Internal Categoricity in Arithmetic and Set Theory

Jouko Väänänen*

Department of Mathematics and Statistics
University of Helsinki
and
Institute for Logic, Language and Computation,
University of Amsterdam

Tong Wang

Institute for Logic, Language and Computation,
University of Amsterdam

April 6, 2012

1 The Set Theory View and the Second Order View

Second order logic was originally considered as an innocuous variant of first order logic in the works of Hilbert. Later study reveals that the analogy with first order logic does not do full justice to second order logic. Quine famously referred to second order logic as "set theory in disguise". Second order logic truly transcends first order logic in terms of strength, and is more appropriate to be compared to (first order) set theory. In second order logic, a large part of set theory becomes essentially logical truth. There is the debate between the "set theory view" and the "second order view" in the foundation of mathematics . The set theory view holds that mathematics is best formalized using first order set theory. The second order view holds that mathematics is best formalized in second order logic.

Two important issues in this debate are completeness and categoricity. It is usually conceived that one merit of the set theory view is that first order logic has a complete proof calculus, while second order logic has not. One merit of the second order view is that second order theories of classical structures (e.g. $\mathbb{N}$, $\mathbb{R}$) are categorical, while first order

*Research partially supported by grant 40734 of the Academy of Finland and by the European Science Foundation EUROCORES LogICCC programme LINT project.

theories allow for non-standard models. More precisely, for a classical structure A, there is a second order sentence θ_A that characterizes A uniquely:

$$A \models \theta_A. \tag{1}$$

$$\forall B \forall C (B \models \theta_A \wedge C \models \theta_A) \to B \cong C. \tag{2}$$

However, a closer inspection shows that the matter is more subtle than a simple trade-off between completeness and categoricity. First of all, the claim that the second order view is inferior to the set theory view because it lacks completeness is unwarranted. It is true that full second order logic does not have a complete proof calculus, but for many reasons it is more reasonable to use *normal second order logic* instead of full second order logic in a foundational quest [5]. Normal second order logic is the extension of the usual logical axioms with the Comprehension Axiom:

CA $\exists X \forall \vec{x}(X\vec{x} \leftrightarrow \phi(\vec{x}))$ for any second order formula ϕ not containing X free.

CA asserts that all definable subsets of the model are in the model. The natural semantics of normal second order logic has as its class of models all normal models, namely Henkin models satisfying all instances of CA [3]. Note that every full model (i.e. model of full second order logic) is normal. Therefore normal second order logic adds to full second order logic the Comprehension Axiom, at the same time its semantics deals with a broader class of models. Importantly, normal second order logic is complete with respect to this extended class of models.

But completeness comes at a price: it leaves the door open for nonstandard models. For example one can, by the usual technique of adding an infinite element, construct a non-standard model of second order arithmetic P^2. Now the obvious opposition disappears: both the set theory view and the second order logic view have a complete underlying logic, and they characterize classical structures to the same level of categoricity [5].

The aim of this paper is to synthesize completeness and categoricity in the second order view. We work within the framework of normal second order logic. We want to restore the idea that second order logic should provide unique characterizations of classical structures. We want something like (1) and (2) to be still true.

Our first innovation is the notion of *internal categoricity.* Internal categoricity is a generalization of the notion of categoricity, proposed by Jouko Väänänen. We say that a theory T is *internally categorical* if all models of T within a common normal model are witnessed isomorphic by the model. We will make this definition more intelligible through examples in what follows. For a detailed account of internal categoricity and the motivation behind it, see [5]. In this paper we prove that second order theories of arithmetic and set theory are internally categorical, although they are not categorical in the classical sense. This fact suggests that nonstandard models and categoricity can exist in harmony. This restores (2).

On the other hand there is the question of consistency. If we take second order logic as a foundation, then the status of (1) is not so clear at first glance. What exactly does it mean that $A \models \theta_A$? It is tempting to say the meaning of $A \models \theta_A$ is given by Tarski's truth definition. However, Tarski's truth definition presupposes that we can read off the truth value of $A \models \theta_A$ in the metatheory—in this case set theory. It would undermine the second order view if the meaning of such basic notions relies on set theory.

Therefore the second task of this paper is to prove the existence of classical structures based on more logical grounds. Suppose T is the second order theory of some classical structure. If we can prove under certain assumptions Γ that there exists a model of T, and that T is internally categorical, then we have, at least to some extent, restored (1) and (2):

$$CA + \Gamma \vdash \exists N (N \models T). \tag{3}$$

$$CA \vdash \forall M \forall N (M \models T \wedge N \models T) \rightarrow M \cong N. \tag{4}$$

In this paper we work out two prime examples of this scheme: arithmetic and set theory. We prove (3) and (4) for P^2 and ZFC^2 respectively. In section 2 we prove internal categoricity of P^2. In section 3 we prove under the assumption that the underlying domain is infinite that there is a model of P^2. In section 4 we prove internal categoricity of ZFC^2. In section 5, the most extensive part of this paper, we prove under certain large domain assumptions that there exists a model of ZFC^2.

2 Internal Categoricity of Arithmetic

The axiom system P^2 is the second order version of Peano Arithmetic [2]. For the purpose of this paper, we consider the relativized version of P^2. Let $L = (N, S, 0)$ be the language of arithmetic. Intuitively N denotes the underlying domain. The axioms of P^2 consist of:

P0 $\forall x (x \in N \rightarrow Sx \in N)$.

P1 $\forall x \in N \neg Sx = 0$.

P2 $\forall x \in N \forall y \in N (Sx = Sy \rightarrow x = y)$.

P3 $\forall X \subset N ((X0 \wedge \forall x \in N (Xx \rightarrow XSx)) \rightarrow X = N)$.

Note that strictly speaking "$x \in N$" and "$X \subset N$" are not part of our language: we use them as abbreviations for Nx and $\forall x (Xx \rightarrow Nx)$. We will use these notations freely whenever ambiguity does not arise.

It is well-known that P^2 characterizes $\mathbb{N}$ up to isomorphism in full second order logic. When it comes to normal second order logic, this is no longer the case. A counterexample can be provided by an application of the completeness theorem. Expand the language with a new constant symbol c, and let Σ be the theory $P^2 \cup \{c > S^n 0 : n \in \mathcal{N}\}$. Clearly Σ is finitely satisfiable. Since normal second order logic has a complete proof system, it satisfies the compactness property. Hence Σ has a model $(M, \mathcal{G})$, which is a non-standard model of P^2. Note that $(M, \mathcal{G})$ is a normal model, but necessarily not a full one: in particular the standard part is not in $\mathcal{G}$, for otherwise it would contradict the induction clause P3.

Now we investigate the notion of internal categoricity. Let $L = (N, S, 0, N', S', 0')$ be two copies of the language of arithmetic. An essential feature of P^2 is that it replaces the induction schema in P by a second order quantification over subsets. This enables P^2 to be a finitely axiomatizable theory. Consequently $(N, S, 0) \models P^2$ can be written as a sentence in the formal language, denote it by $P^2(N, S, 0)$. Similarly $P^2(N', S', 0')$ has the intuitive meaning that $(N', S', 0') \models P^2$. That R is an isomorphism from $(N', S', 0')$ to $(N', S', 0')$ can also be written as a second order sentence $ISO(R, N, S, 0, N', S', 0')$, we spare ourselves with the details here. We say that P^2 is *internally categorical* if whenever a normal model contains two copies of P^2, i.e. $(M, \mathcal{G}) \models P^2(N, S, 0) \wedge P^2(N', S', 0')$, this model "sees" that these two copies are isomorphic, i.e. $(M, \mathcal{G}) \models \exists R\ ISO(R, N, S, 0, N', S', 0')$. We prove that P^2 is internal categorical under this definition.

Theorem 1. *Let $L = \{N, S, 0, N', S', 0'\}$ be two copies of the language of arithmetic. Consider the Comprehension Axiom in the language L. Then*

$$CA \vdash (P^2(N, S, 0) \wedge P^2(N', S', 0')) \rightarrow \exists R\ ISO(R, N, S, 0, N', S', 0').$$

Proof. Suppose $(M, \mathcal{G}) \models CA$, and that:

1. $(M, \mathcal{G}) \models P^2(N, S, 0)$,
2. $(M, \mathcal{G}) \models P^2(N', S', 0')$.
3. $N, S, N', S' \in \mathcal{G}$.

We want to show that there is an $R \in \mathcal{G}$ such that $R : (N, S, 0) \cong (N', S', 0')$. Let $R = \cap\{P \in \mathcal{G} : P00' \wedge \forall x \in N \forall y \in N'(Pxy \rightarrow PSxSy)\}$. Note that R is in fact a definable subset of M. For any $c, d \in M$,

$$Rcd \leftrightarrow \forall P((P00' \wedge \forall x \in N \forall y \in N'(Pxy \rightarrow PSxSy)) \rightarrow Pcd).$$

By the Comprehension Axiom, $R \in \mathcal{G}$.

It is easy to verify that $R00'$, and that $\forall x \in N \forall y \in N'(Rxy \rightarrow RSxSy)$. From these we prove that R is an isomorphism from $(N, S, 0)$ to $(N', S', 0')$.

(i) Totality. By the definition of R, $0 \in dom(R)$ and $\forall x \in N(x \in dom(R) \rightarrow Sx \in dom(R))$. Hence by $P3$, $dom(R) = N$.

(ii) Functionality. Let $X = \{x \in N : \exists! y\ Rxy\}$. We prove $X = N$ by induction. For the base case, suppose $R00'$ and $R0a$ for $a \neq 0'$. Now we define $R' = R - \{a, b\}$. Note that R' also satisfies $R'00' \wedge \forall x \in N \forall y \in N'(R'xy \rightarrow R'SxSy)$, contradicting the minimality of R. The induction case is similar in nature.

(iii) Surjectivity. Dual to totality.
(iv) Injectivity. Dual to functionality.
(v) Homomorphism. It follows directly from the definition of R.

□

For an investigation of the *reverse-mathematics* status of second-order categoricity of arithmetic, see [4].

3 Consistency of Arithmetic

Our next goal is the consistency of P^2. CA alone cannot prove that there is a model of arithmetic: in particular, all finite models are models of CA, but they cannot be models of P^2. Therefore we make the additional assumption that the underlying domain is infinite. More precisely, we assume that the model contains a non-surjective injective mapping F.

Theorem 2. *Let (M, G) be a normal model. Suppose $(M, G) \models \exists F \exists z(\forall x \forall y(Fx = Fy \rightarrow x = y) \wedge z \notin ran(F))$. Then*

$$(M, G) \models \exists N \exists S \exists a \; P^2(N, S, a).$$

Proof. Pick $F \in \mathcal{G}$, $a \in M$ such that F is injective and a is not in the range of F. We take the closure of a under F. Let

$$N = \cap\{X \subset M : Xa \wedge \forall x(Xx \rightarrow XFx)\}.$$

By the same reasoning as in the previous proof, N is definable, and the Comprehension Axiom makes sure that $N \in \mathcal{G}$. We claim that $(N, F, a) \models P^2$.

(i) F is injective on N. This is clear.
(ii) $a \notin ran(F)$. This is also clear.
(iii) The induction clause. Suppose $X \subset N$, $a \in X$ and X is closed under F. If X is a proper subset of N, it would contradict the minimality of N. Therefore $X = N$.

□

4 Internal Categoricity of Set Theory

In the same fashion as P^2, we have the second order counterpart of the ZFC axioms of set theory. Consider the second order language of set theory consisting of a single non-logical symbol $\in$. ZFC^2 has the same axioms for Extensionality, Union, Pairing, Power Set, Infinity, Regularity and Choice as ZFC. As for Separation and Replacement, ZFC^2 replaces the axiom schemata with second order quantifications. Separation now reads:

Sep $\forall X \forall x \exists y(\forall z(z \in y \leftrightarrow (z \in x \wedge Xz)))$.

And similarly for Replacement. Note that ZFC^2 is a finite set of axioms.

It is not reasonable to require internal categoricity of ZFC^2 straightaway. Zermelo proved that the natural models of ZFC^2 are exactly the V_κ's for κ an inaccessible cardinal [6]. For two different inaccessibles κ and λ, V_κ and V_λ are not isomorphic. However, if we assume that two models of ZFC^2 are "of the same height", i.e. there is an isomorphism between ordinals in V_κ and ordinals in V_λ, we can prove that they are internally isomorphic.

Theorem 3. *Let $L = (V, E, V', E')$ be two copies of the language of set theory. Let $(M, \mathcal{G})$ be a normal model. Suppose*

1. $(M, \mathcal{G}) \models ZFC^2(V, E)$,
2. $(M, \mathcal{G}) \models ZFC^2(V', E')$,
3. $V, V', E, E' \in \mathcal{G}$,
4. $(M, \mathcal{G}) \models \exists\pi\, ISO(\pi, Ord, Ord')$. *$Ord$ and Ord' denote ordinals in (V, E) and (V', E') respectively.*

Then

$$(M, \mathcal{G}) \models \exists R\, ISO(R, V, E, V', E').$$

Proof. Pick an isomorphism $\pi \in \mathcal{G} : Ord \cong Ord'$. We define an isomorphism between V and V' by a back-and-forth clause. Let

$$R = \cap\{P \in \mathcal{G} : \pi \subset P, \forall x \in V \forall y \in V'((\forall zEx\exists uE'yPzu \wedge \forall uE'y\exists zExPzu) \to Pxy)\}.$$

We may think of R as the minimal extension of π respecting E and E'. By construction R is in $\mathcal{G}$. Moreover R satisfies:

(a) $\pi \subset R$

(b) $\forall x \in V \forall y \in V'((\forall zEx\exists uE'yRzu \wedge \forall uE'y\exists zExRzu) \leftrightarrow Rxy)$.

The forward direction in (b) follows from the definition of R. The converse follows from the minimality of R. The property (b) gives us a general criterion to decide whether Rxy. Now we proceed to prove that R is an isomorphism in three steps.

(i) The relation R, when defined, is an isomorphism onto its image.

First we prove that R is functional. Let $x \in V$ be an E-minimal element such that there are $y, y' \in V'$ with $y \neq y', Rxy$ and Rxy'. By Extensionality, without lost of generality there is $u \in V'$ such that $uE'y$, $\neg uE'y'$. Since we have Rxy, by property (b) above there is zEx such that Rzu. Similarly, there is $u'E'y'$ such that Rzu'. Since $\neg uE'y$, $u \neq u'$. Now we have Rzu and Rzu' for zEx, contradicting the minimality of x.

By exchanging the role of V and V' in the above argument we can prove that R is injective. That R respects the relations E and E' is clear. Hence R is an isomorphism onto its image when it is defined.

(ii) For each $\alpha \in Ord$, $R : V_\alpha \to V'_{\pi(\alpha)}$ is an isomorphism.

For the base case $R : \varnothing \to \varnothing'$ this is trivially true.

For the successor case, suppose $R : V_\alpha \to V'_{\pi(\alpha)}$ is an isomorphism, we aim to show that so is $R : V_{\alpha+1} \to V'_{\pi(\alpha+1)}$. Thanks to (i) it suffices to prove that R is defined on $V_{\alpha+1}$, and that $R \upharpoonright V_{\alpha+1}$ is onto $V'_{\pi(\alpha+1)}$. Pick an arbitrary $y \in \mathcal{P}(V_\alpha)$, i.e. $Ext(y) \subset V_\alpha$. By induction hypothesis $R(Ext(y)) \subset V_{\pi(\alpha)}$. Let $zE\mathcal{P}(V'_{\pi(\alpha)})$ be such that $Ext(z) = R(Ext(y))$. It is straightforward to check by property (b) that Ryz. Since y is arbitrary, R is defined on $V_{\alpha+1}$. Symmetrically we can prove that $R \upharpoonright V_{\alpha+1}$ is onto $V'_{\pi(\alpha+1)}$.

The limit case is straightforward.

(iii) The relation $R : V \to V'$ is an isomorphism.

Step (ii) implies that $R : V \to V'$ is an embedding. Since π is an isomorphism between the ordinals in V and in V', the dual argument using π^{-1} shows that R^{-1} is also an embedding. Therefore R is an isomorphism from V to V'. □

5 Consistency of Set Theory

In this section we seek to establish the consistency of ZFC^2 on certain large domain assumptions. The key ingredient is a power set operation which generates the set theoretic structure by iteration. Let $(M, \mathcal{G})$ be a normal model. To cope with the relevant set theoretical terminology, in this section we refer to elements x of M as "sets" and subsets X of M as "classes". When x is an element of M and E a binary relation, we denote by $Ext(x) = \{y \in M : yEx\}$ the extension of x.

Let $\mu(X)$ be the formula saying that X is of smaller cardinality than the universe, or in brief "X is small" :

$$\mu(X) =_{def} \neg\exists F(\text{“}F \text{ is injective”} \wedge \forall x\ XFx).$$

Let $\eta(X, Y, E)$ be the formula saying that Y behaves like the power set of X, with E taken to be the intended membership relation:

$$\eta(X, Y, E) =_{def} \forall x \forall y (xEy \to Xx \wedge Yy)$$

$$\wedge \forall x, y \in Y ((\forall z \in X\ zEx \leftrightarrow zEy) \to x = y)$$

$$\wedge \forall Z \subset X \exists y \in Y \forall z (Zz \leftrightarrow zEy).$$

Then the following sentence says that every small class has a power set:

$$\forall X(\mu(X) \to \exists Y, E\ (\mu(Y) \wedge \eta(X, Y, E))). \tag{5}$$

We can also express that the cardinality of the universe is inaccessible, namely the union of a family of small sets indexed by a small set is always small.

$$\forall X \forall R(\mu(X) \wedge \forall x \in X \mu(R(x, -)) \to \mu(R(X)). \tag{6}$$

Here $R(x,-)$ denotes the image of x under R, $R(X)$ denotes the image of the class X under R. Note that $R(X) = \cup_{x \in X} R(x,-)$.

As mentioned before, the natural models of ZFC^2 are the V_κ's for κ an inaccessible cardinal. For such κ we have that $|V_\kappa| = \kappa$. (5) implies that the cardinality of the universe is a strong limit cardinal, (6) implies that the cardinality of the universe is a regular cardinal. Quite naturally, our first guess is that (5) and (6) together imply that there is a model of ZFC^2:

$$(5) + (6) \vdash \exists M, E\ ZFC^2(M, E).$$

Tempting as it is, a moment's reflection shows that this plan does not work. It is not enough to postulate that there is a power set for each small set, but we also need these power sets to be compatible with each other, in order to glue them together and generate the set theoretic structure. For example, suppose $X \subset Y$, ideally we should have that $\mathcal{P}(X) \subset \mathcal{P}(Y)$, yet this is not entailed by (5) and (6). In order to remedy this defect, we will make a compatibility assumption on the power set operation. The price we pay is that the resulting postulate amounts to assuming the existence of a third order object. However, the moral we draw from Gödel's Incompleteness Theorems is that there are always constraints in proving consistency. We have to content ourselves with this situation at the moment.

5.1 The Postulates

Suppose $(M, \mathcal{G})$ is a normal model. We make the following assumptions:

(a) There is a function (P, E) defined on small subsets of M, such that for each small class X it associates another small class $P(X)$ and a relation E_X

$$X \longmapsto P(X), E_X$$

such that

- i $\eta(X, P(X), E_X)$, η is as defined in the previous section;
- ii $\forall X \forall Y (X \subset Y \to P(X) \subset P(Y))$;
- iii $\forall X \forall Y (X \subset Y \to \forall y \in P(X) \{u \in X : u E_X y\} = \{u \in Y : u E_Y y\})$.

(b) The cardinality of the universe is regular, i.e. $\forall X \forall R (\mu(X) \wedge \forall x \in X \mu(R(x,-)) \to \mu(R(X))$.

(c) Each E_X is well-founded.

(d) There is a transitive class X such that it is infinite, and it is small relative to the universe.

We also assume the Axiom of Choice in the metatheory. In other words we always assume the existence of Skolem functions.

AC $\forall R(\forall x \exists y Rxy \to \exists F \forall x Rx Fx)$.

A few words on postulate (a). The requirements (i) and (ii) are quite natural. As for (iii), it expresses a strong compatibility condition. The E_X's provide us with "local fragments" of the intended membership relation, and our goal is to patch them into a global membership relation. For any X small and $y \in P(X)$, define *the extension of* y *in* X to be $Ext_X(y) = \{u \in X : uE_X y\}$. Requirement (iii) says that whenever Y is a supset of X and y is a member of $P(X)$ (and hence by (ii) a member of P(Y)), $Ext_Y(y)$ is equal to $Ext_X(x)$. Upon reflection this is indeed what it is like in the real set theoretic universe. For any set x there corresponds a power set $\mathcal{P}(x)$, and this $\mathcal{P}(x)$ will not change as we regard x as the subset of varying supsets.

Postulate (d) is thus formulated so that the model we construct contains an infinite object. Without this assumption, the structure we end up with might very well be something like V_ω, satisfying every axiom of ZFC^2 except the Axiom of Infinity. Transitivity here is rather a technical assumption in order to guarantee that the addition of an infinite object does not spoil other constructions. We will see its use in the proofs below. At the moment it is not clear how transitivity can be expressed in our language. In order not to distract the reader from the storyline, we postpone the treatment of this point to the next section.

The reader might have sensed where we are going. We construct a model of ZFC^2 in much the same way we did for P^2. Pick a witness C for postulate (d). Define V to be the closure of C under the power set operation P.

$$V = \cap\{N \in \mathcal{G} : C \subset N, \forall X \subset N(\mu(X) \to P(X) \subset N)\}.$$

Define the binary relation E to be the union of the local relations.

$$E = \cup_{X \subset V, \mu(X)} E_X.$$

The Comprehension Axiom implies that $V, E \in \mathcal{G}$.

Before concluding this section we make an important remark. The Comprehension Axiom depends on the language we use. In order for our postulates to achieve its power in full, we have to allow P and E to appear in instances of CA. They become part of our language.

5.2 The Axioms of ZFC^2

Now we set out to prove that $(V, E) \models ZFC^2$. In this section we mean by a formula its relativization to V unless otherwise specified.

First several lemmata. By the definition of E as the union of the E_X's, we have that $Ext(y) = \cup_X Ext_X(y)$. The following lemma shows that it suffices to consider one such X.

Lemma 1. *For any y and any X such that $y \in P(X)$, $Ext(y) = Ext_X(y)$. Equivalently, for any u, uEy if and only if uE_Xy.*

Proof. By definition uE_Xy implies uEy. Now we prove the other direction. Suppose uEy, then uE'_Xy for some X'. Suppose also that u is not in X. Consider the set $X \cup X'$. By postulate (b) on regularity, the union of two small sets is small, hence P is defined on $X \cup X'$. By postulate (a-iii) we have $Ext_X(y) = Ext_{X \cup X'}(y) = Ext_{X'}(y)$. But this is a contradiction, since $u \in Ext_{X'}(y)$ yet $u \notin Ext_X(y)$. □

Lemma 2 (Comprehension)**.** $(V, E) \models \forall X(\mu(X) \to \exists y\ Ext(y) = X)$.

Proof. Take $y \in P(X)$ such that $Ext_X(y) = X$, this can be done because of (a-i). By Lemma 1, $X = Ext_X(y) = Ext(y)$. □

Henceforth we refer to this lemma as the comprehension lemma (not to be confused with the Comprehension Axiom). Together with CA, this lemma tells us that for any class X, if it is definable and is small, then there is a set x with extension X. This will be our key apparatus in proving existential claims in the ZFC^2 axioms.

Lemma 3 (Every set is small)**.** $(V, E) \models \forall x\mu(Ext(x))$.

Proof. We distinguish between two cases. First suppose $x \notin C$. Recall C is a fixed witness for postulate (d). Suppose $Ext(x)$ is not small. Then $x \notin P(X)$ for all small X, for otherwise $Ext(x) = Ext_X(x) \subset X$ and hence is small. Now we consider the model $V' = V - \{x\}$. V' is also closed under P, since x is not in any $P(X)$. Moreover V' still contains C for $x \notin C$. This contradicts the construction of V as the minimal such class. On the other hand if $x \in C$, by the transitivity of C we have that $Ext(x) \subset C$, and hence $Ext(x)$ is small. □

Now we can start verifying the axioms of set theory. Let's begin with some simpler ones.

Theorem 4 (Extensionality)**.** $(V, E) \models \forall x \forall y(\forall z(zEx \leftrightarrow zEy) \to x = y)$.

Proof. Suppose $Ext(x) = Ext(y) = Z$. By Lemma 1, $Ext(x) = Ext_X(x)$ for some small X, and $Ext(y) = Ext_Y(y)$ for some small Y. Consider the set $X \cup Y$. We have that $Ext_{X \cup Y}(x) = Ext_X(x) = Ext(x) = Ext(y) = Ext_Y(y) = Ext_{X \cup Y}(y)$. We know that Extensionality holds locally by (a-i), hence $x = y$. □

Theorem 5 (Power Set)**.** $(V, E) \models \forall x \exists y \forall z(zEy \leftrightarrow z \subset x)$.

There is a bit of abuse of notation here. Previously by the symbol $\subset$ we refer to the subset relation in the metatheory, while here it refers to the subset relation derived from E. For the sake of readability, we stick to the usual notation instead of inventing another symbol.

Proof. The power set axiom amounts to $\exists y(\forall z(zEy \leftrightarrow Xz))$, where X is defined by $\forall z(Xz \leftrightarrow \forall w(wEz \to wEx))$. For any $z \in X$, $Ext(z) \subset Ext(x)$, therefore $z \in P(Ext(x))$. Since z is arbitrary, we have that $X \subset P(Ext(x))$, and hence X is small. By the comprehension lemma there is a y such that $Ext(y) = X$. □

The Axioms of Paring, Separation and Replacement follow from the comprehension lemma. In each case, it is rather straightforward to verify that the set we desire is small.

Theorem 6 (Paring)**.** $(V, E) \models \forall x \forall y \exists z \forall w(wEz \leftrightarrow w = x \vee w = y)$.

Theorem 7 (Seperation)**.** $(V, E) \models \forall X \forall x \exists y(\forall z(zEy \leftrightarrow (zEx \wedge Xz)))$.

Theorem 8 (Replacement)**.** *If a class F is a function, then for every set x, $F(x)$ is a set.*

Proof. Suppose x is a set and F a functional class. Let $Y = \{y \in V : \exists zEx\ Fz = y\}$. By the Axiom of Choice there is a functional class G that associates to each $y \in Y$ one of its pre-images, i.e. $F(Gy) = y$. By functionality of F, G is an injective mapping from Y into $Ext(x)$. Since $Ext(x)$ is small, Y is also small, therefore there is a set y with extension Y by the comprehension lemma. This is the set we want.

□

The Axiom of Union requires the postulate on regularity.

Theorem 9 (Union)**.** $(V, E) \models \forall x \exists y \forall z(zEx \leftrightarrow \exists wEx\ zEw)$.

Proof. By the regularity postulate (b) the class $\cup_{wEx} Ext(w)$ is small, and hence it is a set by the comprehension lemma. □

The proof of the Axiom of Regularity is based on postulate (c).

Theorem 10 (Regularity)**.** $(V, E) \models$ *"Every set has an E-minimal element".*

Proof. Take an arbitrary element x, suppose $Ext(x)$ does not have a minimal element. Then there is an infinite descending chain $...x_n E x_{n-1} .. x_2 E x_1 E x_0$ in $Ext(x)$. Denote $Z = \{x_0, x_1, ..., x_n, ...\}$. Now for each x_n in Z $(n > 0)$ there is a small set Y_{x_n} such that $x_n E_{Y_{x_n}} x_{n-1}$. Take $Y = \cup_{z \in Z} Y_z$. Z is a subset of Ext(x) and hence is small, therefore by regularity Y is also small. Moreover $E_{x_n} \subset E_Y$ for all $x_n \in Z$. Then the x_n's has become an infinite descending chain in the relation E_Y, contradicting postulate (c). Note we have tacitly used the Axiom of Choice in the metatheory. □

Proving the Axiom of Infinite requires more work. But before that, we have to fulfil our promise and show that transitivity is indeed expressible in our language. Let the following sentence say a class X is transitive:

$$\phi(X) =_{def} \forall x(Xx \to \forall y(\exists Y y E_Y x \to Xy)).$$

$\phi(X)$ is equivalent to a universal formula:

$$\forall x \forall y \forall Y (Xx \land yE_Y x \to Xy).$$

Consider the witness C for postulate (d). Postulate (d) asserts that $(M, \mathcal{G})$ sees C as transitive:

$$(M, \mathcal{G}) \models \phi(C).$$

Since $\phi(X)$ is equivalent to a universal formula, it is preserved to the submodel V of M. Therefore V also sees C as transitive:

$$(V, E) \models \phi(C). \tag{7}$$

Now we can tackle the Axiom of Infinity. As mentioned before, postulate (b) makes sure that ω is still considered as a "small" stage in the construction, so that the model we get does not stop at V_ω but ends somewhere beyond.

Theorem 11 (Infinity)**.** $(V, E) \models$ *"There exists an infinite set"*.

Proof. Since C is infinite, there is an injective class function F from C to its proper subset. Define a class X as follows:

$$\forall x \ (Xx \leftrightarrow \exists y \exists z (x = (y, z) \land Fy = z)).$$

Here (y, x) denotes the ordered pair of y and z. X has cardinality less or equal to $C \times C$, and hence is small by regularity. Let f be the set comprehended by X, and c the set comprehended by C. Then f is an injective function (not in the metatheory but in the intended set theory of (V, E)) from c to its proper subset, therefore c is an infinite set in V.

□

Finally, the Axiom of Choice. But haven't we already assumed the Axiom of Choice? We have assumed AC in the metatheory. Now what we will prove is the internalized AC in the model (V, E). This is similar in spirit to what we just did with the Axiom of Infinity: we assume the existence of an infinite object in the metatheory, the proof consists of internalizing the infinite object into (V, E).

Theorem 12 (Choice)**.** $(V, E) \models$ *"For any set x, if the empty set is not a member of x, then there is a choice function on x"*.

Proof. Suppose $\emptyset \notin x$. Then for each $y \in Ext(x)$, there is $z \in Ext(y)$. By AC in the metatheory, there is a class function F such that for all $y \in Ext(x)$, $Fy \in Ext(y)$, i.e. $FyEy$. Similar to the proof of the Axiom of Infinity, let f be the set comprehended by F. This f is a choice function on x. □

Here are all the axioms. In retrospect, what we have done is very similar in spirit to what Burgess does in [1]. In that paper Burgess proves ZFC^2 under the following two assumptions:

(a) There are just as many individuals as small classes.

(b) There are indescribably many individuals.

Assumption (a) is essentially what we have achieved with Lemma 2 and Lemma 3: to each small class there corresponds a set (i.e. an individual), and the extension of each set is a small class. The connection between (b) and our approach remains a topic for further investigation.

References

[1] John Burgess. Axioms of Infinity as the Starting Point for Rigorous Mathematics. *Annals of the Japan Association for Philosophy of Science* special issue on Infinity in Mathematics and Philosophy, (to appear).

[2] Richard Dedekind. What are the meaning and the purpose of numbers? (was sind und was sollen die zahlen?). *Braunschweig*. Vieweg & Sohn, 1888.

[3] Leon Henkin. Completeness in the theory of types. *J. Symbolic Logic*, 15:81–91, 1950.

[4] Stephen Simpson and Keita Yokoyama. On the reverse mathematics of Peano categoricity (unpublished).

[5] Jouko Väänänen. Second-order logic or set theory? *Bull. Symbolic Logic*, 18(1), 91-121, 2012.

[6] Ernst Zermelo. Über Grenzzahlen und Mengenbereiche. Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae*, 16:29–47, 1930.