (How this concept arises as part of what becomes intuitive is not addressed here even though it is a very interesting aspect.) The metaphor of a microscope is classical in nonstandard analysis, but zooming in or out is also a useful alternative. Using a computer and a beamer, students can be shown what happens when one zooms in on the curve of a function. Then they are asked what would appear under a superpowerful zoom. And then by zooming out, faraway parts become visible. Occasionally, students will have difficulty in "guessing" that the reciprocal of an extremely small number is extremely large. This shortcoming would also be a drawback in a more tradi-

[11] In very short, this work is related to nonstandard analysis of which there are two main trends: Robinson's hyperreal number system [Robinson, 1961], where no axioms are added but an extension of the reals results from a rather complicated construction; and Nelson's IST [Nelson, 1977] in which extra axioms are added. Hrbacek's approach, FRIST [Hrbacek, 2007], extends Péraire's RIST [Péraire, 1992] which is an extension of Nelson's ideas.

tional approach with limits so this type of exercise is probably necessary at a preliminary stage for analysis in general.

A discussion with students on how small can positive numbers be, or how long is an instant, always lead to answers involving infinitesimals. Our school has students of 84 nationalities and so many different cultures and we have not yet had a single student for whom the concept of infinitesimal is unknown. Students appear to feel safe about the fact that their guesses are correct; it is possible to tell them that (in their study of calculus) there will be almost no new concept, but we will formalize these and they will discover the power of this formalization; they will reach new conclusions, yet at a very fundamental level, they already know the essential concepts. (Which does not mean that everything will be easy...)

Using the yet undefined term of "vanishingly small" the following exercise introduces all the fundamental concepts for the derivative.

Make a sketch and calculate the average slope over an interval of vanishingly small length h of the function $f : x \mapsto x^2$ centered on $\langle 1, f(1) \rangle$.

When you zoom out so that h becomes vanishingly small, what can you say about the value of the slope?

Here, the slope is obviously $2 + h$ and when zooming out, the slope becomes indistinguishable from 2.

The key innovation of Hrbacek's approach (with respect to other nonstandard analysis approaches) is hinted at in the following exercise. There are not only two levels: "ordinary numbers" and numbers which are visible only through a metaphorical microscope; there are always other "finer" levels which are invisible without an extra microscope.

Now consider that you have shrunk and that you are of the size of the h of the above exercise. Imagine also an extremely small positive value k – *extremely small relative to your new size!* what can you say about h? What can you say about k^2? What can you say about $2 + h$? What can you say about $2 + k$?

Students' first guess is almost always correct: h is not ultrasmall anymore because the context has changed. $2 + h$ is not extremely close to 2 but $2 + k$ is, and k^2 is really very small!

After introductory exercises, formalization may start.

Axiom 1: Levels. Given real numbers $x_1, \ldots, x_n$ there is a coarsest level at which $x_1, \ldots, x_n$ appear. If a number appears at a level, it appears at all finer levels. Every number defined in a unique way, without using the concept of level, appears at the coarsest level.

Definition: ultrasmall and ultralarge. Relative to a level, a real number is ultrasmall if it is nonzero and its absolute value is smaller than any positive number appearing at that level. A real number is ultralarge if its absolute value is larger than any positive number appearing at that level. Two numbers are ultraclose if their difference is ultrasmall or zero.

Axiom 2 : Ultrasmall and Ultralarge numbers. Relative to any level, there exist ultralarge and ultrasmall numbers.

Of course, this implies that an ultrasmall number does not appear at the context level.

Note for mathematicians: axioms 1 and 2 are about the level of individual numbers, not about the set of numbers. The definition is later extended to say that an object appears at the level of the parameters needed to define it. Levels are not needed to define the real numbers as a whole, hence the set *of real numbers appears at coarsest level. As does the interval* $[0;1]$*. Nonetheless, these sets contain numbers of all levels (due to axiom 2). Intuitively, the interval* $[0;1]$ *is defined using 0 and 1 but it contains ultrasmall numbers: a set and its elements do not share the same properties. Students do not have the reflex of collecting objects into sets, so this causes difficulty only to the newcomer mathematician.*

We have observed that for some students, the concept of "extremely large" seems more natural at first than "extremely small". Yet when considering the reciprocal, the extremely small becomes easier to grasp because, as a student put it "we know where it is" whereas the extremely large are somewhere far away. The concept of extremely large has a somewhat fleeting quality to it, quantities "moving away". Extremely small values don't move around. They have no space for that. So once these extremely small quantities have been grasped, it is possible to use the reciprocal again and this time, extremely large numbers, being linked to "fixed" extremely small numbers, can be understood as numbers which are not moving away.

The terms "infinitesimal" and "infinitely large" are not used for several reasons. Infinite sets are sets whose cardinality cannot be given by a natural number. Here, we have ultralarge integers: they are finite yet huge. So to avoid ambiguities we will not keep the usage common to nonstandard analysis practice and change to "ultralarge". Then of course, the reciprocal "ultrasmall" seems the obvious word, as is the word "ultraclose". In fact, the "ultra" prefix allows to define other concepts in a very natural way. Between x and $x+h$, if h is ultrasmall then the tangent can be an "ultraprecise" approximation to the function, the integral is defined using "ultrathin" slices, etc. Another reason is that even though students have a good intuition about infinity, some consider (not wrongly) that infinity plus one is infinity. As we will consider that if N is ultralarge then $N+1 > N$, we avoid misunderstandings by introducing a new word and saying that it is "almost" what they think it intuitively is. Also, if N is greater than zero, then $-N$ is less than N, it is smaller, whereas we need to talk about magnitudes as given by distances to zero. By defining ultralarge to be very big in absolute value and ultrasmall to be very close to zero, we are in fact closer to many people's intuitive use of the word.

Definition: Context level. A context level of a property, a function or a set, is a level at which all its parameters appear.

The parameters are easily listed. A context level is thus easily determined: the words "ultrasmall" and "ultralarge" will always be understood relatively to the context level and if a and b are ultraclose relative to the context level this is written

$$a \simeq b.$$

(A context level is understood, not explicitly given. This fact will prove to be important.)

Note for mathematicians: We say a context level and not the context level. In effect, we start by fixing a context at which all parameters of a function appear, then we consider the context level, that which has been fixed.

Relative to a context level at which a appears, if h is ultrasmall, then $a + h \simeq a$. Intuitively, it seems reasonable that $(a + h)^2 \simeq a^2$ which means that we must be able to show that $2ah \simeq 0$ and $h^2 \simeq 0$. This last one is obvious as, wlog for positive h, we have $0 < h < 1 \Rightarrow 0 < h^2 < h < 1$, but the first statement needs an extra axiom:

Axiom 3 : Closure. $f(x)$ appears at the level of f and x.

Closure means that the result of operations appear at the level of the numbers involved in the operation. Numbers of finer level do not appear unless they have been explicitly summoned.

Note for mathematicians: We say that numbers appear at a given level, that they are of a level and not that numbers are in levels. By the closure principle, we get that $n + 1$ appears at the same level as n, yet we cannot use this argument to prove that all whole numbers appear at the coarsest level. ***Levels are not sets*** *and induction cannot be used on collections which are not sets. There are many cases of collections which are not sets (the collection of all sets for instance) but usually we do not need to consider these. Here, we are suddenly exposed to these unusual objects. But we will never really need to collect them and only use the fact that ultrasmall numbers, for instance, do not appear at the context level.*

There are two other rules which apply but are not necessarily explicitly specified to students: if a reference is made to levels, it may only be to the context level. If this rule is not followed, pathological objects can be defined, but as all definitions are given here relative to the context level, a student would need to invent her own definitions in order to define such illegal objects. We do not use a symbol to express that a appears at the level of b (such a symbol exists in the full theory). Everything has been written in such a way as to make illegal references very difficult. ($a \simeq b$ for instance, automatically refers to a context level because it has been defined that way, hence ultraclose relative to another level than a context level would require the invention of another symbol.)

Another axiom which needs to be mentioned but requires no specific comment for teaching is that if a statement is true when referring to a context level, it remains true when referring to any context level. It expresses the fact that it is not very important which level is used as a context level provided it is fine enough. Hence when studying two different functions which do not appear at the same level, a context level will be any level at least as fine as the finer of the two, yet all results found considering this context level about the coarser function remains true. For instance, the derivative does not depend on which context level is used. This principle is analogous to nonstandard analysis' transfer.

First rules. If (relative to some context level) ε is ultrasmall, its reciprocal $1/\varepsilon$ is ultralarge. By contradiction: Without loss of generality assume $\varepsilon > 0$ and that $1/\varepsilon < b$ for some $b > 0$ appearing at the context level. Then we would have $1/b < \varepsilon$. By closure,

$1/b$ appears at the context level and this contradicts that ε is ultrasmall. Hence $1/\varepsilon$ is ultralarge.

If $a > 0$ is not ultralarge and ε is ultrasmall (relative to some context level), then $a \cdot \varepsilon \simeq 0$. If a appears at a context level and $\varepsilon \simeq 0$ (also assumed to be positive), then $a \cdot \varepsilon \simeq 0$ (otherwise $a \cdot \varepsilon \geq b$ for some b appearing at the context level and $\varepsilon \geq b/a$: a contradiction. Then if a is not ultralarge it is less than some c appearing at the context level hence $a \cdot \varepsilon < c \cdot \varepsilon \simeq 0$.

With these rules, if a appears at a context level and N is ultralarge, then

$$(a + 1/N)^2 = a^2 + \underbrace{2a/N}_{\simeq 0} + \underbrace{1/N^2}_{\simeq 0} \simeq a^2$$

Rules such as $a \simeq x$ and $b \simeq y$ imply $a \cdot b \simeq x \cdot y$ (provided a and b are not ultralarge) are simple exercises, using decompositions such as $x = a + \varepsilon$ with $\varepsilon \simeq 0$.

An important property is that if a and b appear at the context level and $a \simeq b$, then $a = b$. By closure, $a - b$ appears at the context level. If $a \simeq b$ then $a - b$ is ultrasmall or zero. As it appears at the context level it cannot be ultrasmall, hence it is zero and $a = b$.

With these axioms and definitions students can start studying derivatives. These properties and rules are almost certainly not completely understood by the students at this stage, but familiarity will be best achieved by working with them. Except for one other axiom which would be necessary for continuity, no extra "black boxes" will be introduced: it is not necessary to introduce that the limit of a sum is the sum of the limits and such. Everything here follows from properties of numbers. As these properties are used again and again, students gradually become proficient.

Derivative

One of the advantages of this approach is that it is possible to start by the derivative without spending time on the difficult concept of limit. Starting with derivatives has the advantage of introducing something really new. For students, continuity is either considered trivial or not understood for the more sophisticated aspects. Similarly with limits. Asymptotes have been studied in an experimental way the year before, using the calculator to find values of, say, $1/x$ for x close to zero. And all teachers know that "new" is always more interesting for students.

Let

$$f : x \mapsto x^2.$$

For the slope of the curve of f at $x = 3$ we look at the variation of the function for an increment $h > 0$ of the independent variable. A context level is a level where the parameters 3 and 2 appear: the coarsest level or any finer level. Let the increment h be ultrasmall — it is ultrasmall, not zero.

The *average slope* between 3 and $3 + h$ is

$$\frac{f(3+h)-f(3)}{h} = \frac{(3+h)^2-3^2}{h} = 6+h.$$

We obtain the same result with a negative increment. 6 and h can both be seen at the level of h hence $6+h$ can be seen at the level of h. The part of $6+h$ that can be seen at the context level (without a microscope) is 6, hence the slope that we can see at the coarsest level is 6 so we *define*

$$f'(3) = 6.$$

This result does not depend on the choice of the ultrasmall h and is called *the derivative of f at $x = 3$*.

General case: when studying the slope of f at x we consider a context level at which x and 2 appear and look at the part of the result which is at the context level, but we use increments which are ultrasmall (relative to the context level).

$$\frac{(x+h)^2-x^2}{h} = 2x+h$$

and the part of the result at the context level is $2x$. And this is for all x, which is the classical result. ($2x$ appears at the level of x because $2x = x+x$: a familiar operation does not change the level of the result.) We write $f'(x) = 2x$ and say that f *is differentiable at* x if $f'(x)$ exists.

At this stage, one might say that there is no big improvement on the classical definition with the "h tends to zero." Yet cognitively there is. Teaching limits, we are continuously confronted with questions such as, yes but does h "reach" zero, and if it does not, where does it stop? "Tending to" and "reaching" are – in a student's mind – completely different concepts, and defining the limit as that which we never reach is well known to be one of the major stumbling blocks to teaching analysis. An explanation which deals with quantifiers is difficult at this stage for two completely different reasons. On the mathematical side, any concept which uses quantifiers is more difficult to understand and if it requires an alternation of quantifiers it will be even harder. On the metaphorical side, the dynamic metaphor that x moves towards a is powerful but is destroyed when quantifiers are added to make the definition rigorous. In our view, this is the major difficulty of analysis, in order to understand well the metaphor, the metaphor has to be broken, and once it is broken, what is there really left? These consideration are what lead us on the path to nonstandard analysis. Appearing at the context level or more metaphorically: the part of the result which can be seen without a microscope are surprisingly well understood and offer no resistance.

A consequence of the definition of the derivative is the **microscope equation**:

Suppose that f is differentiable at a. Then for any $\Delta x \simeq 0$ there exists $\varepsilon \simeq 0$ such that

$$\Delta f(a) = f(a+\Delta x) - f(a) = f'(a) \cdot \Delta x + \varepsilon \cdot \Delta x.$$

A context level is a level at which f and a appear.

This is called the Increment Equation for f at a with increment Δx. It is clear that $\Delta f(a) \simeq 0$, since $f'(a)$ appears at a context level by definition.

The increment equation can be used to prove theorems which are otherwise quite difficult at introductory level such as the chain rule. Note in particular that this proof also covers the case where $g'(a) = 0$.

Chain Rule

Let f and g be functions such that g is differentiable at a and f is differentiable at $g(a)$. Then the composition $f \circ g$, defined by $(f \circ g)(x) = f(g(x))$, is differentiable at a and

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

proof

Let Δx be ultrasmall. By the Increment Equation for g at a with increment Δx, we have

$$\Delta g(a) = g'(a)\Delta x + \varepsilon \cdot \Delta x.$$

Note that $\Delta g(a) \simeq 0$, since $g'(a)$ appears at a context level. Thus, by the Increment Equation for f at $g(a)$ with increment $\Delta g(a)$, we obtain

$$\Delta f \circ g(a) = f(g(a + \Delta x)) - f(g(a)) = f'(g(a)) \cdot \Delta g(a) + \delta \cdot \Delta g(a).$$

Dividing by Δx yields

$$f'(g(a)) \cdot \frac{\Delta g(a)}{\Delta x} + \underbrace{\delta \cdot \underbrace{\frac{\Delta g(a)}{\Delta x}}_{\simeq g'(a)}}_{\simeq 0} \simeq f'(g(a)) \cdot g'(a) \qquad \square$$

Continuity

Whereas the definition of the derivative remains very close to the definition with limits, the definition of continuity becomes drastically simpler. Here we use the traditional writing dx for an ultrasmall increment of the variable with the explicit condition that $dx \simeq 0$ but $dx \neq 0$.

The function is continuous at a if, for ultrasmall dx

$$\Delta f(a) = f(a + dx) - f(a) \simeq 0.$$

(Ultrasmall obviously refers to a context level which is a level where all parameters of f and a appear.)

We check the continuity of $f : x \mapsto x^2$ at $x = 3$. A context is given by the parameters 2 and 3. Let dx be ultrasmall, then $(3 + dx)^2 - 3^2 = 6 \cdot dx + dx^2$. This difference is ultrasmall because $6 \cdot dx$ and dx^2 are ultrasmall hence their sum also. Therefore f is continuous at 3.

In general take dx ultrasmall with respect to a level where a and 2 appear, then $(a + dx)^2 - a^2 = 2a \cdot dx + dx^2$ is ultrasmall. Notice that by the microscope equation it is immediate that any function differentiable at a is continuous at a.

The continuity of the sine function (which is often omitted at introductory level) is straightforward. By Pythagoras, $\Delta \sin(\theta)$ is less than the chord which (being a straight line) is less (in absolute value) than the arclength which is θ. If the arc is ultrasmall, then $\Delta \sin(\theta)$ is also ultrasmall.

At Geneva high school it is customary to omit proofs of theorems such as the extreme value and intermediate value theorems which are given as "black boxes". In this setting, higher level students can prove these using an extra axiom.

Axiom 4: Neighbor. If (relative to some context level) a is not ultralarge, then there is a number at the context level which is ultraclose to a.

This axiom is used here as equivalent to the completeness of the real numbers. For students, it is not as easy to understand as the other axioms as they have no notion of the completeness of the real numbers or of density or compactness. This is why we do not necessarily attempt to use it with students who take mathematics at compulsory level.

Yet at this stage the students have encountered the neighbor many times in practice: the derivative is the neighbor of the quotient. The neighbor is unique (which they have also seen in practice) because, as was pointed out in the first rules, if a and b appear at a context level and $a \simeq b$ then $a = b$.

The neighbor principle can be used to define transcendent functions such as $x \mapsto a^x$. The students know how to define a^n for integer n and how to extend the definition to $a^{p/q}$. For irrational x, a context level is given by a and x. Let $p/q \simeq x$. Then a^x is defined to be the context neighbor of $a^{p/q}$. The fact that this is well defined is maybe lengthy but not specifically difficult – and may be omitted at introductory level.

An important property of the neighbor is that if a and b appear at the context level and if $x \in [a, b]$ then the neighbor of x is in $[a, b]$. We leave this exercise to the reader.

Intermediate Value Theorem

Let f be a continuous function on $[a; b]$ with $f(a) < 0$ and $f(b) > 0$. Then there exists c, with $a < c < b$ such that $f(c) = 0$.

proof Context is given by a, b and f. Let N be ultralarge, then $dx = \frac{b-a}{N}$ is ultrasmall. Consider $x_i = a + i \cdot dx$, then $a = x_0$ and $x_N = b$. There is a point with least index $j > 0$ such that $f(x_{j+1}) \geq 0$. By choice of j we have

$$f(x_j) < 0 \leq f(x_{j+1}).$$

Let c be the context neighbor of x_j (it exists because x_j is between a and b) and $c \in [a; b]$. Then $x_j \simeq c$ and $c \simeq x_{j+1}$ because $x_j \simeq x_{j+1}$. By continuity of f at c we have

$$f(c) \simeq f(x_j) < 0 \quad \text{and} \quad f(c) \simeq f(x_{j+1}) \geq 0.$$

Hence $f(c) \simeq 0$, but as 0 and $f(c)$ appear at the context level , we conclude that $f(c) = 0$. □

Even though the proof is not difficult, it is not easy for students. The statement of the theorem itself is much more obvious than its proof and many feel that the proof obscures the understanding. It is worth showing that a function $f : \mathbb{Q} \to \mathbb{Q}$ does not satisfy this theorem. The existence of the context neighbor is all we need to characterize the real numbers. In a similar way, we can also show that a continuous function attains its maximum and minimum on a closed interval.

Limits

The concept of limit is not necessary but it is possible to define it in terms of ultrasmall values. Students need to be prepared to meet limits, and also, in Geneva it is part of the syllabus. Here, the limit, instead of being the central concept becomes merely a sort of abbreviation.

For the limit of f when x tends to a, a context level is given by a and f. There is a limit if there is a value L appearing at the context level, such that $f(x) \simeq L$ whenever $x \simeq a$, written in the usual way. It is thus possible to redefine the derivative as $\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$. The interesting aspect of this rewriting is that students will have seen the classical definition and should thus not be taken aback if, during further studies, they encounter more traditional teaching.

Integral

Rather than present the complete construction of the integral, we prefer to show how it is possible to prove that the area under a positive function is given by the antiderivative.

Consider the area under a positive continuous function f between a and x. This area will be written $A(x)$. The context is given by a, f and x. Take $dx > 0$ ultrasmall. On $[x, x + dx]$ the function has a minimum at m and a maximum at M. Hence $f(m) \cdot dx \leq \Delta A(x) \leq f(M) \cdot dx$. Divide by dx, then $f(m) \leq \frac{\Delta A(x)}{dx} \leq f(M)$. As f is continuous at x we have $f(m) \simeq f(x) \simeq f(M)$ hence $\frac{\Delta A(x)}{dx} \simeq f(x)$ and we conclude that $A'(x) = f(x)$.

This is a classical way to show the relation between antiderivative and area but in a classical setting, either it is given merely as a metaphor or it is given with the full limit definition and again, the advantages of the metaphor are lost. Here, the fact that an ultrathin slice is ultraclose to the area of an ultrathin rectangle is felt as obvious. What is used as a metaphor becomes formally correct. As usual, at high school, we assume that this area exists. The importance of Hrbacek's new approach is crucial here. Because context levels depends on x and there are always ultrasmall numbers relative to any context level, the definition of the derivative is the same, whether a point is of the coarsest level or of any level i.e., the "proof" above is valid for any point in an interval.

For the integral of a continuous function from a to b, a context level is given by f, a and b. Let $dx = (b - a)/N$ for ultralarge N and $x_i = a + i \cdot dx$. The integral is defined as the context neighbor of the sum of ultrathin slices:

$$\underbrace{\int_a^b f(x)\cdot dx}_{\text{at the context level}} \simeq \sum_{i=0}^{N-1} f(x_i)\cdot dx.$$

The method

The method can be summarized in the following manner. The two main techniques of analysis are: If a property of a function f is local (a property of f at a) then this property is studied by observing an ultrasmall neighborhood of $\langle a; f(a)\rangle$, i.e., by using $a + dx$. If the property is global (on an interval I) then this property is studied by dividing the interval into an ultralarge numbers of even parts.

This requires first to identify a context level. This is given by the list of parameters. Then it is possible to use ultrasmall or ultralarge numbers with respect to that level. We keep the part of the result which belongs to the initial level provided it does not depend on the choice of the ultrasmall or ultralarge number which was used.

Some consequences are:

- The writing
$$f'(x) = \frac{df(x)}{dx}$$
which is often found in mathematics books, in this context, is given the meaning it seems to have: it is a quotient and the different parts can therefore be separated.

- The abstract writing $\int_a^b f(x)\ dx$ can be rewritten with a multiplication symbol
$$\int_a^b f(x)\cdot dx$$
because it is a sum (up to an ultrasmall) of products of the form $f(x)\cdot dx$ where dx is ultrasmall with respect to a level at which a, b and all parameters needed to define f appear.

- In differential equations, the manipulations are algebraic on real values and it is not necessary to introduce abstract forms.

In our view, the main improvement over Robinsonian nonstandard analysis (even the "diet" versions by Keisler [Keisler, 2000] and Stroyan [Stroyan, 1997]) is that we stay within the real numbers. The construction is axiomatic, hence abstract for students, but quite straightforward. No complicated machinery is needed. Our syllabus requires that we introduce the study of the real number system and it is quite problematic to say that in order to do so, we use hyperreals which we feel cannot be well understood if one doesn't understand well the real numbers first. The main improvement over Nelsonian nonstandard analysis is that is is easier to describe which statements are acceptable (those that refer to a context level or to no level) and the improvement over both approaches is the simplicity with which the derivative can be defined at all points.

A handbook where the method is used to cover a complete course on one variable analysis is in preparation [Hrbacek et al., 2010b].

Students

Students respond well to analysis with ultrasmall numbers. Symbols have more meaning and because the technical difficulties are not so high, it is possible to keep a good level of rigor. University professors sometimes complain that students are rather weak in algebraic techniques. Here, the method is more algebraic than a classical approach and often reveals these weaknesses in computational technique. This side effect is more an advantage than a drawback: systematically working with algebraic methods improves their skills in opposition to the rather "magical" side of computing with limits when the full force of formalism is out of reach.

We do not consider that students should never study the classical ε-δ approach. It remains a fundamental tool and appears in almost all mathematics courses. Having different descriptions of the same objects helps understand them better, and it seems preferable to start with the simpler one, as its teaching can also be a way to introduce the necessity of rigor which is the essence of mathematics. The transition from the definition of limit with ultrasmall numbers and the classical definition of limit can be done rigorously by showing their equivalence or more intuitively by observing that they yield the same concepts. This last method is what is usually assumed for almost all mathematical concepts seen at high school and redefined at university. For most students it means encountering mathematical rigor for the first time, which makes the early days of science studies hard. Our former students tell us that for them, it was simply another form of rigor, but nothing fundamentally different.

The use of ultrasmall quantities should not be seen as an attempt to avoid classical mathematics or as a criticism of some forms of teaching. We believe it is an efficient way to introduce analysis and make students understand the concepts. At a higher level, it also provides new insights. We started hesitatingly, but our students convinced us.

Appendix D

Real numbers in the classroom

This article was published in the Mathematical Teaching and Research Journal, August 2010. Its original title was "Proofs which are not proofs" and issued from considerations on many misundestandings about what real numbers *really* are. It is referred to as [O'Donovan, 2010] in the bibliography.

Abstract: *When introducing concepts of analysis some "proofs" are sometimes presented to show in a more or less intuitive way that series do converge, or that* $0.999... = 1$ *prior to defining real numbers. We analyse some of these so-called proofs and show why in fact they beg the question and conclude that it is not possible, for instance, to* prove *that* $0.999... = 1$ *but that it is a reasonable* definition*.*

Consider the well known series

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$$

The dots at the end are left undefined but carry some flavor of "etc." There are graphical proofs which show that 2 is a least upper bound for such a never-ending sum. Yet these proofs do not *prove* that this series is *equal* to this value.

In mathematics when we ignore the value or the existence of something, it is customary to call it x, manipulate such a symbol and if a numeric value is found to be equal to x then we say that we have found a solution and that x exists. Therefore we write

$$\begin{aligned} x &= 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots \\ &= 1 + \frac{1}{2} \cdot (1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots) \\ &= 1 + \frac{1}{2} \cdot x \end{aligned}$$

and conclude that $x = 2$.

The first question that should arise is about the meaning of the dots as never-ending addition and yet writing down a closing parenthesis after this non-ending sequence of additions.

$$1 + \frac{1}{2} \cdot (1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \underbrace{\ldots}_{?} \underbrace{)}_{?}$$

Consider another application of the same method:

$$\begin{aligned} u &= 1 + 2 + 4 + 8 + \ldots \\ &= 1 + 2 \cdot (1 + 2 + 4 + \ldots) \\ &= 1 + 2 \cdot u \end{aligned}$$

and conclude that $u = -1$. Isn't it appealing that a sum of positive numbers can be negative?

Another classical example used to find the fraction corresponding to a result with repeated sequence of digits:

$$\begin{aligned} v &= 0.3333\ldots \\ 10 \cdot v &= 3.3333\ldots \\ & \quad \text{subtract first line from second} \\ 9 \cdot v &= 3 \end{aligned}$$

and conclude that $v = \frac{3}{9} = \frac{1}{3}$. We also have:

$$\begin{aligned} w &= 0.99999\ldots \\ 10 \cdot w &= 9.9999\ldots \\ \text{hence} & \\ 9 \cdot w &= 9 \end{aligned}$$

and conclude that $w = 1$.

But again, consider another absurd situation:

$$\begin{aligned} z &= 9 + 9 \cdot 10 + 9 \cdot 10^2 + 9 \cdot 10^3 \ldots \\ 10 \cdot z &= 9 \cdot 10 + 9 \cdot 10^2 + 9 \cdot 10^3 \ldots \\ \text{hence} & \\ -9 \cdot z &= 9 \end{aligned}$$

and conclude that $z = -1$.

Of course, we also have

$$\begin{aligned} z &= 9 + 9 \cdot 10 + 9 \cdot 10^2 + 9 \cdot 10^3 \ldots \\ &= 9 + 10 \cdot (9 + 9 \cdot 10 + 9 \cdot 10^2 + 9 \cdot 10^3 \ldots \\ &= 9 + 10 \cdot z \end{aligned}$$

and conclude that $z = -1$. (Two different methods which yield the same result: this must be true!)

Maybe we shouldn't forget that addition starts to the right hand side and that the absence of a right hand side beginning might be worth considering as a reason for all these absurd results. Sometimes the method works and sometimes it doesn't without any obvious criterion if one does not have the tools of the limits. Only intimidation can tell which arguments are correct and which are not.

An interesting aspect of all this is that it makes us realise that assuming a number to exist and finding a numerical value to assign to this number does *not* prove that this number exists – even if we frequently do so in class.

Some say that anything goes to find a solution provided we can check that it is the solution. Sadly this fails in most cases shown here. For x and u above, it is not clear at all how one can *check* the result. For v, the division of 1 by 3 does actually produce the required string of digits, but for w the division of 9 by 9 yields 1 and not the required string of nines. Similarly for z. Finding and checking a correct solution does not prove that the proof was correct as False $\Rightarrow$ False as well as False $\Rightarrow$ True.

These methods need that we first show that indeed the objects under consideration do exist. Showing that the series for x is bounded above by 2 and eventually exceeds any lower value does not prove it exists. After all, it never *reaches* this final value. Thus the explanation about convergence cannot be avoided if one wants to explain what the writing of 0.333... and 0.999... stand for: limits. And real numbers are where limits live.

0.999... is not an infinite string of similar digits. This would induce very difficult questions about what kind of infinity is used. Nonstandard analysis shows that there is a universe in which there can be different infinite lengths to such strings and that if one "stops" at one such infinity, then 0.999...[*stop*] $\neq 1$, whatever size of infinity is involved. Hence even in nonstandard analysis, it is not a good idea to try to say how many nines there are. It doesn't matter how many there are because it is defined as a limit, and limits do not in fact use infinity (at least in the classical ε-δ method).

These examples may be used to initiate discussions on what real number are but cannot *prove* anything except maybe show the necessity of clear definitions. $0.999\ldots$ is 1 by definition of what the dots stand for if we say that they stand for limits. If one says that it means that we go on and on and on, then $0.999\ldots$ simply gets closer and closer and closer...

As a conclusive remark, we would like to indicate another interpretation of what 0.999... could stand for – an interpretation we have nowhere found in the literature even though we were once told that Lebesgue shared this view. It is that 0.999... is a sort of mathematical spelling mistake. Any repeating sequence of digits is a rational number so 0.999... is a rational number yet no division will ever produce $0.999\ldots$ (or any non-ending sequence of nines) hence 0.999... is not a rational number. Therefore 0.999... is not the writing of a number. For many students this would probably be more comfortable. We would also have uniqueness of decimal representations. On the other hand, we could also start by this observation then comment that mathematicians don't like the idea of a sequence of digits not representing a number, hence it is *defined* to be a number. Which is to say that it is defined to be a limit — which is what we do when we write the dots then close the brackets.

Bibliography

[Andreev and Gordon, 2006] Andreev, P. and Gordon, E. I. (2006). A theory of hyperfinite sets. *Annals of Pure and Applied Logic*, 143:3–19.

[Benešová et al., 1989] Benešová, M., Vojtášková, B., and Mlček, J., editors (1989). *Proceedings of the First Symposium: Mathematics in the Alternative Set Theory.*

[Benninghofen, 1984] Benninghofen, B. (1984). *Models and Sets*, volume 1103, chapter Superinfinitesimals and the calculus, pages 9–52. Springer-Verlag.

[Benninghofen and Richter, 1987] Benninghofen, B. and Richter, M. M. (1987). *A general theory of superinfinitesimals.* Fund. Math.

[Davis, 1977] Davis, M. (1977). *Applied Nonstandard Analysis.* John Wiley & Sons.

[Diener and Diener, 1995] Diener, F. and Diener, M., editors (1995). *Nonstandard Analysis in Practice.* Springer-Verlag.

[Gödel, 1992] Gödel, K. (1992). *On Formally undecidable Propositions of Principia Mathematica and related systems (1930).* Dover Publications.

[Gordon, 1997] Gordon, E. I. (1997). Nonstandard methods in commutative harmonic analysis. *American Mathematical Society.*

[Hrbacek, 1978] Hrbacek, K. (1978). Axiomatic foundations for nonstandard analysis. *Fundamenta Mathematicae*, (98):1–19.

[Hrbacek, 1979] Hrbacek, K. (1979). Nonstandard set theory. *American Mathematical Monthly*, 86(8):1–19.

[Hrbacek, 2004] Hrbacek, K. (2004). Internally iterated ultrapowers, in *Nonstandard Models of Arithmetic and Set Theory. Contemporary Mathematics*, 361:87–120.

[Hrbacek, 2007] Hrbacek, K. (2007). Stratified analysis? In Neves, V. and van den Berg, I., editors, *The Strength of Nonstandard Analysis*, pages 47–63. Springer.

[Hrbacek, 2009] Hrbacek, K. (2009). Relative set theory: Internal view. *Journal of Logic and Analysis*, 1.

[Hrbacek, 2010] Hrbacek, K. (2010). Relative set theory: Some external issues. *Journal of Logic and Analysis.*

[Hrbacek and Jech, 1999] Hrbacek, K. and Jech, T. (1999). *Introduction to Set Theory.* CRC Press.

[Hrbacek et al., 2010a] Hrbacek, K., Lessmann, O., and O'Donovan, R. (2010a). Analysis with ultrasmall numbers. *American Mathematical Monthly*, 117(9).

[Hrbacek et al., 2010b] Hrbacek, K., Lessmann, O., and O'Donovan, R. (2010b). *Analysis With Ultrasmall Numbers.* preprint.

[Kanovei and Reeken, 2004] Kanovei, V. and Reeken, R. (2004). *Nonstandard Analysis, Axiomatically.* Springer-Verlag.

[Keisler, 2000] Keisler, H. J. (2000). *Elementary Calculus.* University of Wisconsin. www.math.wisc.edu/~keisler/calc.html.

[Koudjeti and van den Berg, 1995] Koudjeti, F. and van den Berg, I. (1995). *Nonstandard Analysis in Practice*, chapter 7: Neutrices, external numbers and external calculus, pages 145–170. Universitext. Springer.

[Krivine, 1972] Krivine, J.-L. (1972). *Théorie axiomatique des ensembles.* Presses Universitaires de France.

[Nelson, 1977] Nelson, E. (1977). Internal set theory. *Bulletin of the American Mathematical Society.*

[Nelson, 2006] Nelson, E. (2006). Warning signs of a possible collapse of contemporariy mathematics. Talk in San Marino.

[O'Donovan, 2007] O'Donovan, R. (2007). Pre-university analysis. In Neves, V. and Van den Berg, I., editors, *The Strength of Nonstandard Analysis.* Springer.

[O'Donovan, 2009] O'Donovan, R. (2009). Teaching analysis with ultrasmall numbers. *Mathematcis Teaching-Research Journal*, 3(3):1–22.
http://www.hostos.cuny.edu/departments/math/mtrj/archives
/volume3/issue3/Donovan.pdf.

[O'Donovan, 2010] O'Donovan, R. (2010). Proofs which are not proofs. *Mathematcis Teaching-Research Journal.*
http://www.hostos.cuny.edu/departments/math/mtrj/archives
/volume3/issue3/Donovan.pdf.

[O'Donovan and Kimber, 2006] O'Donovan, R. and Kimber, J. (2006). Nonstandard analysis at pre-university level. In N.J. Cutland, M. Di Nasso, D. R., editor, *Nonstandard Methods and Applications in Mathematics*, Lecture Notes in Logic 25. Association for Symbolic Logic.

[Péraire, 1992] Péraire, Y. (1992). Théorie relative des ensembles internes. *Osaka Journal of Mathematics*, 29:267–297.

[Péraire, 1996] Péraire, Y. (1996). Formules absolues dans la théorie relative des ensembles internes. *Rivista di matematica pura ed applicata*, 19:27–55.

[Robinson, 1961] Robinson, A. (1961). *Non-standard Analysis.* Princeton University Press.

[Russel, 1903] Russel, B. (1903). *The Principles of Mathematics.* Allen and Unwin.

[Stroyan, 1997] Stroyan, K. D. (1997). *Mathematical Background, Foundations of Infinitesimal Calculus.* Academic Press.
www.math.uiowa.edu/%7Estroyan/backgndctlc.html.

[Vopěnka, 1979] Vopěnka, P. (1979). *Mathematics in the Alternative Set Theory.* Teubner-Texte zur Mathematik.

[Vopěnka and Hájek, 1972] Vopěnka, P. and Hájek, P. (1972). *The Theory of Semisets.* Academia.

[Wildberger, 1996] Wildberger, N. (1996). Should we believe?
http://web.maths.unsw.edu.au/~norman/Views.htm.

[Zadeh, 1965] Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8:338–353.