ABOUT THE COEXISTENCE OF “CLASSICAL SETS” WITH “NON-CLASSICAL” ONES : A SURVEY

Abstract. This is a survey of some possible extensions of ZF to a larger universe, closer to the “naive set theory” (the universes discussed here concern, roughly speaking: stratified sets, partial sets, positive sets, paradoxical sets and double sets).

Mathematics Subject Classification: 03B50, 03B30, 03B53, 03E35, 03E70.
Keywords: set theory, consistency, paraconsistency, extensionality, choice.

0. Introduction

The aim of this paper is to give a (non-exhaustive) survey of the possibility of extending the “usual” Zermelo-Frankel universe to a larger one, closer to the “naive set theory”. As is well-known, the ZF-universe is satisfying for most mathematicians, but excludes many “intuitive” sets and operations, as, e.g. the Russell set, complementation, “filters” of type \( \{ x \mid a \in x \} \), etc. This gives the impression to more “philosophically oriented” people that ZF in a sense “solves” the paradoxes by just refusing to look at them. Several “alternative” set theories have been proposed, and we will suggest here that one can see them, not as being intended to replace ZF (what most mathematicians would refuse to do), or as being in conflict with ZF, but rather as being “reasonable” extensions of ZF.

As in those theories more “classically paradoxical sets and operations” are available, this should satisfy both “mathematicians” and “philosophers”.

The theories we intend to discuss are:

(1) Quine’s “New Foundations” : NF.
Roland Hinnion

(2) “Partial sets” theory (that can be seen as “Frege in partial logic”).
(3) “Positive comprehension”, in its performant version: $GPK_{\infty}$.
(4) “Paraconsistent Frege”.
(5) Kisielewicz’s “Double extension set theory”.

Roughly speaking these theories (however profoundly different) all have in common difficulties to cumulate the following “qualities” (that $ZF$ cumulates rather “spontaneously”):

a) be “naturally” modelizable (say in some “reasonable” extension of $ZF$)

b) “contain” a transitive class that modelizes $ZF$ (with axiom of choice

   if possible) in a sufficiently “standard” way,

   c) be compatible with the axiom of choice $AC$ ($\equiv$ any set is well-orderable).

We will make more precise each of these points and discuss them for each of the concerned theories. Whenever models are involved, our metatheory will be $ZFC$ ($\equiv ZF + AC$) or an adequate extension of $ZFC$.

The paper is kept as less technical as possible, with references to precise results, proofs and details.

Let us first of all precise the concept of “containing $ZF$ in a standard way”,

for a theory $T$ expressed in the language $L$ of $ZF$ (i.e. with variables for sets

and proper symbols $\in$ and $=$), in classical logic (for the theories (2) (4) (5)

we will need specific adaptations). We will use the notion of “class” in the

usual way, i.e. as “definable collection of sets”; so a class is a meta-object,

which use is intended only to allow traditionnal set-theoretical abbreviations.

When we say that the theory $T$ contains $ZF$ in a standard way, we mean

that the following wishes have been fulfilled: one can define a class $C = \{x \mid \varphi(x, \vec{y})\}$ in $T$ (where $\vec{y}$ is an $n$-tuple $y_1 \ldots y_n$ of parameters), realizing

the following conditions:

(i) $C$ is transitive, i.e. $t \in x \& x \in C \rightarrow t \in C$ (where “$z \in C$” is the abbreviation of “$\varphi(z, \vec{y})$”).

(ii) $C$ provided with $\in$ (more precisely: with the restriction of $\in$ to $C$)

models $ZF$ (if possible also $AC$), i.e. makes any concerned axiom $\sigma$

true when relativized to $C$ (to relativize, bound each quantifier by $C$;

the parameters $\vec{y}$ should also be in $C$).

(iii) any subclass (in the sense of $T$) of a set in $C$ is a set in $C$, i.e. $X \subset a$

   $\& a \in C \rightarrow \exists y \in C \ X = y$. 

(iv) if a set \( f \) (in the universe \( V \) of \( T \)) is a function, \( x \in C \) and \( \forall t \in x \ f(t) \in C \), then \( f \in C \) (and so by (ii), also \( \{ f(t) \mid t \in x \} \in C \)).

Some comments:

1) Condition (iv) extends the “replacement principle” in \( C \) to all functions \( f \) (of \( V \)) of type \( f \subset C \times C \).

2) Conditions (i) (ii) simply express that \( C \) should be a transitive interpretation of \( ZF \).

3) Conditions (iii) (iv) are intended to guarantee “standardness” for a lot of notions (a notion being \textbf{standard} iff it concerns “objects” of \( C \) and has the same meaning in \( C \) and in \( V \)); for example:

- the “powerset” \( \mathcal{P}a := \{ x \mid x \subset a \} \); with condition (iii), we have indeed:
  \[
  \{ x \in C \mid x \subset a \} = \{ x \in V \mid x \subset a \}, \text{ for } a \in C \text{ and } V := \{ y \mid y = y \},
  \]

- “to be countable”, i.e. in bijection with the least infinite Von Neumann ordinal \( \omega \) (use condition (iv)),

- “to be well-founded”,

- “to be a Von Neumann ordinal”,

- etc...

It would indeed be frustrating that (for example) some \( a \in C \) would be considered as countable when seen in \( V \), and uncountable in \( C \)!

1. Quine’s “New Foundations”

The system \( NF \) [20] can be seen as an “adaptation” of Russell’s “type theory” to the first order language (of \( ZF \)) \( \mathcal{L} \) (see [7], [15] for a detailed bibliography).

The idea is to restrict the “naive comprehension principle”:

\[
\exists a \ \forall t \ (t \in a \leftrightarrow \varphi(t, \vec{y}))
\]

to those formulas \( \varphi \) (in \( \mathcal{L} \)) that are “stratified”, i.e. for which one can associate to each variable “\( v \)” (in \( \varphi \)) an integer \( j_v \) (the “type” of \( v \); notice that all occurrences of \( v \) should receive the same type) in such a way that \( j_y = j_x + 1 \) whenever “\( x \in y \)” appears in \( \varphi \) and \( j_z = j_t \) whenever “\( z = t \)” appears in \( \varphi \).
Typically \( x \in x \) is not stratified (so that one avoids the straightforward Russell-paradox argumentation), while (for example) \( \forall t (t \in x \rightarrow t \in y) \) is stratified. Further NF admits also the (classical) axiom of extensionality:

\[
\text{EXT} \equiv (\forall t (t \in x \leftrightarrow t \in y)) \rightarrow x = y.
\]

In 1953, Specker gave his famous proof of \( \neg AC \) in NF [21]. Notice that the well-known equivalent forms of AC (equivalent in ZF) stay equivalent in NF (see [15, Chapter 14]), so that this result is actually a very “hard” one (at last for mathematicians). The positive side of this however is that NF proves an (adequate) “axiom of infinity”; more precisely this axiom does not refer to the existence of the least infinite Von Neumann ordinal \( \omega \), but states that \( V \notin FIN \), where \( V := \{x|x = x\} \) is a set here (because \( x = x \) is stratified) and \( FIN \) is the set of all finite sets:

\[
FIN := \cap\{b|\emptyset \in b \& \forall z \in b \forall t \ z \cup \{t\} \in b\}.
\]

This axiom permits to prove that the set \( \mathbb{N} \) of all Frege-integers (i.e. equipotence classes of finite sets; for example 2 is the set of all true pairs) satisfies the Peano-arithmetic axioms, so is very satisfactory from the mathematical point of view.

The most “negative” aspect of NF however is still the question of its consistency. The problem is related to the (full) axiom EXT, as Jensen [16] showed that “slightly weakening” EXT suffices to get a system, \( NFU \equiv \text{stratified comprehension + extensionality for the non-empty sets, i.e.} \)

\[
\exists z \in x \rightarrow (\forall t (t \in x \leftrightarrow t \in y) \rightarrow x = y)),
\]

that one can modelize and that is even compatible with AC and stronger axioms of infinity, essentially in the same way as ZF [15, in particular Chapter 1].

All this gives the impression that perhaps NFU is the “good” version of “stratified comprehension”, which is the thesis adopted in [15] and reinforced by the easy way in which NFU can “incorporate” ZF. But what can be said about “ZF in NF (the “true” one, not NFU)” ? This is actually realizable (with the “standardness” expectations explained in section 1), at the price of assuming the existence of some large cardinal; more precisely:

**Theorem.** If “\( NF+ \) there exists a strongly inaccessible, strongly cantorian, uncountable cardinal \( \kappa \)” is consistent, then there exists a model \( M \) of NF that contains ZF in a standard way (as defined in section 1).
Before we define the involved notions, let us mention that this results from [11, theorem p.25] and [13, section 3], and that one gets \( ZFC \) instead of \( ZF \) if \( \kappa \) is assumed to be a “well-ordered cardinal”.

Some definitions and comments:

- the class \( C \) that interprets \( ZF \) in \( M \) is even a set in \( M \) (here),
- the notion of “cardinal” should be understood in Frege’s sense (as an “equipotence class”):
  \[
  \text{Card } x := \{y | y \text{ can be put in bijection with } x\}.
  \]
  When \( x \) is a well-orderable set, its cardinal is said to be a “well-ordered cardinal”,
- a cardinal \( \kappa \) is “strongly inacessible” iff
  \[
  \forall y (\text{Card } y < \kappa \to \text{Card } \mathcal{P} y < \kappa) \land \\
  \forall a ((\text{Card } a < \kappa \land \forall b \in a \text{ Card } b < \kappa) \\
  \to \text{Card } \cup a < \kappa);
  \]
  this is the “natural” adaptation to \( NF \) of this familiar notion of \( ZF \) (notice that some authors just say “inaccessible”, e.g. [2]),
- in \( NF \), a set \( x \) is called “strongly cantorian” when the collection \( \{(t, \{t\}) | t \in x\} \) (that is not defined a priori by a stratified formula) is a set. Strongly cantorian sets have a more “normal” behaviour than others (see e.g. [15]). By extension a cardinal \( \text{Card } x \) is called “strongly cantorian” when \( x \) is a “strongly cantorian set”,
- naturally, so far nobody really knows about the compatibility of \( NF \) with the type of “large cardinal” involved in the Theorem, but the risk of inconsistency exists also for “\( ZF + \text{ such a } \kappa \)”; it should also be mentioned that traditionnaly “strong inaccessibility” is a rather “weak” assumption in the hierarchy of “large cardinals”,
- one can get “\( ZF \) in \( NF \)” at a “lower price”, but then not in a “standard way” (as defined in section 1); in [13, “characterization theorem”, p. 524] for example, it is shown that the consistency of “\( NF + \text{ Rosser’s axiom (stating that } \mathcal{N} \text{ is a strongly cantorian set, with } \mathcal{N} \text{ the set of all cardinals of finite sets) } + \text{Con } ZF \) (the arithmetical statement that translates “\( ZF \text{ is consistent}”) suffices to guarantee the consistency of “\( NF + \text{ there exists a transitive set } M \text{ that modelizes } ZF \)”); that model \( M \) (munished with \( \in \) ) realizes our wishes (i) (ii) (see section 1), but not (iii), (iv), because \( M \) itself is countable (in \( NF \) !
2. Partial sets

This kind of concept was introduced by Gilmore (see the pioneer paper [10]), who proposed two versions (\(PST\) and \(PST^+\)) of a theory of “partial sets”, both in a language with variables (for the “partial sets”), primitive symbols \(\in, \notin, =\) and an abstraction operator. Those theories negate the corresponding axiom of extensionality [10], [12], precisely because of the presence of the abstraction operator. So it was very natural to hope that some adaptations of these theories to a first-order language without abstractor could be compatible with extensionality. Actually these adaptations can also be seen as particular versions of “Frege” (i.e. full comprehension) with extensionality, in some “partial logic”; this presentation has been introduced in [12] and further studied in detail in [14], where models for some versions, but not all, are given; in particular the question of a model is still open for the “most natural” version, called \(F2\) in [14] (this one is very close to \(PST^+\)).

To give some intuition about all this, we only describe here what a model of “\(F2\) in partial logic” should be. Actually one wants a structure of type

\[ M = (A, \in^+_M, \in^-_M, =^+_M, =^-_M), \]

where \(A\) is a set and \(\in^+_M, \in^-_M, =^+_M, =^-_M\) are binary relations on \(A\) (\(\in^+_M\) corresponds to “\(\in\)” in [10], \(\in^-_M\) to “\(\notin\)”, \(=^+_M\) to “\(=\)” and \(=^-_M\) to “\(\neq\)”). Further \(=^+_M\) should be the “true” equality on \(A\), and the following conditions should be fulfilled:

- **Extensionality**:

  \[ \forall t ((t \in^+_M x \leftrightarrow t \in^+_M y) \& (t \in^-_M x \leftrightarrow t \in^-_M y)) \leftrightarrow x =^+_M y \]

  and

  \[ \exists t ((t \in^+_M x \& t \in^-_M y) \lor (t \in^-_M x \& t \in^+_M y)) \leftrightarrow x =^-_M y \]

- **Comprehension**:

  \[ \exists a \forall t ((t \in^+ a \leftrightarrow \varphi(t, \vec{y})) \& (t \in^- a \leftrightarrow \varphi(t, \vec{y})) \]

should be realized in \(M\), for any positive formula (i.e. build up from the atomic formulas of type \(x \in^+ y, x \in^- y, x =^+ y, x =^- y\), only using \(\forall, \exists, \&\), \(\lor\); \(\varphi\) is the “dual” of \(\varphi\) obtained via the rules:

- \(\overline{x \in^+ y} \) is \(x \in^- y\), \(\overline{\forall x \psi} \) is \(\exists x \overline{\psi}\), etc...
3. Positive comprehension

The unsolved questions about “partial sets” naturally suggested (apparently) simpler problems: what about “positive comprehension”, but in the language $\mathcal{L}$ of ZF and with classical logic? More precisely, what can be said about the theory $T$ whose axioms are:

- Extensionality:
  \[ \forall t (t \in x \leftrightarrow t \in y) \rightarrow x = y \]

- Positive comprehension:
  \[ \exists a \ \forall t (t \in a \leftrightarrow \varphi(t, \vec{y})) \]
  for any positive formula $\varphi$ in $\mathcal{L}$.

- Empty set axiom (intended to avoid the trivial model with one element):
  \[ \exists x \ \forall t \neg t \in x. \]

Now, Gilmore’s technique (to modelize PST) can be easily adapted to furnish a model of $T$, but alas without extensionality [12, Remarque p.310]. The solution for the full $T$ came only later [8], [13], thanks to improvements of Malitz’s pioneer work [19], that uses “topological” techniques. The models even allow “generalized positive” comprehension (i.e. extensions of $T$) and have for now been studied in detail by several authors.

We will restrict ourselves here to those results that are linked to the aim of this paper (see section 1).

Call $GPF$ (“generalized positive formulas”) the formulas obtained from the atomic ones (of type $x \in y$, $x = y$), only using $\forall$, $\exists$, $\&, \lor$, $\forall x \in y$ (“universal bounded quantification”), and quantification of type $\forall x$ such that $\theta(x)$, where $\theta$ is an arbitrary formula (not necessarily positive!) with exactly one free variable (namely ‘$x$’). The system $GPK^+$ has as axioms the empty set axiom and extensionality (of $T$), and the following comprehension scheme:

\[ \exists a \ \forall t (t \in a \leftrightarrow \varphi(t, \vec{y})), \]
for any formula $\varphi$ of type

$$\forall z (\theta(z, \vec{y}) \rightarrow \psi(t, z, \vec{y})),$$

where $\psi$ is $GPF$ and $\theta$ does not contain the variable “$t$”. In particular one has this comprehension when $\varphi$ is $GPF$ (just take $\psi \equiv \varphi$ and $\theta \equiv z = z$). So this is really “generalized” positive comprehension, as $\varphi$ may contain many “negative” parts (compare with $T$). The system $GPK^+$ can be enriched via an adequate infinity axiom, namely “$\exists \omega$” that states the existence of an infinite Von Neumann ordinal; the resulting system is denoted $GPK^+_\infty$.

This appears to be a kind of “optimal natural” version of “positive comprehension” and has been formulated and studied by Esser [4]. Actually $GPK^+$ (without the infinity axiom) has models in $ZFC$ [8], [13], which satisfy even “spontaneously” $AC$ [9]. But, if one wants models of $GPK^+$ that “contain $ZF$ in a standard way”, one has to modelize $GPK^+_{\infty}$, and this theory disproves $AC$ (see [6]); this important result should be put in parallel with the analogue situation for $NF$ [21]); further the “price” for a model of $GPK^+_{\infty}$ is a “weakly compact, uncountable cardinal” (in the hierarchy of “large cardinals” the existence of that type of cardinal $\kappa$ is still considered as a “rather low” assumption; the model that is constructed contains actually all the well-founded sets of rank $< \kappa$, and these modelize $ZFC$ in the desired “standard way”; for details, see [8]). The exact interpretative power of $GPK^+_{\infty}$ has been determined by Esser [4], [5]; it corresponds to “Kelley-Morse + the class of all ordinals behaves like a weakly compact cardinal”.

### 4. Paradoxical sets

The “topological” models for $GPK^+$ evoked in section 3 inspired adaptations to modelize “paraconsistent Frege”. Without extensionality Gilmore’s technique suffices (the corresponding consistency result was independently noticed by Crabbé [3]), but to overcome the lack of extensionality the “topological” techniques were necessary. The system “paraconsistent Frege” that we want to discuss here can be described in several ways. In [14] it is “$F2$ in $Pd$-Logic”. This “$Pd$-Logic” (“$Pd$” for “paradoxical”) corresponds actually exactly to the logic $CLuNs^1$, familiar (at least) to the Flemish-Polish “paraconistency tradition” (see e.g. [1]). In order to describe the system as simply as possible, we adopt the presentation of section 2 (the situation here is a kind of “dual” of the situation there).

---

1 I wish to thank Prof. D. Batens who immediately recognized this correspondence.
So a model of “$F_2$ in Pd-Logic” should simply be like a model of “$F_2$ in partial logic” (see section 2), except that the “partial sets principle” should be replaced by the following “paradoxical sets principle”: $\in^+_M \cup \in^-_M$ and $=^+_M \cup =^-_M$ are both $A \times A$ ($A$ is the universe of the model). To avoid the trivial model with one element, one has also to add (here) a condition like: $\exists x \forall t \neg t \in^+_M x$.

For the reader familiar with $\mathcal{CLuNs}$, the system can be described like this: the logic is $\mathcal{CLuNs}$, with classical negation $\neg$ and the non-classical negation $\sim$; further the non-logical axioms express the following version of “Frege with extensionality”:

- **Extensionality**: 
  $$\forall t (t \in x \leftrightarrow t \in y) \leftrightarrow x = y$$
  and
  $$\sim (\forall t (t \in x \leftrightarrow t \in y)) \leftrightarrow \sim (x = y)$$
  (where $\sim$ is classical, i.e. $A \sim B$ is $\neg A \lor B$).
- **Comprehension**: 
  $$\exists a \forall t (t \in a \leftrightarrow \varphi(t, \vec{y})),$$
  for $\varphi$ build up positively (i.e. via $\forall$, $\exists$, $\&$, $\lor$) from basic formulas of type: $x \in y$, $\sim (x \in y)$, $x = y$, $\sim (x = y)$.
- **Empty set axiom**: 
  $$\exists x \forall t \neg t \in x.$$  

The link with the previous description in terms of “a model” is simply that: $x \in y$ corresponds to $x \in^+_M y$, $\sim (x \in y)$ to $x \in^-_M y$, $x = y$ to $x =^+_M y$ and $\sim (x = y)$ to $x =^-_M y$.

Actually $ZFC$ is strong enough to modelize this form of “paradoxical Frege” [14]. The construction uses a “projective limit” with $\omega$ levels; actually, going beyond $\omega$, with the aim of getting also “$ZF$ in Frege” presents several technical difficulties, not yet completely solved. Also the question of $AC$ has not been studied so far.

However, as the situation is (apparently) close (but not identical) to the one of $GPK^-_{\omega}$, we think that it is reasonable to conjecture\(^2\) that, in “ZFC+” there exists an uncountable, weakly compact cardinal” one can construct a

\(^2\)This was written in 2000; underwhile T. Libert axiomatized the adequate version of “paraconsistent Frege” (and called it “HyperFrege”), and O. Esser proved the conjecture.
model of “paraconsistent Frege containing ZFC in a standard way”; as for $GPK^\omega_\infty$, such a model (obtained via projective limits) will probably negate $AC$ (on its whole universe), but accept $AC$ on the $ZF$-part.

**Remark.** The “projective limit” construction, even restricted to $\omega$ levels, is problematic when adapted to the “partial sets” case, putting in evidence the “not really dual” relation between “$Pd$-logic” and “$Pt$-logic” (see [14]).

5. **Kisielewicz’s “Double sets”**

Another alternative approach to set theory is to be found in the “double sets” concept, first presented in a language with class variables [17] and later improved in the sense of a welcome simplification of the axioms as well as of the language (first-order this time) [18], while keeping the desired strenght of the system(s), which lies in its aim to avoid paradoxes in a very original and simple syntactical way and to get $ZF$ in it!

We only describe here some of the basic ideas involved. The underlying logic is classical, the language is first-order (with equality) but has two kinds of membership: “$\epsilon$” and “$\epsilon^\prime$”.

A **regular** set $x$ is a set realizing:

$$\forall t (t \epsilon x \leftrightarrow t \epsilon^\prime x).$$

Notice that this is very different from the “partial sets” and “paradoxical sets” concepts, where a “regular” (or “normal”, or “classical”) set should realize:

$$\forall t (t \epsilon^\prime x \leftrightarrow \neg t \epsilon x);$$

so $\epsilon$ has to do with *complements*, while here $\epsilon^\prime$ has to do with another extension.

The axiomatization of “double set” theories is inspired by the following remark: consider $R := \{x | \neg x \in x\}$ (the “Russell set”): if instead of a comprehension sentence like $\forall t (t \in R \leftrightarrow \neg t \in t)$, one has this: $\forall t (t \epsilon R \leftrightarrow \neg t \epsilon t)$, the paradox (at least in its immediate form) vanishes!

“The questions of the consistency of these theories as well as their behaviour w.r.t. $AC$ are, however, still open.”

---

3 I wish to thank Prof. J. Perzanowski who let us know the existence of this very original “non-classical alternative” for set theory.

4 This was written in 2000. Underwhile (precisely in 2002) Randall Holmes succeeded
References


In proving the inconsistency of these systems, except for the weakest variant, that might be consistent and still strong enough to “contain ZF”: this is the subject of papers by R. Holmes that should appear in a close future.


Roland Hinnion
Université Libre de Bruxelles, CP 211
Bd. du Triomphe
B-1050 Brussels, Belgium
rhinnion@ulb.ac.be