Russell's paradox and *free* zig zag solutions

Ludovica Conti

FINO - Northwestern Philosophy Consortium
University of Pavia

June 29th, 2019 Anogeia

Plan of the talk

1. The debate about Russell's Paradox
   - Russell's Paradox
   - Cantorian vs Predicativist explanations
   - *Zig zag* solutions

2. Extensionalist explanation and free *zig zag* solutions
   - Extensionalist explanation
   - Negative free logic and Russell's paradox
   - Free Fregean theories

Paradox: minimal version of a contradiction's derivation.
$\rightarrow$ list of all and only necessary premises;
$\rightarrow$ elimination (or relevant change) of each of them is sufficient to avoid the contradiction.

Explanation: instruction for a solution.
Expl. 1: selection of the specific *guilty* premise:
what premise we have to change to solve the paradox.
Expl. 2: indication of the *guilt* itself:
how we have to change a (selected) premise to solve the paradox.

Solution: specific change of the derivation which
- follows from an explanation;
- is sufficient to avoid the contradiction;
- is able to preserve as much as possible the derivational power of the theory.

1. $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \forall x(Xx \leftrightarrow Yx))$ (BLV)
2. $\exists X \forall x(Xx \leftrightarrow \exists Y(x = \epsilon Y \wedge \neg Yx))$. Call this concept R. (CA)
3. $\exists x(x = \epsilon R)$ (2, AT)
4. $\neg R \epsilon R$ (A)
5. $R \epsilon R$ (2,4)
6. $\neg R \epsilon R \rightarrow R \epsilon R$ (4,5)
7. $R \epsilon R$ (A)
8. $\exists Y(\epsilon R = \epsilon Y \wedge \neg Y \epsilon R)$ (2,7)
9. $\neg R \epsilon R$ (1,8)
10. $R \epsilon R \rightarrow \neg R \epsilon R$ (7,9)
11. $R \epsilon R \leftrightarrow \neg R \epsilon R$ (6,10)

Traditional debate

"Boolos and I are agreed that Frege's theory would be rendered consistent if either (i) Axiom V were deleted, or (ii) only first-order quantification were admitted. The substance of our disagreement is therefore restricted to the question which is the snow and which the yodel, in his metaphor, or which the match and which the matchbox." (Dummett 1993)

Cantorian explanation:

Expl.1: BLVb $\forall X \forall Y(\epsilon X = \epsilon Y \rightarrow \forall x(Xx \leftrightarrow Yx))$

Expl.2: injectivity of the extensionality function

Predicativist explanation:

Expl.1: CA: $\exists X \forall x(Xx \leftrightarrow \phi(x))$ - where $\phi(x)$ does not contain X free

Expl.2: impredicativity of concepts' specification

Cantorian explanation

Expl.1: BLVb $\forall X \forall Y(\epsilon X = \epsilon Y \rightarrow \forall x(Xx \leftrightarrow Yx))$
Expl.2: injectivity of extensionality function, namely a (semantic and syntactic) incompatibility with Cantor's theorem.

Semantic argument: the existential assumption of an injective function (derivable from BLVb) from the concepts' domain to the objects' one imposes an unsatisfiable cardinality request - namely that the object's domain has (at least) the same cardinality of the concept's domain.

Syntactic argument: the existential assumption of an injective function from the concepts' domain to the objects' one (derivable from BLVb) is inconsistent with Cantor's theorem

Semantic argument

[The existential assumption of an injective function (derivable from BLVb) from the concepts' domain to the objects' one imposes an unsatisfiable cardinality request - namely that object's domain has (at least) the same cardinality of concept's domain]

Limitation (Heck 1996):
given the incompleteness of pure second-order logic, the unsatisfiability of BLVb (in standard models of the language) is not an explanation of the inconsistency.

Objection:
there are some secondary models (e.g. Henkin's models) in which the objects' domain and the concepts' domain have the same cardinality: in these models BLV is unsatisfiable, even if the cardinalities of the second-order's domain and the first-order's one admit an injection.

Syntactic argument

The real contradiction of Frege's system arises between two theorems:
- existential generalisation of BLVb ($\exists \iota \forall X \forall Y(\iota X = \iota Y \rightarrow \forall x(Xx \leftrightarrow Yx))$
- Cantor's theorem ($\neg \exists \iota \forall X \forall Y(\iota X = \iota Y \rightarrow \forall x(Xx \leftrightarrow Yx))$).

Russell's contradiction ($R \epsilon R \leftrightarrow \neg R \epsilon R$) is only a subordinate consequence

Objections:

1) presupposes a different reconstruction of the paradoxical derivation

2) anything follows from a contradiction (*ex falso quodlibet*)
Russell's contradiction follows from the original contradiction in the same way in which anything follows from this contradiction:
it is not clear why $R \epsilon R \leftrightarrow \neg R \epsilon R$ is the proper *symptom* of that contradiction

3) both these propositions are theorems, so the original contradiction has to be looked for in the axioms or assumptions from which they follows:
$\exists \iota \forall X \forall Y(\iota X = \iota Y \rightarrow \forall x(Xx \leftrightarrow Yx))$ follows from $\exists$-I, BLV;
$\neg \exists \iota \forall X \forall Y(\iota X = \iota Y \rightarrow \forall x(Xx \leftrightarrow Yx))$ follows from $HOL^{=}$(with CA),

Cantorian solution

The Cantorian solution - intended as fixing cardinalities or weakening standard BLVb just in order to avoid the alleged original contradiction (with Cantor's theorem) - is not sufficient to avoid the contradiction: there is a derivation of the same contradiction from a restricted version of BLV (*Definable- BLV*) that is compatible with Cantor's theorem (Paseau 2015).

1. $\forall X(\forall x(Xx \leftrightarrow \phi x) \rightarrow \forall Y(\epsilon X = \epsilon Y) \leftrightarrow \forall x(Xx \leftrightarrow Yx))$ (Def.-BLV)
2. $\exists X \forall x(Xx \leftrightarrow \exists Y(x = \epsilon Y) \wedge \neg Yx))$. Call this concept R. (CA)
3. $\exists x(x = \epsilon R)$ (2, AT)
4. $\forall x((Xx \leftrightarrow \exists Y(x = \epsilon Y) \wedge \neg Yx)) \rightarrow \forall Y(\epsilon R = \epsilon Y \leftrightarrow \forall x(Rx \leftrightarrow Yx))$
5. $\forall Y(\epsilon R = \epsilon Y \leftrightarrow \forall x(Rx \leftrightarrow Yx))$ (2,4, MP)
6. $\neg R(\epsilon R)$ (A)
7. ...

13. $R(\epsilon R) \leftrightarrow \neg R(\epsilon R)$ (9,12)

Predicativist explanation

Expl.1: CA: $\exists X \forall x(Xx \leftrightarrow \phi(x))$ - where $\phi(x)$ does not contain X free
Expl.2: impredicativity of concepts' specification

Arguments: the inconsistency follows from the specification of Russell's concept because of its impredicativity - intended as implicit and vicious circularity, source of indefinite extensibility, lack of definitional guarantees (...)

Predicativist arguments are not very strong because there are several other impredicative but consistent abstraction principles
(cfr. HP:$\forall F \forall G(\sharp F = \sharp G \leftrightarrow F \approx G)$).

Predicativist Solution

Solutions: predicative restrictions of the comprehension's formula of CA
Predicative Subsystems of *Grundgesetze*
(Cfr. Heck 1996, Wehemeier 1999, Ferreira-Wehemeier 2002)

Predicativist solutions work but are very weak:
- avoid the contradiction but
- allow to derive only Robinson Arithmetic Q
(prevent the derivation of Peano Arithmetic PA - first goal of the original Fregean proposal).

Predicativist Expl. 1 is correct because CA is a necessary condition of Russell's paradox;
Predicativist Expl. 2 admits objections because it identifies a feature that is necessary not only for the contradiction but also for the derivation of PA.

Zig zag solutions

Russell's *zig zag* proposal:
"In the zigzag theory, we start from the suggestion that propositional functions determine classes when they are fairly simple, and only fail to do so when they are complicated and recondite"

General idea:
Not all *propositional functions* (open formulas) determine *classes* (extensions).

In our terms - admitted that every open formulas specifies a concept:
- the full second-order domain is specified (unlike predicativist solutions);
- the correlation between concepts and extensions is injective (unlike cantorian solutions);
- the correlation between concepts and extensions is not total.

Carving the correlation. First way.

I) CARVING CORRELATION BY A DISTINCTION ON THE CONCEPTS' DOMAIN.

every open formulas specifies a concept but there are two sort of open formulas:

- *predicative* formulas that specifies concepts related to extensions;

- *not-predicative* formulas that specifies concepts that go *zig zag* between the extensions.

Simplifying: there are two sort of concepts - defined by formulas - predicative and not-predicative.

What it means to be *predicative*?

(Russell) fairly simple

(Boccuni 2010- Ferreira 2018) A definition is said to be predicative if it does not quantify over a totality to which the entity being defined belongs. Otherwise the definition is said to be impredicative.

A comprehension axiom schema is said to be predicative if the comprehension formula ϕ(x) contains no bounded second-order variables, and impredicative otherwise.

PE *Predicative Extensions* (Ferreira 2018)

Two sorted second order language (primitive symbols):

- denumerably many first-order variables: x, y, z;
- denumerably many PREDICATIVE second-order variables: F, G, H;
- denumerably many IMPREDICATIVE second-order variables: *F, G, H*;
- logical constants: $\neg, \wedge, \vee, \rightarrow$;
- quantifiers for each order and sort of variables:$\exists$ x, $\exists$ F, $\exists$ *F*;
- operator term-forming (ˆ) applied to open formulas.

Syntax:

- Complex singular terms: if $\phi(x)$ is a PREDICATIVE formula, $\hat{}x.\phi(x)$ is a (complex) singular term;
- Atomic formulae: if Π is a (PREDICATIVE or IMPREDICATIVE) second order variable and x is a first order variable, $\Pi(x)$ is an atomic formulas;
- Complex formulae by usual inductive definition.

Axioms of PE:

- Second order Logic;

- Predicative comprehension axiom schema: $\exists F \forall x (Fx \leftrightarrow \phi(x))$ - where $\phi(x)$ is a PREDICATIVE formula (without F free);

- Impredicative comprehension axiom schema: $\exists F \forall x (Fx \leftrightarrow \phi(x))$ - where $\phi(x)$ is a IMPREDICATIVE formula;

- schematic Basic Law V: $\hat{}x.\phi(x) = \hat{}x.\psi(x) \leftrightarrow \forall x(\phi(x) \leftrightarrow \psi(x))$
$\rightarrow$ Automatically restricted to PREDICATIVE formulas $(\phi(x), \psi(x))$.

PG *Plural Grundgesetze* (Boccuni 2010)

Second order with two sorted first order language:

- denumerably many first-order SINGULAR variables: x, y, z;
- denumerably many first-order PLURAL variables: xx, yy, zz;
- denumerably many second-order variables (*conceptual variables*): *X, Y, Z*;
- logical conectives: $\neg, \wedge, \vee, \rightarrow$ - quantifiers for each order and sort of variables:$\exists x, \exists xx, \exists X$;
- relational constant η, between fol SINGULAR and fol PLURAL variables;
- operator term-forming ($\hat{}$) applied to open formulas.

Syntax:

- Complex singular terms: if $\phi(x)$ includes free sol variables, bounded fol PLURAL variables, free or bounded fol SINGULAR variables, $\epsilon x.\phi(x)$ is a (complex) singular term;
- Atomic formulae: if a and b are terms, aa is a fol PLURAL term and F is a sol term, a = b, aηaa, Fa are atomic formulae;
- Complex formulae by usual inductive definition.

Axioms of PG:

- Second order Logic;

- "Predicative" comprehension axiom schema for sol variables:
$\exists F \forall x (Fx \leftrightarrow \phi(x))$ -
where $\phi(x)$ includes free sol variables, bound fol PLURAL variables, free or bound fol SINGULAR variables;
*It is *more than predicative*: $\phi(x)$ contain neither bound sol variables nor free fol PLURAL variables.

- Impredicative comprehension axiom schema for fol PLURAL variables:
$\exists xx \forall x (x \eta xx \leftrightarrow \phi(x))$ -
where $\phi(x)$ is a IMPREDICATIVE second-order formula (without xx free);

- schematic Basic Law V: $\hat{}x.\phi(x) = \hat{}x.\psi(x) \leftrightarrow \forall x(\phi(x) \leftrightarrow \psi(x))$
$\rightarrow$ Automatically restricted to "predicative" formulas containing neither bound sol variables nor free fol PLURAL variables - the same that specifies sol variables.
* There is, for every concept (specified sol variable), a corresponding complex singular term and vice versa.

Some differences:

1) Predicativity:

- PE involve a standard definition of predicativity ($\phi(x)$ is predicative if and only if contains no bound second order variables) and use this notion as primitive tool to distinguish the two sorts of second order variables.
- PG involve two primitive sorts of non-singular variables (fol PLURAL and sol variables) - allowing to distinguish two forms of reference - and use this richer vocabulary to characterize a finer grained definition of predicativity ($\phi(x)$ is predicative if and only if contains no bound second order variables and no free first order free variables).

2) Zig zag:

- PE is a classical zig zag theory: some (predicative) concepts are correlated with extensions and other (impredicative) concepts have not correlated extensions.
- PG is less - strictly speaking - zig zag and more fregean theory: every concepts (which is "predicatively" specified) has a correlated extension.

Shared zig zag idea:

- Syntactic agreement about a zig zag feature: $\neg AT$

$\neg\forall X\exists x(x = \hat{}x.\phi(x))$ / $\neg\forall xx\exists x(x = \hat{}x.\phi(x))$ - where $\phi(x)$ is the impredicative formula which specifies X or xx

- Semantic agreement about a zig zag model (M=<D, I>)

D: set of natural numbers = *objectual domain*.
(domain for fol variables of PE and fol SINGULAR variables of PG)

Π(D): power set of D = *conceptual domain*
(domain for IMPREDICATIVE sol variables of PE and fol PLURAL variables of PG)

$\pi(D) \subseteq \Pi(D)$: countable subset of power set of D = *predicative subset of conceptual domain*
(domain for PREDICATIVE sol variables of PE and every sol variables of PG)

I(ϵ): partial (injective) function f: $\pi(D) \rightarrow D$
= *zig zag correlation* from conceptual into objectual domain.

From the (first) *zig zag* solutions to an explanation.

First zig zag proposals:
theories with two versions of comprehension axiom schema (CA), which involve two sorts of second-order variables, a restricted application of the term-forming operator and a correspondent restriction of BLV.

Explanatory suggestion:
The reason why the full second order domain cannot be the domain of the extensionality function is
NOT the specification of its members via CA
the classical assumption that each of them is correlated to an extension.

1. $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \forall x(Xx \leftrightarrow Yx))$ (BLV)
2. $\exists X \forall x(Xx \leftrightarrow \exists Y(x = \epsilon Y \wedge \neg Yx))$. Call this concept R. (CA)
3. $\exists x(x = \epsilon R)$ (2, AT)
4. $\neg R \epsilon R$ (A)
5. $R \epsilon R$ (2,4)
6. $\neg R \epsilon R \rightarrow R \epsilon R$ (4,5)
7. $R \epsilon R$ (A)
8. $\exists Y(\epsilon R = \epsilon Y \wedge \neg Y \epsilon R)$ (2,7)
9. $\neg R \iota R$ (1,8)
10. $R \epsilon R \rightarrow \neg R \epsilon R$ (7,9)
11. $R \epsilon R \leftrightarrow \neg R \epsilon R$ (6,10)

Abstraction's Principle: $@X = @Y \leftrightarrow X \equiv Y$

Abstracts' Theorem: $\forall X \exists x(x = @X)$:

1)

1) t = t	[FOL$^=$]
2) $\forall x(x = x)$	[1, UI]
3) $@X = @X$	[2, UE]
4) $\exists x(x = @X)$	[3, EI]
5) $\forall X \exists x(x = @X)$	[4, UI]

2)

1) $X \equiv X$	[FOL$^=$]
2) $@(X) = @(X)$	[1, AP]
3) $\exists x(x = @(X))$	[2, EI]
4) $\forall X \exists x(x = @(X))$	[3, UI]

Extensionalist explanation

Problematic premise: Abstracts' Theorem $\forall X \exists x(x = \epsilon X)$

which allows to derive,
from Russell's concept $(\exists X \forall x(Xx \leftrightarrow \exists Y(x = \epsilon Y \wedge \neg Yx))$, namely R)
Russell's extension $(\exists x(x = \epsilon R)$, namely r).

Expl.1: AT $[\forall X \exists x(x = @X)] \rightarrow FOL^{=}$ (classical first-order theory of identity and quantification)
Expl.2: (classical logic) assumption that every singular term must be denoting and every function must be total

Argument: classical logic - whose AT is a theorem - by the un-restricted formulation of quantification and identity rules prevents a possible feature of the abstraction, namely that there are some concepts without a correlated abstract.

Carving the correlation. Second option.

I) CARVING CORRELATION INTO THE CORRELATION ITSELF.

Extensionalist idea:
The domain of the extensionality function is a proper subset of the second-order domain.

New zig zag solution:
a restriction of the zig- zag correlation between concepts and extensions obtained by working on the correlation itself

by substituting classical logic with a negative free logic and
by moving the restrictions, traditionally imposed on the comprehension axiom schema, on the right hand of BLV.

Negative free language

Standard second-order language (primitive symbols):

- denumerably many first-order variables: x, y, z;
- denumerably many second-order variables: X, Y, Z;
- logical constants: $\neg, \wedge, \vee, \Sigma, \exists$;
- identity: $=$;
- function symbol: ϵ.

Definable symbols:

- universal "un-restricted" quantifier FOL Π: $\Pi x Ax =_{def} \neg\Sigma x \neg Ax$;
- universal "restricted" quantifier FOL $\forall$: $\forall x Ax =_{def} \neg\exists x \neg Ax$;
- predicative monadic existential predicate E!: E!a $=_{def} \exists x(x = a)$.

Negative free logic (axiomatic version)

Theory:

- impredicative comprehension axiom schema (CA: $\exists X \forall x (Xx \leftrightarrow \phi(x))$);
- standard second-order logic for second-order quantification and standard first-order logic for "unrestricted" first-order quantification;
- free (non inclusive) logic (with identity) for "restricted" quantification and identity:

N1) $\forall v \alpha \rightarrow (E!t \rightarrow \alpha(t/v))$;

N2) $\exists v E!v$;

N3) $s = t \rightarrow (\alpha(s) \rightarrow \alpha(t//s))$;

N4) $\forall v(v = v)$;

N5) $Pt_1, ...t_n \rightarrow E!t_i(with 1 \leq i \leq n)$.

Negative free logic interpretation

Model M = <D, D_0, I>:

- D: domain of "restricted" quantification ($D \subseteq D_0$);
- D_0: domain of "unrestricted" quantification;
- I: total interpretation function on D_0.

I(ϵ): partial injective function from a subset of the powerset of D_0 in D.

No Extensions' Theorem:

Abstracts' theorem (standard/free version):

1) $t = t$	[$FOL^=$]	Not in FL
2) $\forall x(x = x)$	[1, IU]	[N4]
3) $@(X) = @(X)$	[2, EU]	EU Not in FL
4) $\exists x(x = @(X))$		
5) $\forall X \exists x(x = @(X))$	[4, IU sol]	

1) $X \equiv X$	[$FOL^=$]	Not (necessarly) in FL
2) $@(X) = @(X)$	[1, BLVr-I]	
3) $\exists x(x = @(X))$	[2, EG]	
4) $\forall X \exists x(x = @(X))$	[3, UG]	

No Russell's Paradox...

Standard version:

1. $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \forall x(Xx \leftrightarrow Yx))$ (BLV)
2. $\exists X \forall x(Xx \leftrightarrow \exists Y(x = \epsilon Y \wedge \neg Yx))$. Call this concept R. (CA)
3. $\exists x(x = \epsilon R)$ (2, AT)
4. $\neg R \epsilon R$ (A)
5. $R \epsilon R$ (2,4)
6. $\neg R \epsilon R \rightarrow R \epsilon R$ (4,5)
7. $R \epsilon R$ (A)
8. $\exists Y(\epsilon R = \epsilon Y \wedge \neg Y \epsilon R)$ (2,7)
9. $\neg R \iota R$ (1,8)
10. $R \epsilon R \rightarrow \neg R \epsilon R$ (7,9)
11. $R \epsilon R \leftrightarrow \neg R \epsilon R$ (6,10)

... but less logicist result

In the classical framework ($SOL^{=}$ + BLV), the logical part of the theory is involved in the specification of ϵ's domain:

- BLVa: $\forall X \forall Y(\forall x(Xx \leftrightarrow Yx)) \rightarrow \epsilon X = \epsilon Y)$: ϵ is a functional correlation;

- BLVb: $\forall X \forall Y(\epsilon X = \epsilon Y \rightarrow \forall x(Xx \leftrightarrow Yx))$: ϵ is an injective correlation.

- AT $\rightarrow FOL^{=}$: $\forall X \exists x(x = \epsilon X)$: ϵ is a total correlation;

In the free framework (FL + BLV), the abstraction principle defines also the ϵ's domain:

Not only:

- BLVa: $\forall X \forall Y(\Pi x(Xx \leftrightarrow Yx)) \rightarrow \epsilon X = \epsilon Y)$: ϵ is a functional correlation;
- BLVb: $\forall X \forall Y(\epsilon X = \epsilon Y \rightarrow \Pi x(Xx \leftrightarrow Yx))$: ϵ is an injective correlation.

But also:

1. $X \equiv X \rightarrow \epsilon(X) = \epsilon(X)$ (BLVa)
2. $X \equiv X$ (A)
3. $\epsilon(X) = \epsilon(X)$ (MP)
4. $\exists x(x = \epsilon(X))$ (IE)

:

ϵ is a partial function.

... a new circularity

If Russell's concept is reflexively co-extensional with itself, (by BLVa) its extension exists, then the contradiction arises.

$R \approx R$

$\epsilon(R) = \epsilon(R)$

$\exists x(x = \epsilon(R))$

...

$\bot$.

Now - because ϵ can be a partial function - we can conclude that Russell's extension does not exist.

But, if Russell's extension does not exist, Russell's concept (by BLVb) should be not reflexively co-extensional with itself.

Instead, if Russell's extension does not exist, Russell's concept seems to be reflexively co-extensional with itself and then, again, able to introduce its problematic extension.

Other Restrictions

The zig zag solutions used to test the extensionalist explanation have to presuppose not only a logic weakening (from classical to free logic) but also a correspondent weakening of the non logical part of the theory (BLVa).

We will compare three *free zig-zag Fregean systems* (E-FL, P-FL, T-FL) which share the logical axioms (FL) and distinguish one another by the different restrictions admitted on the right hand of BLV:
1) E-BLV:$\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow E!\epsilon X \wedge E!\epsilon Y \wedge \Pi x(Xx \leftrightarrow Yx))$
2) P-BLV:$\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \phi X \wedge \phi Y \wedge \Pi x(Xx \leftrightarrow Yx))$ - where ϕ means *predicative*;
3) T-BLV:$\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \phi X \wedge \phi Y \wedge \Pi x(Xx \leftrightarrow Yx))$ - where ϕ means *positive*.

E-FL

Theory:
FL;
E-BLV:$\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow E!\epsilon X \wedge E!\epsilon Y \wedge \Pi x(Xx \leftrightarrow Yx))$.

We are able to define Frege Arithmetic's vocabulary:
$0 = \sharp([\lambda x.x \neq x]) = \epsilon(\lambda x.\exists X(x = \epsilon(X) \wedge X \approx [\lambda x.x \neq x])$
$1 = \sharp([\lambda x.x = 0]) = \epsilon(\lambda x.\exists X(x = \epsilon(X) \wedge X \approx [\lambda x.x = 0]))$
P(x,y) $= \exists X \exists z(Xz \wedge y = \sharp(X) \wedge x = \sharp([\lambda w.Xw \wedge w \neq z]))$
H(X, R) $= \forall x \forall y(Rxy \rightarrow (Xx \rightarrow Xy))$
R*(xy) $= \forall X((\forall z(Rxz \rightarrow Xz) \wedge Er(X, R)) \rightarrow Xy)$
R"(x,y) = R*$(x, y) \vee x = y$
Nx = P"(0, x)

However E-BLV is too weak to derive the existence of the extensions - namely to derive that the denotations of number terms belong to D. So we are not able to follows Frege's strategy: deriving HP from E-BLV and deriving Peano axioms - as theorems - from HP.

P-FL

Theory:
FL;
P-BLV: $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \phi X \wedge \phi Y \wedge \forall x(Xx \leftrightarrow Yx))$ - where ϕ means *predicative* (the formula that specifies X does not contains bound second ordered variables).

We are not able to define Frege Arithmetic's vocabulary because Frege definitions of cardinal numbers are impredicative ($0 =_{def} \sharp([\lambda x.x \neq x]) = \epsilon(\lambda x.\exists X(x = \epsilon(X) \wedge X \approx [\lambda x.x \neq x]))$.

We directly adopt the second strategy and define the Set Arithmetic's vocabulary:
$0 =_{def} \epsilon(\lambda x.x \neq x)$
$Sn =_{def} \epsilon(\lambda x.x = n)$
$I(X) =_{def} X0 \wedge \forall z(Xz \rightarrow X\epsilon(\lambda x.x = z))$
$\mathbb{N}x =_{def} \forall X(I(X) \rightarrow Xx)$

Theorem

$\mathbb{N}0$.

Proof.

1. $E!0$ (P-BLVa)
2. $\forall X((I(X) \rightarrow X0)$ (def. I(X))
3. $\mathbb{N}0$ (2, def. $\mathbb{N}$)

□

Theorem

$\forall x(Sx \neq 0)$.

Proof.

1. $E!0$	(P-BLVa)
2. $\exists y(E!Sy \wedge y = a)$	(def. S, P-BLVa)
3. a $= 0$	(A)
4. $\epsilon(\lambda x.x = y) = \epsilon(\lambda x.x \neq x)$	(def. S, def. 0)
5. $\forall z([\lambda x.x = y](z) \leftrightarrow [\lambda x.x \neq x](z))$,	(P-BLVb)
6. $\forall z(z = y \leftrightarrow z \neq z)$.	$(\lambda - conv)$
7. $\forall z E!y(z = y \leftrightarrow z \neq z)$	(T1)
8. $y = y \leftrightarrow y \neq y \perp$	(N1)
9. $\neg\exists y(Sy = 0)$	(3,8)
10. $\forall y \neg(Sy = 0)$	(9)
11. $\forall y(Sy \neq 0)$	(10, def. $\neq$)

□

Theorem

$\forall y \forall z(\epsilon(\lambda x.x = y) = \epsilon(\lambda x.x = z) \rightarrow y = z)$.

Proof.

1. $\epsilon(\lambda x.x = y) = \epsilon(\lambda x.x = z)$ (A)
2. $\forall x(\lambda x.x = y)(x) \leftrightarrow (\lambda x.x = z)(x)$ (P-BLVb)
3. $(\lambda x.x = y)(a) \leftrightarrow (\lambda x.x = z)(a)$ (N1, E!a)
4. $a = y \leftrightarrow a = z$ $(\lambda - conv)$
5. $y = z$ (N3)

□

Theorem

$\forall X(X0 \wedge \forall y(Xy \rightarrow X\epsilon(\lambda x.x = y)) \rightarrow \forall x(Xx))$.

Proof.

It follows from the definition of $\mathbb{N}$. □

Theorem

$\forall y \exists x(x = \epsilon([\lambda x.x = y))$

Proof.

1) $\forall y \exists X \forall x(Xx \leftrightarrow x = y)$.
We call this concept $[\lambda x.x = y]$ [AC]
2) $\phi([\lambda x.x = y]) \wedge \Pi x([\lambda x.x = y](x) \leftrightarrow [\lambda x.x = y](x))$
$\rightarrow \epsilon(\lambda x.x = y) = \epsilon(\lambda x.x = y))$ [P-BLVa]
3) $\phi([\lambda x.x = y]) \wedge \Pi x([\lambda x.x = y](x) \leftrightarrow [\lambda x.x = y](x))$ [free $FOL^{=}$]
4) $\epsilon(\lambda x.x = y) = \epsilon(\lambda x.x = y))$ [2,3, MP]
5) $\exists x(x = \epsilon(\lambda x.x = y))$ [4,T2]
6) $\forall y \exists x(x = \epsilon(\lambda x.x = y))$ [1-5, IU]

□

T-FL

Theory:
FL;
T-BLV:$\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow \phi X \wedge \phi Y \wedge \forall x(Xx \leftrightarrow Yx))$ - where ϕ means *positive* (the formula that specifies X contains bound second-order variables only in a even number of negation symbols - considering formulas reduced to its primitive form: $\exists X \exists x(\neg Xx)$ is not positive; $\exists X \exists x(Xx \wedge \neg(x \neq x))$ is positive).

We are able to define Frege Arithmetic's vocabulary because Frege definitions of cardinal numbers are *positive*:

$0 = \sharp([\lambda x.x \neq x]) = \epsilon(\lambda x.\exists X(x = \epsilon(X) \wedge X \approx [\lambda x.x \neq x])$
$1 = \sharp([\lambda x.x = 0]) = \epsilon(\lambda x.\exists X(x = \epsilon(X) \wedge X \approx [\lambda x.x = 0]))$
$\mathrm{P(x,y)} = \exists X \exists z(Xz \wedge y = \sharp(X) \wedge x = \sharp([\lambda w.Xw \wedge w \neq z]))$
$\mathrm{H(X,\ R)} = \forall x \forall y(Rxy \rightarrow (Xx \rightarrow Xy))$
$\mathrm{R^*(xy)} = \forall X((\forall z(Rxz \rightarrow Xz) \wedge Er(X, R)) \rightarrow Xy)$
$\mathrm{R''(x,y)} = \mathrm{R^*}(x, y) \vee x = y$
$\mathrm{Nx} = \mathrm{P''(0,\ x)}$

Definition

$x \in y = \exists X(y = \epsilon(X) \wedge Xx)$ - $(\text{o}\ \forall X(y = \epsilon(X) \rightarrow Xx))$

Definition

$X \approx Y = \exists R(\Pi x(Xx \rightarrow \exists! y(Yy \wedge Rxy)) \wedge \Pi y(Yy \rightarrow \exists! x(Xx \wedge Ryx)))$

Lemma

$\forall X \forall Y(\Pi x(Xx \leftrightarrow Yx) \rightarrow X \approx Y)$

Lemma

$\forall X(X \approx X)$

Lemma

$\forall X \forall Y(X \approx Y \rightarrow Y \approx X)$

Lemma

$\forall X \forall Y \forall Z(X \approx Y \wedge Y \approx Z \rightarrow X \approx X)$

Derivation of HP from T-BLV

Theorem (First theorem about extensions - weak positive version)

$\forall X(\phi(X) \rightarrow \exists x(x = \epsilon(X)).$

Proof.

1. $\phi(X)$ (A)
2. $\Pi x(Xx \leftrightarrow Xx) \wedge \phi(X) \rightarrow \epsilon(X) = \epsilon(X)$ (T-BLVa)
3. $\Pi x(Xx \leftrightarrow Xx)$ (rifl. co-ext)
4. $\epsilon(X) = \epsilon(X)$ (1, 2, 3, MP)
5. $\exists x(x = \epsilon(X))$ (4, T2)

$\square$

Lemma

$\forall X(\exists x(x = \epsilon(X) \to \phi(X)))$.

Proof.

1.	$\exists x(x = \epsilon(X))$	(A)
2.	$\epsilon(X) = \epsilon(X)$	(1, N4)
3.	$\Pi x(\epsilon(X) = \epsilon(X) \to Xx \leftrightarrow Xx \wedge \phi(X))$	(T-BLVb)
4.	$\phi(X)$	(2,3 MP)

□

Theorem

$\forall X(\phi(X) \rightarrow \Pi x(x \in \epsilon(X) \leftrightarrow Xx)).$

Proof.

1. $\phi(F)$	(A)
2. $a \in \epsilon(F)$	(A)
3. $\exists X(\epsilon(F) = \epsilon(H) \wedge Ha)$	(2, def. $\in$)
4. $\epsilon(F) = \epsilon(G) \wedge Ga$	(3, Lemma 2.2, EE)
5. $\Pi x(Fx \leftrightarrow Gx) \wedge \phi(F) \wedge \phi(G)$	(4, T-BLVb)
6. Fa	(4, 5)
7. Fa	(A)
8. $\exists x(x = \epsilon(F))$	(1, 7, Teorema 2.1)
9. $\epsilon(F) = \epsilon(F)$	(8, N4)
10. $\epsilon(F) = \epsilon(F)$	(7, 9, I$\wedge$)
11. $\exists X(\epsilon(F) = \epsilon(X) \wedge Xa)$	(10, IE)
12. $a \in \epsilon(F)$	(11, def. $\in$)
13. $\forall X(\phi(X) \rightarrow \Pi x(x \in \epsilon(X) \leftrightarrow Xx))$	(1, 2-6, 7-12)

□

Lemma

$\forall X \forall Y(\epsilon(Y) \in \sharp(X) \leftrightarrow [\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X)]\epsilon(Y)).$

Proof.

1. $\forall X(\phi(X) \rightarrow \bigvee x(x \in \epsilon(X) \leftrightarrow Xx))$
2. $\forall X(\phi([\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X])))$
3. $\forall X \Pi x(x \in \epsilon([\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X]) \leftrightarrow [\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X]$
4. $\forall X \Pi x(x \in \sharp(X)) \leftrightarrow [\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X]x))$
5. $\forall X(\epsilon(F) \in \sharp(X) \leftrightarrow [\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X)]\epsilon(F))$
6. $\forall X \forall Y(\epsilon(Y) \in \sharp(X) \leftrightarrow [\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx X)]\epsilon(Y))$

□

Lemma

$\forall Y(\phi(Y) \rightarrow \forall X(\epsilon(Y) \in \sharp(X) \leftrightarrow Y \approx X))$.

Proof.

1, $\phi(G)$	(A)
2. $\epsilon(G) \in \sharp(F)$	(A)
3. $[\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx F)](\epsilon(G)$	(2, Lemma 2.4)
4. $\exists Z(\epsilon(G) = \epsilon(Z) \wedge Z \approx F)$	(3, $\lambda - conv.$)
5. $\epsilon(G) = \epsilon H \wedge H \approx F$	(4, Lemma 2.2, EE)
6. $\Pi x(Gx \leftrightarrow Hx) \wedge \phi(G) \wedge \phi(H)$	(5, T-BLVb)
7. $G \approx H$	(6, Lemma 1.1)
8. $G \approx F$	(5, 7 trans. $\approx$)
9. $G \approx F$	(A)
10. $\epsilon(G) = \epsilon(G) \wedge G \approx F$	(1, N4, 9, I$\wedge$)
11. $\exists Z(\epsilon(G) = \epsilon(Z) \wedge Z \approx F)$	(IE)
12. $[\lambda x.\exists Z(x = \epsilon(Z) \wedge Z \approx F)](\epsilon(G)$	(11, $\lambda - conv.$)
13. $\epsilon(G) \in \sharp(F)$	(12, Lemma 2.4)

□

Theorem

$X \approx Y \rightarrow \sharp(X) = \sharp(Y)$

Proof.

1. $F \approx G$
2. $[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]a$
3. $\exists X(a = \epsilon X \wedge X \approx F)$
4. $a = \epsilon(H) \wedge H \approx F$ (3, Lemn
5. $a = \epsilon(H) \wedge H \approx G$ (
6. $\exists X(a = \epsilon X \wedge X \approx G)$
7. $[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]a$
8. $[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]a \rightarrow [\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]a$

□

9. $[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]a$
10. $\exists X(a = \epsilon X \wedge X \approx G)$
11. $a = \epsilon(H) \wedge H \approx G$ (10
12. $a = \epsilon(H) \wedge H \approx F$ (1,
13. $\exists X(a = \epsilon X \wedge X \approx F)$
14. $[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]a$
15. $[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]a \rightarrow [\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]a$

16. $\Pi x(\exists x(x = a) \rightarrow$
$[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]x \leftrightarrow [\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]x$
17. $\Pi x(\neg\exists x(x = a) \rightarrow$
$[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]x \leftrightarrow [\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]x$
18. $\Pi x([\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]x \leftrightarrow [\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G]x$
19 $\forall Z(\phi([\lambda x.\exists X(x = \epsilon(X) \wedge X \approx Z])$
20. $\epsilon([\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]) = \epsilon([\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G])$
21. $\sharp(F) = \sharp(G)$

Theorem

$\phi(X) \wedge \phi(Y) \rightarrow (\sharp(X) = \sharp(Y) \rightarrow X \approx Y)$

Proof.

1. $\phi(F) \wedge \phi(G)$ (A)
2. $\sharp(F) = \sharp(G)$ (A)
3. $\exists x(x = \sharp(F))$ (2, T1)
4. $\exists x(x = \sharp(F))$ (2, T1)
5. $F \approx F$ (rifl. $\approx$)
6. $\epsilon(F) \in \sharp(F)$ (1, 5, Lemma 2.5)
7. $\epsilon(F) \in \sharp(G)$ (5, N3)
8. $F \approx G$ (1, 7, Lemma 2.5)

□

Theorem

$\sharp(X) = \sharp(Y) \to X \approx Y$

Proof.

1. $\sharp(F) = \sharp(G)$
2. $\exists x(x = \sharp(F))$
3. $\exists x(x = \sharp(F))$
4. $\epsilon([[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F]) = \epsilon([[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G])$
5. $\Pi x(([\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F)](x) \leftrightarrow [\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G)](x))$
$\wedge \phi([\lambda x.\exists X(x = \epsilon(X) \wedge X \approx F])) \wedge \phi([[\lambda x.\exists X(x = \epsilon(X) \wedge X \approx G])))$
6. $\Pi x(\exists X(x = \epsilon(X) \wedge X \approx F) \leftrightarrow \exists X(x = \epsilon(X) \wedge X \approx G))$
7. $\Pi x(x = \epsilon(H) \wedge H \approx F \leftrightarrow x = \epsilon(H) \wedge H \approx G)$
8. $\Pi x(H \approx F \leftrightarrow H \approx G)$
9. $F \approx G$

□

Conclusions:

1. Extensionalist explanation partially works: classical identity and quantification axioms are necessary to derive the standard version of Russell's Paradox but the success of zig zag theories also depends on other restrictions
2. The advantage of zig zag theories over other solutions (i.e. predicative theories) depends on identifying that the origin of the paradox does not concern the domain of second-order logic but the domain of the extensionality function

1. An advantage of *free* zig zag theories consists of emphasising the role of classical first-order logic in the standard derivation of Russell's contradiction
2. Another advantage of *free* zig zag theories consists of showing that we can obtain a same result (PA/FA) with a weaker logic theory: usually the solutions that involve restrictions on the abstraction principle presuppose the adoption of a stronger logic theory (plural or modal logic); instead, if we identify the source of the contradiction in the interaction between classical first-order logic and the non-logical part of the theory, we can weaken both these parts of Frege's system

Bibliography:

Antonelli, A and May R. (2005). Frege's Other Program, Notre Dame Journal of Formal Logic, Vol. 46, 1, 1-17.

Boccuni, F. (2010). Plural Grundgesetze, Studia Logica, 96, 2, 3015-330.

Boccuni, F. (2011). On the Consistency of a Plural Theory of Frege's Grundgesetze, Studia Logica, 97, 3, 329-345.

Boolos, G. (1987). Saving Frege from the Contradiction, Proceedings of Aristotelian Society, 87, 137-151.

Burgess, J. (2005). Fixing Frege. Princeton University Press, Princeton.

Cocchiarella, N. B. (1992). Cantor's power-set Theorem versus Frege's double correlation Thesis, History and Philosophy of logic, 13, 179-201.

Dummett, M. (1991). Frege, Philosophy of Mathematics, Oxford University Press, Oxford.

Ferreira, F. (2018). Zigzag and Fregean Arithmetic, in Tahiri, H. (eds), The Philosophers and Mathematics, Springer, Cham, vol. 43, 81-100.

Ferreira, F. and K. F. Wehmeier (2002). On the Consistency of the Fragment of Frege's Grundgesetze. Journal of Philosophical Logic.

Frege, G. (1903). Grundgesetze der Arithmetik, II, Verlag Hermann Pohle.

Heck, R. (1996). The Consistency of Predicative Fragments CA_1^1 of Frege's Grundgesetze der Arithmetik. History and Philosophy of Logic 17, 209–220.

Payne, J. (2013). Abstraction relations need not be reflexive, Thought, 2, 137 - 147.

Paseau, A. C. (2015). Did Frege commit a Cardinal Sin?, Analysis 75 (3), 379–386.

Quine W. V. (1955). On Frege's way out, Mind, 64, 145-159.

Russell, B. (1903). On Some Difficulties in the Theory of Transfinite Numbers and Order Types, in Russell B. (1973). Essays in Analysys, New York.

Uzquiano, G. (forthcoming). Impredicativity and Paradox.

Wehmeier, K. F. (1999). Consistent Fragments of Grundgesetze and the Existence of Non-Logical Objects. Synthese 121(3), 309–328.