Paraconsistent set theory by predicating on consistency

Walter Carnielli and Marcelo E. Coniglio
Department of Philosophy
Institute of Philosophy and Human Sciences (IFCH)
and
Centre for Logic, Epistemology and the History of Science (CLE)
State University of Campinas (UNICAMP) - Brazil
e-mail: {carniell,coniglio}@cle.unicamp.br

Abstract

This paper intends to contribute to the debate about the uses of paraconsistent reasoning in the foundations of set theory, by means of employing the logics of formal inconsistency (LFIs) and by considering consistent and inconsistent sentences, as well as consistent and inconsistent sets. We establish the basis for new paraconsistent set-theories (such as *ZFmbC* and *ZFCil*) under this perspective and establish their non-triviality, provided that *ZF* is consistent. By recalling how George Cantor himself, in his efforts towards founding set theory more than a century ago, not only used a form of 'inconsistent sets' in his mathematical reasoning, but regarded contradictions as beneficial, we argue that Cantor's handling of inconsistent collections can be related to ours.

1 A new look at antinomic sets

Ever since the discovery of the paradoxes, the history of contemporary set theory has centered around attempts to rescue Cantor's naive theory from triviality, traditionally by placing the blame on the Principle of (unrestricted) Abstraction (or Principle of Comprehension). Unrestricted abstraction (which allows sets to be defined by arbitrary conditions) plus the axiom of extensionality, and plus the laws of the underlying logic where the theory of sets is expressed, leads to a contradiction when a weird collection such as the Russell set or similar constructions are defined. The problem is not the weird collections by themselves – set-theorists are used to strange objects like large cardinals, measurable cardinals and the like, and, in fact, hypothesizing on large cardinals enables us to investigate the capabilities of possible extensions of *ZFC*. The problem is that some weird sets, such as Russell's, entail a contradiction, and in classical logic a contradiction entails everything.

One way of escaping this mathematical Armagedon is to consider weaker forms of separation, by patching in the Principle of Comprehension, but this seems to remedy the problem just temporarily, since (as is well-known) there is no possibility of proving the absolute consistency of set theory. A radical departure from this position, the paraconsistent one, advocates that the underlying paradoxes are not necessarily to be solved, but that they can be made part of the rational arsenal of the working mathematician, as well as of the philosopher interested in foundations of mathematics. This kind of rational accommodation can be achieved if the underlying consequence relation is paraconsistent, instead of classical – in such a case, contradictions remain, so to speak, quarantined, and the triviality disease does not spread all over the universe.

Much has been said and written on paraconsistent set theory, chiefly because of Russell's antinomy and its alleged dangerous consequences. In the usual classical set theory, classical negation plus an unrestricted application of the comprehension (or separation) scheme may cause deductive trivialization of all mathematics, if one takes seriously the dogma that mathematics must be founded on sethood.

Paraconsistent set theory, in a nutshell, is the theoretical move of maintaining weird sets as much as we can – we call them *inconsistent sets* for reasons which will be made clear in the following – and weakening the underlying logic governing sets so as to avoid the disastrous consequences of inconsistent sets. This move is in frank opposition with traditional strategies, which maintain the underlying logic and weaken the Principle of Abstraction. E. Zermelo's proposal in 1908 of an axiomatization of set theory by replacing the Principle of Abstraction with a weaker existence axiom, the axiom of separation (*Aussonderung*), gave birth to *ZF* set theory and type theory, but represented a radical departure from Cantor's intuitions. Indeed, Cantor regarded the antinomies as positive results which fully complemented the advance of his investigation. Not only Cantor, but Hadamard and Jourdain also held similar ideas. Somehow, Cantor had already conceived a form of inconsistent sets (cf. [18]):

> "A collection [Vielheit] can be so constituted that the assumption of a 'unification' of all its elements into a whole leads to a contradiction, so that it is impossible to conceive of the collection as a unity, as a 'completed object'. Such collections I call absolute infinite or inconsistent collections."
>
> (Cantor, letter to Dedekind, 1899)

What may be surprising in many ways is that Cantor indeed reasoned with such 'inconsistent sets', and regarded contradictions as beneficial, incorporating them into his philosophy of the infinite. By taking such inconsistent collections into account, Cantor concluded that every consistent set had to be power-equivalent to a definite Aleph. As J. W. Dauben, in the most comprehensive biography of Cantor to date ([18], p. 244), puts it:

> "In one theorem, based on the inconsistency of the system of all transfinite numbers, Cantor had succeeded in resolving several perplexing and long-standing problems of set theory."

Can Cantor's approach be related to paraconsistent set theory? And if so, to which one? Paraconsistent set theory has been around since at least fifty years proposed by N. da Costa, who inaugurated the topic in his [12]. In the 1980s, A. Arruda and N. da Costa studied the systems NF_n, a kind of weakened version of Quine's *NF*; several results had been published by them between 1964 and 1975. The first proponents of the paraconsistent approach to set theory have found some puzzling problems, however. Arruda showed in 1980 (cf. [1] and [13]) that, without a restriction to the comprehension axiom, all the set-theoretical systems in the hierarchy NF_n are trivial. She also proposed alternative systems which apparently fix the problem, conjecturing their equiconsistency with *NF*; additional results were obtained by Arruda and Batens in 1982 (cf. [2]). In addition, in 1986 da Costa himself (in [13]) proposed an improved version of his paraconsistent set theory by introducing a hierarchy called CHU_n which is obtained from Church's set theory *CHU* (cf. [11]) by changing the underlying first-order logic by $C_n^=$ (da Costa's first-order logic C_n^* plus equality), and by adding two versions of the axioms involving negation, one using strong (classical) negation and another using weak (paraconsistent) negation (cf. [14]). He proves that *CHU* is consistent if and only if each CHU_n is non-trivial. An interesting feature of CHU_n is that both the universal set and the Russell set are allowed. More details of this proposal and other features and applications of the diverse da Costa's paraconsistent set theories can be found in [15].

Another approach to paraconsistent set theory was proposed in [31] by R. Routley (who changed in 1983 his name to R. Sylvan), by introducing a system for set theory based on relevant logic in which the full comprehension principle holds. In 1989, R. Brady (see [4]) proved that there is a non-trivial model for Routley's paraconsistent set theory (including some contradictions such as Russell's), and that therefore is definitely non-trivial. Brady's results represent an important step in the development of Routley's set theory, not only by producing a model of dialetheic set theory which shows its non-triviality, but because Routley's set theory includes the metaphysical perspective that the universe is comprised of, aside from real objects, nonexistent, contradictory or even absurd objects that one can imagine or think about. Seen from this perspective, paraconsistent set theory is instrumental for the whole of rationality, far beyond the foundations of mathematics.

In a more recent investigation (cf. [23]), the question of the existence of natural models for a paraconsistent version of naive set theory is reconsidered. There it is proved that allowing the equality relation in formulae defining sets (within an extensional universe) compels the use of non-monotonic operators, and that fixed points can be used to obtain certain kind of models. Already in 1994, a naive set theory with paraconsistent basis was investigated in [20]; more recently, in 2010, Z. Weber (cf. [34]), by revisiting Sylvan's ideas, introduced an axiomatic system for naive set theory (i.e., with a full comprehension principle) in a paraconsistent setting based on relevant logic. In [33] a non-trivial (but inconsistent) set theory based on unrestricted comprehension was also proposed, formalized by means of an adaptive logic.

The main difficulty with all such views on set theory from a paraconsistent viewpoint is that they seem to be all quite unconnected and apparently *ad hoc.*

However, we show that in essence they are really similar. In this paper we take up the discussion, showing that a wide approach to antinomic set theory can be given by means of the paraconsistent LFIs (Logics of Formal Inconsistency as found in [9]).

We propose here a new axiomatic paraconsistent set theory based on the first-order version of some (paraconsistent) LFIs, by admitting that sets, as well as sentences, can be either consistent or inconsistent. A salient feature of such a paraconsistent set theory, inherited from LFIs, is that only consistent and contradictory objects will explode into triviality. Moreover, if we declare that all sets and sentences are consistent, we immediately obtain traditional *ZF* set theory, and nothing new.

This permits one to recover in principle, within our paraconsistent set theory, all the concepts and definitions concerned with the predicativist program for the foundations of mathematics, in particular the variant proposed by A. Avron in [3]. One could also think, although this direction has not been yet explored, about defining some new constructs by referring to constructs which were introduced by previous definitions by allowing some degrees of circularity (impredicativity). This approach would introduce a form of controlled predicativism, that might be interesting by itself. The inclusion of a consistency predicate for sets constitutes one of the main differences with previous approaches to paraconsistent set theory such as da Costa's proposals.

Inconsistent objects such as the ones Cantor was fond of, on the other hand, can be used to entail beneficial consequences, as Cantor himself pointed out. Our inconsistent sets are, in their way, akin to the inconsistent sets conceived by Cantor, in the sense that the essentially unfinished sets devised by Cantor can be seen as a particular kind of the inconsistent sets our theory supports.

2 (First-order) Logics of Formal Inconsistency

Logics of Formal Inconsistency (LFIs), introduced in [10] (see also [9]) are paraconsistent logics (that is, having a non-explosive negation $\neg$) having (primitive or defined) connectives $\circ$ and $\bullet$ describing the notion of consistency and inconsistency, respectively. The main point behind the LFIs is that contradictions involving $\neg$ do not necessarily trivialize (in deductive terms) a system or a theory, but a contradiction on a realm that is taken to be consistent does entail any conclusion. That is, a theory containing a contradictory pair of sentences $\alpha, \neg\alpha$ is not necessarily trivial in deductive terms, but any theory containing $\alpha, \neg\alpha, \circ\alpha$ is always deductively trivial. Contradiction and inconsistency are therefore not coincident notions in LFIs, and neither are non-contradiction and consistency.

The logic *mbC*, one of the basic LFIs, fulfills this rationale by adjoining to positive classical logic axiom $(ax10)$ concerning negation and axiom $(bc1)$ concerning consistency (the latter regarded as a primitive unary connective denoted by $\circ$):

$(ax1)$ $\qquad \varphi \to (\psi \to \varphi)$

($ax2$) $(\varphi \to \psi) \to ((\varphi \to (\psi \to \gamma)) \to (\varphi \to \gamma))$

($ax3$) $\varphi \to (\psi \to (\varphi \wedge \psi))$

($ax4$) $(\varphi \wedge \psi) \to \varphi$

($ax5$) $(\varphi \wedge \psi) \to \psi$

($ax6$) $\varphi \to (\varphi \vee \psi)$

($ax7$) $\psi \to (\varphi \vee \psi)$

($ax8$) $(\varphi \to \gamma) \to ((\psi \to \gamma) \to ((\varphi \vee \psi) \to \gamma))$

($ax9$) $\varphi \vee (\varphi \to \psi)$

($ax10$) $\varphi \vee \neg\varphi$

($bc1$) $\circ\varphi \to (\varphi \to (\neg\varphi \to \psi))$

Inference Rule:

Modus ponens

The logic mbC can be extended to first-order languages by adding the following axioms and inference rules, obtaning the logic $QmbC$ (cf. [28]):

($ax11$) $\varphi[x/t] \to \exists x\varphi$, if t is a term free for x in φ

($ax12$) $\forall x\varphi \to \varphi[x/t]$, if t is a term free for x in φ

($ax13$) If φ is a variant of ψ, then $\varphi \to \psi$ is an axiom[1]

Inference Rules:

(I-$\forall$) $\varphi \to \psi$ / $\varphi \to \forall x\psi$, if x isn't free in φ

(I-$\exists$) $\varphi \to \psi$ / $\exists x\varphi \to \psi$, if x isn't free in ψ.

Axiom ($ax13$) seems redundant, since the usual equivalences between $Gx\varphi$ and $Gy\varphi[x/y]$ (where $\varphi[x/y]$ denotes the formula obtained from φ by substituting all the free occurrences of x by y, provided that y is free for x in φ) and between $Gx\varphi$ and φ (if x is not free in φ) can be proved in $QmbC$ as usual, for $G = \forall$ or

[1] We say that φ is a variant of ψ if φ is obtained from ψ by means of addition or deletion of void quantifiers, or by the renaming of bound variables (keeping the same free variables in the same places).

$G = \exists$. The problem arises because $\circ$ and $\neg$ do not preserve logical equivalences (cf. [9]), and thus the inclusion of $(ax13)$ is necessary in order to simplify the calculus.

The basic feature of *mbC* is that there exists α and β such that

$$\alpha, \neg\alpha \nvdash \beta$$

but, for every α and β, it holds that

$$\alpha, \neg\alpha, \circ\alpha \vdash \beta.$$

In *mbC* a bottom formula can be defined as $\bot_\alpha = \alpha \wedge \neg\alpha \wedge \circ\alpha$. On this basis, the strong (classical) negation is definable in *mbC* as $\sim_\alpha \psi =_{def} (\psi \to \bot_\alpha)$. Note that $\bot_\alpha$ and $\bot_\beta$ are equivalent, as well as $\sim_\alpha \psi$ and $\sim_\beta \psi$, for every α, β and ψ. In order to avoid confusions, from now on we will consider $\bot = \bot_p$ and $\sim\psi = \sim_p \psi$ for a fixed propositional variable p. In the case of first-order languages, p will be replaced by the formula $\forall x(x = x)$, for a fixed variable x.

It is also useful to recall the following properties of *mbC*:

Proposition 2.1 *In mbC it holds that: if* $\Gamma, \varphi \vdash \bot$ *then* $\Gamma \vdash \neg\varphi$*. In particular,* $\sim\varphi \vdash \neg\varphi$ *and* $\circ\varphi \vdash \neg(\varphi \wedge \neg\varphi)$*, but the converses do not hold in general. Additionally,* $(\varphi \wedge \neg\varphi) \vdash \neg\circ\varphi$ *holds in mbC, but the converse does not hold in general. On the other hand,* $\varphi \to \psi \vdash \neg\psi \to \neg\varphi$ *does not hold in general.*

Concerning the first-order calculus, it is worth noting that the quantifiers $\forall$ and $\exists$ are interdefinable in *QmbC* as usual, but using the strong negation $\sim$ instead of the paraconsistent negation $\neg$. Thus, $\exists x\varphi$ is equivalent to $\sim\forall x \sim\varphi$ in *QmbC*, but it is not the case that $\exists x\varphi$ and $\neg\forall x\neg\varphi$ are equivalent in *QmbC*.[2]

3 Predicating on consistency

The basic system of paraconsistent set theory proposed here is called *ZFmbC*, and consists of the first-order version *QmbC* of *mbC* as introduced in the previous section, over a first-order language containing two binary predicates "=" (for equality) and "$\in$" (for membership), and a unary predicate C (for consistency of sets), together with the following axioms, organized in five groups:

1) The Leibniz axiom for equality:

$(Leib)$ $\quad (x = y) \to (\varphi \to \varphi[x/y])$

2) All the axioms of *ZF*, other than the Regularity Axiom (the notation for the axioms of this group will be explained below):

$(A1)$ $\quad \forall z(z \in x \leftrightarrow z \in y) \to (x = y)$ (extensionality)

[2]It is worth noting that this phenomenon occurs with first-order paraconsistent logics in general.

$(A2)$ $\exists y \forall x(x \in y \leftrightarrow \forall z(z \in x \to z \in a))$ (power set)

$(A3)$ $\exists y \forall x(x \in y \leftrightarrow \exists z((x \in z) \wedge (z \in a))$ (union)

$(A4)$ $\exists w((\emptyset^* \in w) \wedge (\forall x)(x \in w \to x \cup \{x\} \in w))$ (infinity)

$(A5)$ $FUN_\psi \to \exists b \forall y(y \in b \leftrightarrow \exists x(x \in a \wedge \psi(x,y)))$ (replacement)

3) The Regularity Axiom:

$(A6)$ $C(x) \to (\exists y(y \in x) \to \exists y(y \in x \wedge {\sim}\exists z(z \in x \wedge z \in y)))$

4) The Unextensionality Axiom:[3]

$(UnExt)$ $(x \neq y) \leftrightarrow \exists z((z \in x) \wedge (z \notin y)) \vee \exists z((z \in y) \wedge (z \notin x))$

5) Axioms for the consistency predicate:

(C_1) $\forall x(C(x) \to \circ(x = x))$

(C_2) $\forall x(\neg\circ(x = x) \to \neg C(x))$

(cp) $\forall x(x \in y \to C(x)) \to C(y)$

Notation used in the *ZF* axioms: In the *axiom of infinity* (A4), $\emptyset^*$ stands for the *strong empty set* $\emptyset^* =_{def} \{x : {\sim}(x = x)\}$; $\{a\} =_{def} \{x : (x = a)\}$; and $a \cup b =_{def} \{x : (x \in a) \vee (x \in b)\}$ (all these terms are *legitimated*, that is, they can be constructed from the given axioms, see Remark 3.1). Finally, for the *replacement schema axiom* (A5), let $\psi(x, y)$ be a formula where variables x and y occur free, and where variable b does not occur free (other variables can occur free in ψ). Let z be a fresh variable. Then FUN_ψ denotes the following formula:

$$FUN_\psi =_{def} \forall x \forall y \forall z(\psi(x,y) \wedge \psi(x,z) \to (y = z)).$$

Remark 3.1 *From (A5) the* Separation Axiom *can be derived in ZFmbC:*

(Sep) $\exists b \forall x((x \in b) \leftrightarrow ((x \in a) \wedge \varphi(x)))$

where $\varphi(x)$ is a formula with variable x occurring free, and where variable b does not occur free (other variables can occur free in φ). Indeed, it is enough to consider $\psi(x,y) =_{def} \varphi(x) \wedge (x = y)$ in (A5), where y does not occur free in $\varphi(x)$ and $y \neq b$. The set b created from (Sep) is denoted by $\{x : (x \in a) \wedge \varphi(x)\}$. The strong empty set $\emptyset^* = \{x : {\sim}(x = x)\}$ is defined in ZFmbC by the Separation axiom and by extensionality, given that $\vdash ((x \in a) \wedge {\sim}(x = x)) \leftrightarrow {\sim}(x = x)$. From $\emptyset^*$, (A2) and (A5) it can be obtained

[3] From now on $x \neq y$ abbreviates the formula $\neg(x = y)$ and $x \notin y$ abbreviates the formula $\neg(x \in y)$, where $\neg$ is the weak (paraconsistent) negation of *mbC*.

$\{a,b\} = \{x : (x = a) \vee (x = b)\}$ for every a and b. In particular $\{a\}$ is legitimated, by taking $a = b$. Finally $a \cup b = \bigcup\{a,b\}$ where $\bigcup x$ is a notation for the set y created from axiom (A3). Observe that all the constructions above work in ZFmbC exactly as in ZF, given that the connectives $\neg$ and $\circ$ (the only ones that do not preserve logical equivalences) are not used in the above derivations.

Some immediate properties of equality and membership, most of them as a consequence of Regularity, are the following:

Proposition 3.2 *Let $n \geq 2$. In ZFmbC it holds that:*

(i) $\vdash \forall x((x \in x) \wedge (x \notin x) \to (x \neq x))$;

(ii) $\vdash \forall x(x = x)$;

(iii) $\vdash \forall x(C(x) \to {\sim}(x \in x))$;

(iv) $\vdash \forall x(C(x) \to (x \notin x))$;

(v) $\vdash \forall x_1 \ldots \forall x_n(\bigwedge_{i=1}^{n} C(x_i) \to {\sim}((x_1 \in x_2) \wedge (x_2 \in x_3) \wedge \ldots \wedge (x_n \in x_1)))$;

(vi) $\vdash \forall x_1 \ldots \forall x_n(\bigwedge_{i=1}^{n} C(x_i) \to \neg((x_1 \in x_2) \wedge (x_2 \in x_3) \wedge \ldots \wedge (x_n \in x_1)))$;

(vii) $\vdash C(\emptyset^*)$;

(viii) $\vdash \circ(\emptyset^* = \emptyset^*)$.

Proof. (i) Assume that $(x \in x) \wedge (x \notin x)$. But then it follows that $\exists z((z \in x) \wedge (z \notin x))$ and so $(x \neq x)$, by $(UnExt)$.

(ii) From the theorem $(z \in x) \leftrightarrow (z \in x)$ it follows that $\forall z((z \in x) \leftrightarrow (z \in x))$ and so $(x = x)$, by the extensionality axiom and by *Modus Ponens*.

(iii) Assume that $C(x)$. Then, it follows that $C(\{x\})$, by (cp) and extensionality. By regularity (replacing x by z such that $z = \{x\}$) and by classical logic it follows that ${\sim}(x \in x)$. The result follows by the Deduction Metatheorem.

(iv) It follows from (iii) and the fact that $\vdash {\sim}\varphi \to \neg\varphi$, by Proposition 2.1.

(v) Assume that $C(x_1) \wedge \ldots \wedge C(x_n)$. Then $C(\{x_1, \ldots, x_n\})$, by (cp) and extensionality. From this and by regularity (replacing x by z such that $z = \{x_1, \ldots, x_n\}$) it follows, by reasoning in classical logic, that ${\sim}((x_1 \in x_2) \wedge (x_2 \in x_3) \wedge \ldots \wedge (x_n \in x_1))$. The result follows by the Deduction Metatheorem.

(vi) By (v) and by $\vdash {\sim}\varphi \to \neg\varphi$ (cf. Proposition 2.1).

(vii) Since $\forall x{\sim}(x \in \emptyset^*)$ then $\forall x((x \in \emptyset^*) \to C(x))$. The result follows by axiom (cp) and by *Modus Ponens*.

(viii) This follows from (vii) and (C_1). ■

Remark 3.3 *Items (iii) and (v) in Proposition 3.2 suggest that ZF is not contained in ZFmbC, despite the latter being able to define the classical negation ${\sim}$. This is a slight distinction from other paraconsistent set theories, such as da Costa's CHU_n systems. Indeed, ZF can be regarded as a subsystem of CHU_n,* *since their underlying paraconsistent logic is an extension of the classical logic, and the axioms of ZF are derivable in such systems. On the other hand, in ZFmbC (and in its extensions) the Regularity Axiom (A6) is weaker than the original one from ZF: it only applies to consistent sets. However, as we shall see in the next proposition, ZF can be seen as a subsystem of ZFmbC up to the assumption that any set is consistent. That is, a kind of* Derivability Adjustment Theorem *(DAT, see [9]) between ZF and ZFmbC can be stated (see Theorem 3.4 below).*

For every formula φ in the language of *ZF* let φ^* be the corresponding formula in the language of *ZFmbC* obtained from φ by replacing every occurrence of $\neg$ by $\sim$. If Γ is a set of formulas in the language of *ZF* then Γ^* denotes the set of all the formulas of the form φ^*, for φ a formula in Γ. Then:

Proposition 3.4 (DAT between *ZF* and *ZFmbC*) *With the notation as above, let Γ be a set of closed formulas of ZF and let φ be a closed formula of ZF. If $\Gamma \vdash \varphi$ holds in ZF then $\forall x C(x), \Gamma^* \vdash \varphi^*$ holds in ZFmbC.*

Proof. Since *ZF* is compact, has a conjunction and satisfies the Deduction Metatheorem for closed formulas, and since $*$ preserves implication $\rightarrow$ and conjunction $\wedge$, it is sufficient to prove the result for theorems of *ZF*. Observe that, if φ is an instance of an axiom of *ZF* other than the Regularity Axiom, then φ^* is derivable in *ZFmbC*. On the other hand, if φ is an instance of the Regularity Axiom of *ZF* , then $\forall x C(x) \vdash \varphi^*$ holds in *ZFmbC*. Since *ZF* and *ZFmbC* have the same inference rules, the result follows easily by induction on the length of a derivation in φ in *ZF*. ■

At this point we can already refer to sets x which are not consistent, that is, satisfying $\neg C(x)$. This does not mean that these are inconsistent sets: indeed, at this stage, within the realm of *ZFmbC*, non-consistency is not the same as inconsistency, both for formulas and for sets. The next proposition shows how non-consistent sets could appear, namely through violation of basic properties of the identity predicate or of the membership predicate, as expressed in Proposition 3.2.

Proposition 3.5 *In ZFmbC it holds that:*

(i) $\vdash (x \neq x) \rightarrow \neg C(x)$*;*

(ii) $\vdash (x \in x) \rightarrow \neg C(x)$*;*

(iii) $\vdash C(x) \rightarrow \neg((x = x) \wedge (x \neq x))$*;*

(iv) $\vdash C(x) \rightarrow \neg((x \in x) \wedge (x \notin x))$*.*

Proof. (i) Assume that $(x \neq x)$. Note that $(x = x)$, by Proposition 3.2(ii), and so $(x = x) \wedge (x \neq x)$. From this it follows that $\neg\circ(x = x)$, by Proposition 2.1. The result follows by (C_2) and by the Deduction Metatheorem, by observing that the inference rules for quantifiers were not used in the derivation.

(ii) Assume that $(x \in x)$. Suppose firstly that $C(x)$. From Proposition 3.2(iv) it follows that $(x \notin x)$, and so $(x \in x) \wedge (x \notin x)$. But then it follows that $(x \neq x)$, by Proposition 3.2(i). Then $\neg C(x)$, by (i). Now, assuming $\neg C(x)$ instead of $C(x)$ it follows clearly that $\neg C(x)$. By (ax10) of *mbC* and the properties of the disjunction it follows that $C(x)$ is a consequence of $(x \in x)$. The result follows by using the Deduction Metatheorem.

(iii) From $C(x)$ and $(x = x) \wedge (x \neq x)$ it follows that $\circ(x = x) \wedge (x = x) \wedge (x \neq x)$, by (C_1), and then it follows $\bot$, by (bc1). But then $C(x) \vdash \neg((x = x) \wedge (x \neq x))$, by Proposition 2.1. The result follows once again by the Deduction Metatheorem.

(iv) From $C(x)$ and $(x \in x) \wedge (x \notin x)$ it follows that $C(x)$ and $(x \neq x)$, by Proposition 3.2(i). From this it follows that $C(x)$ and $(x = x) \wedge (x \neq x)$, which leads to $\bot$, by the proof of item (iii). Then $C(x) \vdash \neg((x \in x) \wedge (x \notin x))$, by Proposition 2.1. The result holds by the Deduction Metatheorem. ■

Proposition 3.6 *In ZFmbC it holds that:*

(i) $\vdash \sim((\emptyset^* = \emptyset^*) \wedge (\emptyset^* \neq \emptyset^*))$*;*

(ii) $\vdash \neg((\emptyset^* = \emptyset^*) \wedge (\emptyset^* \neq \emptyset^*))$*.*

Proof. (i) From $(\emptyset^* = \emptyset^*) \wedge (\emptyset^* \neq \emptyset^*)$ it follows that $(\emptyset^* \neq \emptyset^*)$ and so $\exists x((x \in \emptyset^*) \wedge (x \notin \emptyset^*))$, by $(UnExt)$. From this, $\exists x((x \in \emptyset^*)$. But $\vdash \forall x \sim(x \in \emptyset^*)$ and so $\vdash \sim\exists x(x \in \emptyset^*)$ (cf. [28]), which leads to the bottom. Thus it follows that $\vdash \sim((\emptyset^* = \emptyset^*) \wedge (\emptyset^* \neq \emptyset^*))$.
(ii) This follows from the fact that $\vdash \sim\varphi \to \neg\varphi$, by Proposition 2.1. ■

The main objective of the present study is to investigate whether or not the set theory *ZFmbC* could support contradictions without trivialization, that is, if it is indeed paraconsistent. Specifically, the following facts would be expected in *ZFmbC*:

$(x = x), (x \neq x) \nvdash \varphi$ for some formula φ, and

$(x \in x), (x \notin x) \nvdash \psi$ for some formula ψ.

On the other hand, we have the following:

Corollary 3.7 *In ZFmbC it holds:*

(i) $C(x), (x = x), (x \neq x) \vdash \varphi$ *for every formula* φ*;*

(ii) $C(x), (x \in x), (x \notin x) \vdash \varphi$ *for every formula* φ*.*

Proof. (i) It is a consequence of (C_1) and (bc1).

(ii) It is a consequence of Proposition 3.2(iii). ■

In order to guarantee that $C(x)$ is a necessary condition in the above derivations (which, in turn, guarantee that *ZFmbC* is in fact paraconsistent) the following question should be answered: is *ZFmbC* non-trivial? As we shall prove in Corollary 3.16, this system, as well as its extensions introduced herein, are indeed non-trivial, provided that *ZF* is consistent.

Different from other approaches to paraconsistent set theory (see, for instance, [2] and [34]) the universal set $V = \{x : x = x\}$ cannot be defined in *ZFmbC*, in order to avoid the triviality of such a set theory. Indeed, if V were definable in *ZFmbC*, then $\forall x(x \in V)$ would be a theorem and so it would be possible to define the set

$$A_\varphi = \{x : (x \in V) \wedge \varphi(x)\}$$

for every formula φ in which A_φ does not occur free, by the Separation (or Comprehension) axiom of *ZF*. In particular, the *strong Russell set*

$$\mathcal{R}^* = \{x : (x \in V) \wedge {\sim}(x \in x)\}$$

would be definable in *ZFmbC*, where $\sim$ is the strong (classical) negation definable in *mbC* as explained above. But then $(\mathcal{R}^* \in \mathcal{R}^*) \leftrightarrow {\sim}(\mathcal{R}^* \in \mathcal{R}^*)$ and so $(\mathcal{R}^* \in \mathcal{R}^*) \wedge {\sim}(\mathcal{R}^* \in \mathcal{R}^*)$ would follow, trivializing *ZFmbC* (this argument was adapted from [1]).

On this basis, one may ask if it is possible to start from a weaker paraconsistent system in which the strong negation is not definable. As Arruda has pointed out in [1], the Curry paradox follows in a paraconsistent set theory without strong negation if one allows the unrestricted version of the Comprehension axiom and considers the set

$$B_\varphi = \{x : (x \in x) \to \varphi\}$$

such that φ is any formula in which B_φ does not occur free. Adapting Arruda's argument once again, we can show the following: even by adopting the Separation Axiom of *ZF* instead of the unrestricted version of the Comprehension Axiom, if the universal set V were definable in a paraconsistent set system without strong negation, then the set

$$C_\varphi = \{x : (x \in V) \wedge ((x \in x) \to \varphi)\}$$

would be definable in that system, by the Separation Axiom of *ZF*. But the latter set is equal to the set B_φ, by extensionality and by the fact that $x \in V$ for every x, and so the system would be trivial.

The next step is to consider extensions of *ZFmbC* by taking stronger LFIs and appropriate axioms for the consistency operator C for sets. Thus, consider for example the system *ZFmCi*, obtained from *ZFmbC* by adding the following axioms:

(ci) $\quad \neg{\circ}\varphi \to (\varphi \wedge \neg\varphi)$

(cc_n) $\quad {\circ}\neg^n{\circ}\varphi \quad (n \geq 0)$

(C_3) $\forall x(\neg C(x) \to \neg\circ(x = x))$

(C_4) $\forall x(\neg C(x) \to \neg\circ(x \in x))$

The first two axioms transform the underlying paraconsistent logic *mbC* into the stronger logic *mCi*, in which the inconsistency operator can be defined by $\bullet\varphi =_{def} \neg\circ\varphi$ and so $\circ\varphi$ is equivalent to $\neg\bullet\varphi$ (see [9]). Axioms (C_3) and (C_4) intend to strengthen C, the consistency operator for sets, in order to obtain an inconsistency operator for sets I dual to C, defined by $I(x) =_{def} \neg C(x)$.

Lemma 3.8 *In mCi it holds:* $\vdash \sim(\varphi \wedge \neg\varphi) \to \circ\varphi$.

Proof. It is easy to prove by using valuations and then completeness of *mCi* with respect to such valuations (cf. [9]). ■

Proposition 3.9 *In ZFmCi it holds that :*

(i) $\vdash C(x) \to \circ(x \in x)$;

(ii) $\vdash \neg\circ(x \in x) \to \neg C(x)$;

(iii) $\vdash \neg C(x) \to (x \neq x)$;

(iv) $\vdash \neg C(x) \to (x \in x)$;

(v) $\vdash \circ(\emptyset^* \in \emptyset^*)$.

Proof. (i) Assume $C(x)$, $(x \in x)$ and $(x \notin x)$. By Proposition 3.2(i) it follows that $C(x)$ and $(x \neq x)$ and so $C(x)$, $(x = x)$ and $(x \neq x)$, by Proposition 3.2(ii). By (C_1) and *Modus Ponens* we get $\circ(x = x)$, $(x = x)$ and $(x \neq x)$ and then it follows $\perp$, by (bc1). By the Deduction Metatheorem and definition of $\sim$ it follows that $C(x) \vdash \sim((x \in x) \wedge (x \notin x))$ and then $C(x) \vdash \circ(x \in x)$, by Lemma 3.8. The result follows by the Deduction Metatheorem.

(ii) Assume $\neg\circ(x \in x)$. By (ci) and *Modus Ponens* it follows that $(x \in x) \wedge (x \notin x)$. Then $(x \neq x)$, by Proposition 3.2(i). Therefore $(x = x) \wedge (x \neq x)$, by Proposition 3.2(ii) and so $\neg\circ(x = x)$, by Proposition 2.1. Then, by (C_2) and *Modus Ponens* it follows that $\neg C(x)$. The result follows by the Deduction Metatheorem.

(iii) Assume $\neg C(x)$. Then $\neg\circ(x \in x)$, by (C_4), and so $(x \in x) \wedge (x \notin x)$, by (ci). By Proposition 3.2(i) it follows that $(x \neq x)$, as desired.

(iv) From $\neg C(x)$ it follows that $(x \in x) \wedge (x \notin x)$, by the proof of item (iii). From this we get $(x \in x)$, by the properties of conjunction.

(v) It follows from item (i) and from Proposition 3.2(vii). ■

Proposition 3.10 *In ZFmCi it holds that :*

(i) $\vdash \sim(C(x) \wedge \neg C(x))$;

(ii) $\vdash \neg(C(x) \wedge \neg C(x))$;

(iii) $\vdash \circ C(x)$.

Proof. (i) Assume $C(x)$. Then $\circ(x = x)$, by (C_1). On the other hand $\neg C(x) \vdash \neg\circ(x = x)$, by (C_3). Thus $C(x) \wedge \neg C(x) \vdash \circ(x = x) \wedge \neg\circ(x = x)$. But $\circ(x = x) \wedge \neg\circ(x = x)$ implies $\bot$, since $\vdash \circ\circ\varphi$ holds in *mCi* for every φ, by axiom (cc_0). The result holds by the Deduction Metatheorem and the definition of $\sim$.

(ii) If follows from (i) and Proposition 2.1.

(iii) If follows from (i) and Lemma 3.8. ∎

Now, recalling that $I(x) =_{def} \neg C(x)$, it is easy to prove the following:

Proposition 3.11 *In ZFmCi it holds that :*

(i) $\vdash \forall x(C(x) \leftrightarrow \neg I(x))$;

(ii) $\vdash \forall x(I(x) \leftrightarrow \neg\circ(x = x))$;

(iii) $\vdash \forall x(I(x) \leftrightarrow \neg\circ(x \in x))$;

(iv) $\vdash \forall x(I(x) \leftrightarrow \bullet(x = x))$;

(v) $\vdash \forall x(I(x) \leftrightarrow \bullet(x \in x))$;

(vi) $\vdash \forall x(I(x) \leftrightarrow ((x = x) \wedge (x \neq x)))$;

(vii) $\vdash \forall x(I(x) \leftrightarrow ((x \in x) \wedge (x \notin x)))$;

(viii) $\vdash \forall x(I(x) \leftrightarrow (x \neq x))$;

(ix) $\vdash \forall x(I(x) \leftrightarrow (x \in x))$.

Recall that the logic *Ci* is obtained from *mCi* by adding the axiom

(cf) $\qquad \neg\neg\varphi \to \varphi$

or, equivalently, by adding to *mbC* the axioms (ci) and (cf) (see [9]). Thus *ZFCi* is defined as the system obtained from *ZFmCi* by adding axiom (cf) or, equivalently, the system obtained from *ZFmbC* by adding axioms (ci), (cf), (C_3) and (C_4). On the other hand, the system *Cil* is obtained from *Ci* by adding the axiom

(cl) $\qquad \neg(\varphi \wedge \neg\varphi) \to \circ\varphi$

(see [9]). Let *ZFCil* be the system obtained from *ZFCi* by adding axiom (cl) plus

(C_5) $\qquad \forall x(\neg((x = x) \wedge (x \neq x)) \to C(x))$

(C_6) $\qquad \forall x(\neg((x \in x) \wedge (x \notin x)) \to C(x))$

Proposition 3.12 *In ZFCil it holds that:*

(i) $\vdash \forall x(C(x) \leftrightarrow \neg((x = x) \wedge (x \neq x)))$;

(ii) $\vdash \forall x(C(x) \leftrightarrow \neg((x \in x) \wedge (x \notin x)))$*;*

(iii) $\vdash \forall x(C(x) \leftrightarrow \circ(x = x))$*;*

(iv) $\vdash \forall x(C(x) \leftrightarrow \circ(x \in x))$*.*

Proof. (i) This is immediate from Proposition 3.5(iii) and axiom (C_5).

(ii) It follows from Proposition 3.5(iv) and axiom (C_6).

(iii) From $\circ(x = x)$ it follows that $\neg((x = x) \wedge (x \neq x))$, by Proposition 2.1, and so we get $C(x)$, by (C_5). The converse is simply (C_1).

(iv) From $\circ(x \in x)$ it follows that $\neg((x \in x) \wedge (x \notin x))$, by Proposition 2.1. Then it follows $C(x)$, by (C_6). The converse follows from Proposition 3.9(i). ■

Remarks 3.13 *(1) At this point, it can be interesting to summarize the list of equivalences relating the consistency predicate* C*, the identity predicate* $=$ *and the membership relation* $\in$ *found up to now in ZFCil:*

$C(x) \leftrightarrow \circ(x = x)$	$C(x) \leftrightarrow \neg((x = x) \wedge (x \neq x))$
$C(x) \leftrightarrow \circ(x \in x)$	$C(x) \leftrightarrow \neg((x \in x) \wedge (x \notin x))$
$\neg C(x) \leftrightarrow \neg\circ(x = x)$	$\neg C(x) \leftrightarrow (x \neq x)$
$\neg C(x) \leftrightarrow \neg\circ(x \in x)$	$\neg C(x) \leftrightarrow (x \in x)$

Additionally, $\vdash C(x) \to (x \notin x)$ *but the converse should not be valid. Some of the equivalences above should not be not valid in the weaker systems ZF*mbC *and ZF*mCi*.*

(2) The fact that $\circ(x \in x)$*, the consistence of* $(x \in x)$*, is equivalent to* x *be a consistent set, expressed by* $C(x)$ *(and, in particular, the fact that* $(\emptyset^* \in \emptyset^*)$ *is a consistent sentence) deserves some explanation. In the realm of LFIs, the consistency of a sentence* φ *means that it is 'reliable', or it has a 'classical' or 'expected' behavior (with respect to the principle of explosion). In semantic terms, it means that* $\circ\varphi$ *is true whenever the truth values of* φ *and* $\neg\varphi$ *are different (it holds from* mCi *on, see Lemma 3.8). In particular, if* φ *is always true and its negation is always false, or if* φ *is always false, then it is a consistent sentence. This is why the assumption* $C(x)$ *of consistency of a given set* x *forces to have* $(x \in x)$ *as being a consistent claim: it must be always false, under such assumption.*

As we can see, the expedient of separating sets and sentences into consistent and inconsistent, and moreover of using very cautious logics (that is, logics with weaker negations, but yet endowed with semantics in the best sense), permits us a finer control of the most basic reasoning underlying the foundations of (even inconsistent) mathematics.

A fundamental question to be answered is the following: Is *ZFCil* (and therefore its subsystems) non-trivial? Of course the answer to this question is related to the consistency of *ZF*, as we will show below.

Recall that a negation $\neg$ in a logic $\mathbf{L}$ is *explosive* if $\varphi, \neg\varphi \vdash \psi$ for every formula φ and ψ. A theory T in a logic $\mathbf{L}$ with an explosive negation $\neg$ is *inconsistent* if $T \vdash \varphi$ and $T \vdash \neg\varphi$ for some formula φ; otherwise it is *consistent*.[4] On the other hand, a theory T in a logic **L** is *trivial* if $T \vdash \varphi$ for every formula φ; otherwise, it is *non-trivial*. In case $\neg$ is explosive in **L**, then T is inconsistent if and only if it is trivial.

As a result, the question of the consistency of *ZF* (seen as a theory in first-order classical logic, where the negation $\neg$ is explosive) is equivalent to the statement of its non-triviality.

On the other hand, a theory based on a paraconsistent logic can be inconsistent but non-trivial. As such, the relevant question about *ZFCil* is to determine its non-triviality (seen as a theory of first-order *Cil*) instead of its inconsistency. As we shall see in Corollary 3.16, the non-triviality of *ZFCil* depends on the consistency of *ZF*, and so paraconsistent set theory based on LFIs is no more unsafe than the set theory *ZF*. This will be established by means of the following technical result, in the same vein of the famous Gödel embedding of classical logic into intuitionistic logic:

Proposition 3.14 *Let t be a mapping from the language of ZFCil to the language of ZF defined recursively as follows: for atomic formulas, $t(x \in y)$ is $(x \in y)$; $t(x = y)$ is $(x = y)$; $t(C(x))$ is $(x = x)$. For complex formulas, $t(\varphi * \psi)$ is $(t(\varphi) * t(\psi))$ for $*$ equals to $\wedge$, $\vee$ or $\to$; $t(\neg\varphi)$ is $\neg t(\varphi)$; $t(\circ\varphi)$ is $(t(\varphi) \leftrightarrow t(\varphi))$; $t(\forall x\varphi)$ is $\forall x\, t(\varphi)$; and $t(\exists x\varphi)$ is $\exists x\, t(\varphi)$. Then, t is an embedding of ZFCil into ZF, that is: if $\vdash \varphi$ in ZFCil then $\vdash t(\varphi)$ in ZF.*

Proof. It is sufficient to observe that, if φ is an instance of an axiom of *ZFCil*, then $t(\varphi)$ is derivable in *ZF*. On the other hand, any application of an inference rule in *ZFCil* is tranformed by t into an application of the same inference rule in *ZF*. Thus, any derivation of φ obtained in *ZFCil* can be transformed into a derivation of $t(\varphi)$ in *ZF*. ■

Theorem 3.15 *If ZF is consistent then ZFCil cannot prove any contradiction. That is: $\varphi \wedge \neg\varphi$ is not provable in ZFCil, for every φ.*

Proof. Suppose that $\varphi \wedge \neg\varphi$ is derivable in *ZFCil* for some formula φ. By Proposition 3.14, the formula $t(\varphi \wedge \neg\varphi)$ is derivable in *ZF*. But this implies that $t(\varphi) \wedge \neg t(\varphi)$ is derivable in *ZF* and so *ZF* is inconsistent. ■

Corollary 3.16 *If ZF is consistent then ZFCil is non-trivial.*

Proof. Suppose that *ZF* is consistent. By Theorem 3.15, no contradiction can be proved in *ZFCil*. Then *ZFCil* is non-trivial, since some formula cannot be derivable in it. ■

Assuming the consistency of *ZF*, Theorem 3.15 shows a basic feature of LFIs: contradictions are not *proved* in these systems, but they can be used as *hypothesis* without trivializing.

[4] We are using here the traditional terminology. If the negation $\neg$ is not explosive (that is, if **L** is paraconsistent) then it would be more appropriate to speak of *contradictory* and *non-contradictory* theories, respectively.

Consider now the following axioms that could be added to *ZFCil*:

$(cNC)_1$ $\exists y \forall x(x \in y \leftrightarrow \circ\varphi(x))$

$(cNC)_2$ $\forall x \circ\varphi(x) \to \exists y \forall x(x \in y \leftrightarrow \varphi(x))$

(crp) $C(y) \to \forall x(x \in y \to C(x))$

where (cNC) stands for 'consistent Naive Comprehension' and (crp) stands for 'consistency retropropagation'. The basic idea is that either $(cNC)_1$ or $(cNC)_2$ could make it possible to recover the main constructions of [34]. However, in the presence of the embedding t from Proposition 3.14, it is easy to see that t maps any instance of $(cNC)_1$ into a sentence that implies the existence of the universal set V in *ZF*. By its turn, any instance of $(cNC)_2$ is mapped by t into a sentence that implies an instance of the unrestricted version of the Comprehension Axiom, and thus able to obtain, in particular, the strong Russell set.

On the other hand, t maps any instance of (crp) into a sentence derivable in *ZF*. This being so, this axiom is the only acceptable one to be added to *ZFCil*, among the three axioms above, as far as we mantain the mapping t as an embedding that guarantees equi-non-triviality.

Another question concerns the existence of the universal set of all consistent sets, in the form

$$\mathsf{cons} = \{x \ : \ C(x)\}.$$

Intuitively, this would constitute a kind of universal set. Moreover, if cons is legitimated then, using again the embedding t, the universal set V would be legitimated in *ZF*. Observe that, in the presence of (cp), it follows that $C(\mathsf{cons})$ and therefore $\mathsf{cons} \in \mathsf{cons}$, trivializing the system in the presence of the Regularity Axiom (see Proposition 3.2(iii)).

4 Can inconsistent sets be regarded as proper classes?

In 1899, Cantor discovered, or became aware of,[5] the famous paradox about the cardinal number of the set of all sets. In intuitive terms, it may be described as follows: on the one hand, there must be a greatest possible cardinal, call it C. On the other hand, if this cardinal number is a set, the cardinal number of the power set of C is strictly larger than the cardinal number of C (this is now known as Cantor's theorem). Hence, either this cardinal number does not exist, or it is not a set. This paradox, together with Burali-Forti's paradox of 1897 (which proves that naïvely constructing "the set of all ordinal numbers"

[5] It is worth noting that in [26] it is shown that Burali-Forti's paradox, generally regarded as the first of the set-theoretical paradoxes, was neither created by Burali-Forti nor by Cantor. It arose gradually and only acquired its contemporary form in the hands of Bertrand Russell in 1903.

leads to an antinomy) led Cantor to formulate a concept called "limitation of size" according to which collections such as the one of all ordinals, or of all sets, was an "inconsistent multiplicity", too large to be a set. Such collections later became known as proper classes. In the words of J. W. Dauben in [19]:

> "Anything that was too large to be comprehended as a well defined, unified, consistent set was declared inconsistent. These were "absolute" collections, and lay beyond the possibility of mathematical determination. This, in essence, is what Cantor communicated first to Hilbert in 1897, and somewhat later to Dedekind in his letters of 1899."

According to [19], Cantor himself believed, in his early period, that the idea of the actual infinite "could not be consistently formulated and so had no place in rigorous mathematics", and he took seriously Kronecker's criticisms [6], which were echoed in his major work on set theory, [8]. This work was a truly mathematical-philosophical investigation of the infinite, as the title indicates, and was where he declared his famous pronouncement that the essence of mathematics is exactly its freedom. Such a sense of freedom, together with his method of using in a positive way what seemed to be paradoxical, as explained by A. Kanamori in the quotation below from [22], is what raises doubts whether an 'inconsistent multiplicity' would coincide with what we today know as a proper class:

> "Cantor in 1899 correspondence with Dedekind considered the collection Ω of all ordinal numbers as in the Burali-Forti Paradox, but he used it *positively* [7] to give mathematical expression to his Absolute. He defined an "absolutely infinite or inconsistent multiplicity" as one into which Ω is injectible, and proposed that these collections be exactly the ones that are not sets. He would thus probe the very limits of sethood using his positive concept of power!"

P. Maddy, in her search for a realistic theory of sets and classes (cf. [24]), tries to provide a theory of both sets and classes (along the lines of J. König), claims for a theory where:

(1) classes should be real, well-defined entities;
(2) classes should be significantly different from sets

Maddy is not alone in questioning whether the Fregean formulation of naive set theory (which was later refuted by Russell's paradox) would really be a faithful interpretation of the Cantorian conception of sets. As A. Weir (cf. [35], p. 766) observes:

[6] L. Kronecker, a prominent German mathematician who had been one of Cantor's teachers, even attacked Cantor personally, calling him a "scientific charlatan" a "renegade" and a "corrupter of the youth"!

[7] Emphasis in the original.

> "...it may well be seriously mistaken to think of Cantor's Mengenlehre as naive..."

As we know today, in *ZF* set theory the notion of class is metalinguistic (actually, *ZF* does not refer to classes at all), whereas in other set theories, such as von Neumann–Bernays–Gödel (*NBG*), the notion of "class" is axiomatized and interpreted as an entity that is not a member of any other entity. Indeed, in *NBG* classes are the basic objects, and a set is defined as a class that is an element of some other class. Like *NBG*, Morse–Kelley (*MK*) set theory admits proper classes as basic objects, but whereas *NBG* is a conservative extension of *ZF*, *MK* is strictly stronger than both *NBG* and *ZF*. Other set theories, such as New Foundations (*NF*) and the theory of semisets, treat the relationship between sets and classes in still different ways. The set theoretical proposals of P. Finsler, centered around his concept of "circle-free" sets, are regarded in principle as incoherent (see [21]). The idea was reworked (apparently independently) by W. Ackermann, and makes for a quite different theory from the standard one (it is considered to be the first genuine 'alternative set theory'). Ackermann's set theory, as explained in [21], is a theory of classes in which some classes are sets, and indeed the notion of set is indefinable – there is no simple definition of which classes are sets. It turns out, however, to be essentially the same theory as *ZF*. Coming now to our question, whether inconsistent sets can be equated (or at least somehow regarded as) proper classes, we see no reason why Cantor's "inconsistent multiplicities" could be so simply reduced to the notion of proper classes. Firstly, as it is clear from the above overabundance of approaches to sets and classes, there is no unique sense of what proper classes would be. One criticism might be that there is even less agreement on what an inconsistent collection would be, but we are not defining what an inconsistent collection is – rather, we are postulating their rational possibility (which is much less than their existence), and showing how they can be handled from a coherent logical standpoint. Our approach views inconsistent situations as possible evidence that something may have gone wrong, as a sign that contradictions may appear, but not as seal of condemnation– only in this way, we believe, can one explore positively the very limits of sethood, as Cantor did.

However, even if not coincident, the notions of being consistent (as formalized here) and being a set are related: in fact, let $set(x) =_{def} \exists y(x \in y)$ be the usual notion of sethood in a theory of classes. In case $\vdash set(x) \to C(x)$, then the following will be obtained: $x \in x \vdash set(x) \vdash C(x) \wedge (x \in x) \vdash \bot$ and so $x \in x$ cannot be satisfied for any entity x. Of course this is the case in classical set (or class) theories because of the Regularity Axiom, but in a paraconsistent set theory $(x \in x)$ and $(x \notin x)$ should not be necessarily excluded. Therefore the non-classical character of our theory would vanish by requiring the relationship between sethood and consistency in the above way. An interesting point is that the converse relation can be added to the theory without any further problems:

(cs) $\quad C(x) \to set(x)$ $\quad$ (*consistency as sethood* axiom).

On the other hand, if $class(x) =_{def} \neg\exists y(x \in y)$ denotes the proper-classhood predicate as usual, it would be possible to consider the following principle:

(pci) $\quad class(x) \to \neg C(x)$ $\quad$ (*proper-classhood as inconsistency* axiom)

together with (or as an alternative to) axiom (cs).

Our theoretical plan for set theory by predicating on consistency has no need to overpopulate the usual formal ontology of sets, so we do not need any kind of strange new objects (aside from the usual ones of *ZF*). If we consider what is called "dialetheia" (cf. [29]) as a sentence which is both true and false, our theory is even free of "dialetheias" since in our setting a sentence can be seen as true and its negation as also true (at least provisionally) without either of them falling into the anathema of falsity. But we might purposedly add, for instance, what may be called the Russell set axiom, and keep harvesting its consequences without falling into triviality:

$(Russ)$ $\qquad \exists x((x \in x) \wedge (x \notin x))$

From Proposition 3.5 it follows that the Russell sets are inconsistent:

Corollary 4.1 *Let $\mathcal{R}$ be a Russell set, that is,* $\vdash (\mathcal{R} \in \mathcal{R}) \wedge (\mathcal{R} \notin \mathcal{R})$. *Then* $\vdash \neg C(\mathcal{R})$.

Postulating the Russell set axiom $(Russ)$ is not really necessary for the development of a paraconsistent set theory. However, we can indulge in a dialetheistic assumption if we want to enrich set theory towards an antinomic extension. Although the postulation of a Russell set takes us closer to Routley's set theory, we could as well remove this axiom from our systems while tolerating inconsistent sets as premises. In this sense, our approach is not metaphysically committed. According to da Costa in [13]:

> "The main concern to paraconsistent set theory is not to make possible the existency, and thereby the investigation of some sets which can cause trouble in naive set theory, such as Russell's set, Russell's relations and the set of all non-k-circular sets ($k = 1, 2, \ldots$). On the contrary, the most important characteristic of paraconsistent theories is that they allow us to handle the extensions of 'inconsistent' predicates which may exist in the real world or are inherent in some universes of discourse in the fields of science and philosophy."

On the other hand, the inclusion of a Russell set in our systems would require a new proof of non-triviality of the system relative to the consistence of *ZF* (cf. Corollary 3.16). Indeed, if *ZF* is consistent then the mapping t defined in Proposition 3.14 would no longer be an embedding between the system extended with $(Russ)$ into *ZF*. In fact, t maps the axiom $(Russ)$ into a sentence not derivable in *ZF* (cf. Theorem 3.15).

Related to the question of postulating the existence of some sets, observe that the existence of the *strong* empty set $\emptyset^*$ such that $\vdash \forall x {\sim}(x \in \emptyset^*)$ is guaranteed by the axioms of *ZFmbC*, and in fact this set plays an important role in the theory. What about the existence of the *weak* empty set? It would be a set $\emptyset$ such that $\vdash \forall x \neg(x \in \emptyset)$, that is, satisfying $\vdash \forall x(x \notin \emptyset)$. Since it is not obvious that the existence of such a set can be proved even in *ZFCil*, its existence should be postulated by an specific axiom:

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

The convenience of postulating the existence of such an object should be ponderated. It is clear that, as happens with Russell sets, its existence is not relevant for the development of our paraconsistent set theory.

5 Models, comparisons and further work

Although models of (classical) set theory can be seen as special cases of models of a first-order language, some models can also be built using proper classes. But what would be the models of *ZFCil* ? In what follows, we explore some preliminary ideas on this question, with the understanding that there is much to be done, and that we are just scratching the surface of this topic here. The symbol "ϵ" will be used for denoting membership in the metalanguage, while "$\in$" will denote the membership relation symbol in the object (first-order) language. T. Libert considers in [23] a class of structures suitable for paraconsistent set theories. Such interpretation structures are pairs of the form $\mathcal{M} = \langle M; [\cdot]_{\mathcal{M}} \rangle$ such that M is a non-empty set and $[\cdot]_{\mathcal{M}} : M \to \wp_p(M)$ is a function, where $\wp_p(M) =_{def} \{\langle X, Y \rangle \ : \ X, Y \subseteq M \text{ and } X \cup Y = M\}$. The equality symbol is interpreted in the standard way. For every $b \epsilon M$, the pair $[b]_{\mathcal{M}} = \langle [b]^+_{\mathcal{M}}, [b]^-_{\mathcal{M}} \rangle$ is such that $[b]^+_{\mathcal{M}} = \{a \epsilon M \ : a \in_{\mathcal{M}} b\}$ is the *positive extension* of b, while $[b]^-_{\mathcal{M}} = \{a \epsilon M \ : a \notin_{\mathcal{M}} b\}$ is its *negative extension.* A *strong extensionality* is required, namely that the function $[\cdot]_{\mathcal{M}}$ be injective, and so for every $a, b \epsilon M$ it holds that: $[a]^+_{\mathcal{M}} = [b]^+_{\mathcal{M}}$ and $[a]^-_{\mathcal{M}} = [b]^-_{\mathcal{M}}$ iff $a = b$. This leads naturally to a 3-valued paraconsistent logic. It is noteworthy to stress the similarity between Libert's interpretation structures and the partial models underlying the theory of quasi-truth, introduced by I. Mikenberg, N. da Costa and R. Chuaqui in [25]. The notion of quasi-truth, also called pragmatic truth and partial truth, is a generalization of Tarski's concept of truth in a structure. O. Bueno and N. da Costa defend in [6], for instance, the view that if scientific theories are taken to be quasi-true, and if the underlying logic is paraconsistent, it is perfectly rational for scientists and mathematicians to entertain theories involving contradictions without triviality, and that such a move provides a new way of thinking about the foundations of science (Bohr's theory of the atom and classical electrodynamics, for instance, are well-known examples discussed in the literature on scientific theories involving contradictions). In [25] a rigorous formal presentation of this idea was given, of which one of the most salient features is the notion of partial structure. It is even possible to think of a theory of quasi-sets (see for instance [5] and [17]) as a theory of indistinguishable objects, motivated mainly by the assumption that one cannot meaningfully apply the notion of identity to quantum particles.

Such observations make it very natural to consider, generalizing the classical state of affairs, special cases of models of first-order paraconsistent language as models of *ZFCil*, and indeed (though we shall not go into details here) exploring such models seem to be a promising approach to understanding the distinction between consistent and inconsistent sentences, the distinctions between consistent and inconsistent sets, and the relations between inconsistent sentences and inconsistent sets.

Another point which would deserve attention is the possibility of studying a variant of our paraconsistent set theory in the direction of the non-well-founded set theories, obtained by replacing the Regularity Axiom (also known in this context as the axiom of foundation) by one or more axioms implying its negation. The resulting would certainly give an interesting theory of circular sets, but nothing has been done on this regard yet.

Part of our interest is to compare our approach with the main alternative solutions, and to show that the Logics of Formal Inconsistency capture the essentials of several other proposals. By admitting that sets, as well as sentences, can be inconsistent, and that only consistent and contradictory objects may cause triviality, our view offers support to proposals such as the one by W. Byers [7], where thinking about math requires creativity that is promoted, rather than neglected, using forms of thought connected to contradiction. Aside from Russell's antinomy, it is natural to extend our considerations on Berry's and to the Burali-Forti's antinomies.

Moreover, a criticism that can be raised against category theory is that most of its mechanisms are just inspired on *ZFC* set theory. There is no reason not to consider a wider version of category theory that would include inconsistent situations, such as (simultaneously) commuting and non commuting diagrams, and the like.

Inconsistent mathematics, and in more general terms paraconsistent category theory, may have interesting applications to quantum mathematics and to the study of quantum measurement, in particular to dynamic systems in which there are discontinuous jumps, as suggested by C. Mortensen in [27].

It is hard to conceive of mathematics today without the completed infinite, and if it had not been stated by a mathematician of the intellectual stature of Kronecker, we would consider as pure insanity his claim (cf. e.g. [19]) that:

> "I don't know what predominates in Cantor's theory philosophy or theology, but I am sure that there is no mathematics there."

Who can be sure where is there mathematics, or what kind of new mathematics will be revealed by yet unthought entities as inconsistent sets?

Acknowledgements: This research was financed by FAPESP (Thematic Project LogCons 2010/51038-0, Brazil) and by individual research grants from The National Council for Scientific and Technological Development (CNPq), Brazil. We are indebted to Rodrigo de Alvarenga Freire for furthering our intuitions on inconsistent sets, and to two anonymous referees for suggestions that helped to improve the paper.

References

[1] A. I. Arruda. The Russell paradox in the systems NFn. In: *Brazilian Conference on Mathematical Logic 3* (Proceedings). A. I. Arruda, N. C. A. da Costa, A. M. Sette (eds.), São Paulo: Sociedade Brasileira de Lógica, 1980, p.1–12

[2] A. I. Arruda and D. Batens. Russell's set versus the universal set in paraconsistent set theory. *Logique et Analyse* 98, 1982, p. 121–133.

[3] A. Avron. A new approach to predicative set theory, In: *Ways of Proof Theory* , R. Schindler (ed.), Ontos Mathematical Logic, Ontos Verlag, 2010, p. 31–63.

[4] R. T. Brady. The Non-Triviality of Dialectical Set Theory. In: *Paraconsistent Logic: Essays on the Inconsistent*, G. Priest, R. Routley and J. Norman (eds.), München: Philosophia Verlag, 1989, p. 437–471.

[5] O. Bueno. Quasi-truth in quasi-set theory. *Synthese* 125(1-2): 33–53, 2000.

[6] O. Bueno and N. C. A. da Costa. Quasi-truth, paraconsistency, and the foundations of science. *Synthese* 154(3): 383–399, 2007.

[7] W. Byers. *How Mathematicians Think: Using Ambiguity, Contradiction, and Paradox to Create Mathematics.* Princeton University Press, 2007.

[8] G. Cantor. *Grundlagen einer allegemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen.* Leipzig: B. O. Teubner, 1883. English translation in Ewald, William B., 1996. *From Kant to Hilbert: A source book in the foundations of mathematics*, Oxford: Oxford University Press, volume 2.

[9] W. A. Carnielli, M. E. Coniglio and J. Marcos. Logics of Formal Inconsistency. In D. Gabbay and F. Guenthner, (eds.) *Handbook of Philosophical Logic (2nd edition)*, volume 14, pages 1–93. Springer, 2007.

[10] W. A. Carnielli and J. Marcos. A taxonomy of **C**-systems. In: *Paraconsistency — The logical way to the inconsistent*, W. A. Carnielli, M. E. Coniglio and I. M. L. D'Ottaviano, (eds.), volume 228 of *Lecture Notes in Pure and Applied Mathematics*, pages 1–94, New York, 2002. Marcel Dekker.

[11] A. Church. Set theory with a universal set. In: *Proceedings of the Tarski Symposium*, L. Henkin (ed.), Providence: A.M.S., 1974, pp. 297–308.

[12] N. C. A. da Costa. Inconsistent Formal Systems (in Portuguese), Habilitation Thesis, 1963. Republished by Editora UFPR, Curitiba, 1993.

[13] N. C. A. da Costa. On paraconsistent set theory. *Logique et Analyse* 29(115): 361–71, 1986.

[14] N. C. A. da Costa. Paraconsistent Mathematics. In D. Batens et al. (eds.), Frontiers of Paraconsistent Logic, Hertfordshire: Research Studies Press, 165–180, 2000.

[15] N. C. A. da Costa, J.-Y. Béziau and O. Bueno. Elementos de Teoria Paraconsistente de Conjuntos (*Elements of Paraconsistent Set Theory*, in Portuguese). Volume 23 of *Coleção CLE*, CLE-Unicamp, Campinas, 1998.

[16] N.C.A. da Costa and S. French, S. The model-theoretic approach in the philosophy of science. *Philosophy of Science* 57(2): 248–265, 1990.

[17] N.C.A. da Costa and D. Krause: Logical and Philosophical Remarks on Quasi-Set Theory. *Logic Journal of the IGPL* 15(5-6): 421–431, 2007.

[18] J. W. Dauben. *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Princeton University Press, 1990.

[19] J. W. Dauben. The Battle for Cantorian Set Theory. In: *Mathematics and the Historians Craft: The Kenneth O. May Lectures*, edited by Michael Kinyon and Glen van Brummelen. New York: Springer Verlag, Canadian Mathematical Society Books in Mathematics, 2005.

[20] R. Hinnion. Naive Set Theory with Extensionality in Partial Logic and in Paradoxical Logic. *Notre Dame Journal of Formal Logic* 35 (1):15–40,1994.

[21] M. R. Holmes. Alternative Axiomatic Set Theories. *The Stanford Encyclopedia of Philosophy* (2012 Edition), Edward N. Zalta (ed.), `http://plato.stanford.edu/entries/settheory-alternative/#ClaTheOveZFC`

[22] A. Kanamori. The Mathematical Development of Set Theory from Cantor to Cohen. *The Bulletin of Symbolic Logic*, 2(1):1–71, 1996.

[23] T. Libert. Models for a paraconsistent set theory. In: Annals of *A Paraconsistent Decagon: The Workshop on Paraconsistent Logic*, Trento, Italy, 2002. *Journal of Applied Logic* 3(1):15–41, 2005.

[24] P. Maddy. Proper classes. *The Journal of Symbolic Logic* 48(1):113–139, 1983.

[25] I. Mickenberg, N. C. A da Costa and R. Chuaqui. Pragmatic truth and approximation to truth. *The Journal of Symbolic Logic* 51(1):201–221, 1986.

[26] G. H. Moore and A. Garciadiego. Burali-Forti's paradox: A reappraisal of its origins. *Historia Mathematica* 8(3):319–350, 1981.

[27] C. Mortensen. Inconsistent Mathematics. *The Stanford Encyclopedia of Philosophy* (2008 Edition), Edward N. Zalta (ed.), `http://plato.stanford.edu/entries/mathematics-inconsistent/`

[28] W. A. Carnielli, M. E. Coniglio, R. Podiacki and T. Rodrigues. Completeness theorems for First-Order Logics of Formal Inconsistency. To appear.

[29] G. Priest and K. Tanaka. Paraconsistent Logic. *The Stanford Encyclopedia of Philosophy* (2009 Edition), Edward N. Zalta (ed.), `http://plato.stanford.edu/archives/sum2009/entries/logic-paraconsistent/`

[30] G. Restall. A Note on Naive Set Theory in LP. *Notre Dame Journal of Formal Logic* 33(3):422–432, 1992.

[31] R. Routley. Ultralogic as Universal? *Relevance Logic Newsletter* 2: 50–90 and 138–175, 1977. Reprinted in [32].

[32] R. Routley. In Exploring Meinong's Jungle and Beyond. Departamental Monograph, Philosophy Department, RSSS, Australian National University, vol. 3, pp. 892–962. Canberra: RSSS, Australian National University, Canberra, 1980.

[33] P. Verdee. Strong, universal and provably non-trivial set theory by means of adaptive logic. *Logic Journal of the IGPL* 21 (1):108–125, 2013, doi: 10.1093/jigpal/jzs025.

[34] Z. Weber. Transfinite Numbers in Paraconsistent Set Theory. *Review of Symbolic Logic* 3(1):71–92, 2010.

[35] A. Weir. Naive Set Theory is Innocent! *Mind* 107(428):763–798, 1998.