A reduction of the NF consistency problem

Athanassios Tzouvaras*

Aristotle University of Thessaloniki,
Department of Mathematics,
541 24 Thessaloniki, Greece.
e-mail:tzouvara@math.auth.gr

Abstract

We give a necessary and sufficient condition in order that a type-shifting automorphism be constructed on a model of the Theory of Simple Types (TST) by forcing. Namely it is proved that, if for every $n \geq 1$ there is a model of TST in the ground model M of ZFC that contains an n-extendible coherent pair, then there is a generic extension $M[G]$ of M that contains a model of TST with a type-shifting automorphism, and hence $M[G]$ contains a model of NF. The converse holds trivially. It is also proved that there exist models of TST containing 1-extendible coherent pairs.

Keywords. New Foundations, type-shifting automorphism, forcing, coherent pair, n-extendible coherent pair.

1 Introduction and preliminaries

The major open problem of the theory NF (New Foundations) is its consistency (relative to that of ZFC). Important reductions of the problem were obtained by E. Specker (who reduced it to the consistency of TST+(Amb)), and by V.N. Grishin (who reduced NF to its fragment NF_4). In this paper we use both of the aforementioned reductions and give a new one using the

*The project is co-funded by the European Social Fund and National Resources (EPEAEK II) PYTHAGORAS II.

language of forcing. Actually the reduction emerged out from an attempt to construct a type-shifting automorphism on a model of TST_4 (the fragment of the theory of types TST up to the fourth level) by forcing. The main implication reads as follows: "If for each $n \geq 1$ there is a model of TST in the ground model M of ZFC that contains an n-extendible coherent pair, then there is a generic extension $M[G]$ of M that contains a model of TST with a type-shifting automorphism, and hence $M[G]$ contains a model of NF". The converse holds trivially: If M contains a model of NF, then it contains also a model of TST having n-extendible coherent pairs, for every $n \geq 1$. Coherent pairs are, roughly, finite approximations of a type-shifting automorphism of a model of TST_4.

Section 1 contains preliminary material. Most of it is well-known, but some facts, though elementary, seem to have been fixed here for the first time. For example the definition 1.1 of level collapse of a model, and lemmas 1.2 and 1.6, which essentially allow one to restrict oneself to standard transitive models of TST. Also lemma 1.9 is just an adaptation of a result of Grishin.

In section 2 we define the notion of n-extendible coherent pair and, by the help of a certain extension of the theory TST, we prove the main theorem 2.8 (and its equivalent version theorem 2.9).

In section 3 we prove that there are models of TST containing 1-extendible (i.e., just extendible) coherent pairs (theorem 3.6).

Section 4, finally, contains some comments on 2-extendible coherent pairs.

1.1 The theory of types and its models

Our metatheory will be ZFC. We often refer to it either as "external world" or as "ground model". $\in$ is the membership relation of the ground model. $\mathbb{N}$ is the set of natural numbers of the ground model.

The language L_{TST} of the Theory of Simple Types (TST) consists of the binary predicate symbol ε and countably many sorts (or types) $S_i(x)$, $i \in \mathbb{N}$. Usually we introduce typed variables x^i, y^j, etc., $i, j \in \mathbb{N}$, where the superscript indicates the type, and x^i stands for $S_i(x)$. Formulas of L_{TST} are built from atomic formulas of the form $x^i \varepsilon x^{i+1}$ and $x^i = y^i$, in the ordinary way. The axioms of TST are the following schemes of comprehension and extensionality:

(Co) $(\exists x^{i+1})(\forall y^i)(y^i \varepsilon x^{i+1} \Leftrightarrow \phi(y^i))$, for every $\phi(y^i) \in L_{\mathrm{TST}}$ possibly with extra free variables.

(Ex) $(\forall x^i)(x^i \varepsilon y^{i+1} \Leftrightarrow x^i \varepsilon z^{i+1}) \;\Rightarrow\; y^{i+1} = z^{i+1}$.

A model of TST is a sequence $\mathcal{A} = (A, A_0, A_1, \ldots, R)$, such that $A = \bigcup_{i\in\mathbb{N}} A_i$, and each A_i interprets the variables of type i. R is a binary relation on A, i.e., $R \subseteq A^2$, that interprets ε. Since $A = \bigcup_{i\in\mathbb{N}} A_i$, we may just write $\mathcal{A} = (A_0, A_1, \ldots, R)$ instead of $\mathcal{A} = (A, A_0, A_1, \ldots, R)$. Note that the axioms above say nothing as to whether the levels A_i of a model are disjoint or not, or whether there are x, y belonging to the same level such that xRy. So R may well be non-well-founded.

For every $i \geq 0$ and $x \in A_{i+1}$ let

$$x_R = \{y \in A_i : yRx\}.$$

$\mathcal{A}$ is said to be *standard* if R is $\in$ (restricted to the sets of the model). Equivalently $\mathcal{A}$ is standard if for every $x \in A_{i+1}$,

$$x_R = x \cap A_i.$$

A standard $\mathcal{A}$ is *transitive* if for every i, $x \in A_{i+1} \Rightarrow x \subseteq A_i$, i.e., if for every $x \in A_{i+1}$,

$$x_R = x.$$

In such a case for all i, $A_{i+1} \subseteq \mathcal{P}(A_i)$.

Standard transitive (henceforth s.t.) models of TST are the most natural and intuitively graspable ones. E.g. for every $X \neq \emptyset$, the sequence $(X, \mathcal{P}(X), \mathcal{P}^2(X), \ldots, \in)$ is a s.t. model of TST. Such a model is called *full* and is denoted by $\langle\langle X\rangle\rangle$. If X is infinite $\langle\langle X\rangle\rangle$ is uncountable. To find a countable model we can take a countable elementary submodel of $\langle\langle X\rangle\rangle$. Such a model is standard but not transitive.

Definition 1.1 Given a model $\mathcal{A} = (A_0, A_1, \ldots, R)$ of TST, define the mapping $\sigma : \bigcup_i A_i \to V$ by induction on i as follows:

$\sigma(a) = a$ for all $a \in A_0$,

$\sigma(x) = \{\sigma(y) : y \in A_i \;\&\; yRx\} = \{\sigma(y) : y \in x_R\}$, for $x \in A_{i+1}$.

Let $B_i = \sigma'' A_i$[1] and $\mathcal{B} = (B_0, B_1, \ldots, \in)$. $\mathcal{B}$ is said to be the *level collapse* of $\mathcal{A}$ and is denoted by $lc(\mathcal{A})$.

[1]Given a mapping $h : X \to Y$, h'' denotes the induced mapping from $\mathcal{P}(X)$ to $\mathcal{P}(Y)$, where $h''x = \{h(y) : y \in x\}$ (T. Forster [1] uses the notation $j`h$ for h''). If h is 1-1 or onto, then so is h''.

Obviously $lc(\mathcal{A})$ is a standard and transitive structure, defined independently of whether R is well-founded or not[2]. In particular, if $\mathcal{A}$ is standard, then for $x \in A_{i+1}$, $\sigma(x) = \{\sigma(y) : y \in x \cap A_i\}$. Since $\mathcal{A} \models \mathrm{Ex}$, it is easy to see that for every i, $\sigma{\restriction}A_i$ is 1-1. But σ need not be 1-1 on the entire $\bigcup_i A_i$. For example, for each i there is an object $x \in A_i$ such that $x_R = \emptyset$. Then for every such x, $\sigma(x) = \emptyset$.

Lemma 1.2 *Let $\mathcal{A} \models$ TST. For every formula $\phi(x_1^{i_1}, \ldots, x_n^{i_n})$ of L_{TST}, with free variables among $x_1^{i_1}, \ldots, x_n^{i_n}$, and every sequence of objects $a_1, \ldots, a_n$ such that $a_k \in A_{i_k}$,*

$$\mathcal{A} \models \phi(a_1, \ldots, a_n) \iff lc(\mathcal{A}) \models \phi(\sigma(a_1), \ldots, \sigma(a_n)).$$

Proof. By induction on the length of ϕ. Consider the atomic formula $x^i \varepsilon y^{i+1}$. Then for every $a \in A_i$ and every $b \in A_{i+1}$ we have, by the definition of σ,

$$\mathcal{A} \models a\varepsilon b \iff aRb \iff \sigma(a) \in \sigma(b) \iff lc(\mathcal{A}) \models \sigma(a)\varepsilon\sigma(b).$$

Similarly for the atomic formula $x^i = y^i$, and any $a, b \in A_i$ we have

$$\mathcal{A} \models a = b \iff a = b \iff lc(\mathcal{A}) \models \sigma(a) = \sigma(b).$$

The other steps of the induction are routine. $\dashv$

It follows from lemma 1.2 that for every model $\mathcal{A}$ of TST, not only is $lc(\mathcal{A})$ a s.t. model of TST, but also $\mathcal{A}$ and $lc(\mathcal{A})$ are "almost isomorphic". In view of this fact, and also of the lemma 1.6 below, talking about models of TST can be practically restricted to talking about s.t. models only. So henceforth, unless otherwise stated, a model of TST will be a s.t. one, and we shall write simply $\mathcal{A} = (A_0, A_1, \ldots)$ instead of $\mathcal{A} = (A_0, A_1, \ldots, \in)$. Moreover, in cases where we write formulas of L_{TST} semi-formally, we conflate ε with $\in$.

Let $\mathcal{A} = (A_0, A_1, \ldots)$ be a model of TST. A set $X \subseteq A_i$ is said to be *definable in* $\mathcal{A}$ if $X = \{x \in A_i : \mathcal{A} \models \phi(x)\}$ for some formula $\phi(x^i)$, possibly

[2]If the relation R of $\mathcal{A}$ is well-founded, then the usual "Mostowski collapse" π can also be defined on A by induction on the R-rank of the elements of A. Namely $\pi(x) = \{\pi(y) : yRx\}$. Then for every i and every $x \in A_i$ and $y \in A_{i+1}$, $xRy \iff \pi(x) \in \pi(y)$. However π may distort heavily $\mathcal{A}$. For example if all elements of A_0 are R-minimal, then $\pi''A_0 = \{\emptyset\}$. So this kind of collapse has no practical use.

with parameters. In view of the axiom (Co), X is definable in $\mathcal{A}$ iff X belongs to (some level of) $\mathcal{A}$. For example if $f : A_i \to A_i$ is a definable function, then $f \in A_{i+3}$. Sets containing objects from various levels of $\mathcal{A}$ cannot belong to $\mathcal{A}$ (being non-stratified objects). Nevertheless we can extend to them also the property of definability, by considering certain stratified copies of them. Such copies are constructed using the "shifting mapping" $\imath(x) = \{x\}$ and its iterations. To be specific, a mapping $f : A_i \to A_{i+1}$ is said to be *definable* if the mapping $g : \imath'' A_i \to A_{i+1}$ such that $g(\imath(x)) = f(x)$ (i.e., $g \circ \imath = f$) is definable. E.g. $\imath$ itself is definable, since the identity id for which $id \circ \imath = \imath$ is definable.

In every structure $\mathcal{A} = (A_0, A_1, \ldots)$, each A_{i+1} plays, roughly, the role of the powerset of A_i. So although bijections $f : A_i \to A_{i+1}$ always exist for countable $\mathcal{A}$, none of them can be definable in $\mathcal{A}$, because otherwise Cantor's diagonal argument reappears.

Lemma 1.3 *Let $\mathcal{A} = (A_0, A_1, \ldots)$ be a model of* TST*. For every i, there is no definable bijection $f : A_i \to A_{i+1}$.*

Proof. Suppose $f : A_i \to A_{i+1}$ is a definable bijection. Then there is a definable bijection $g : \imath'' A_i \to A_{i+1}$. By (Co) there is $y \in A_{i+1}$ such that $y = \{x \in A_i : x \notin g(\imath(x))\}$. Hence there is $c \in A_i$ such that $g(\imath(c)) = y$. Then for every $x \in A_i$,

$$x \in y \iff x \in g(\imath(c)) \iff x \notin g(\imath(x)).$$

For $x = c$ we have $c \in g(\imath(c)) \iff c \notin g(\imath(c))$, a contradiction. $\dashv$

The discussion in subsequent sections involves the property of "finiteness" of sets of a model $\mathcal{A}$. Every model of TST has an internal notion of finiteness, expressed by a formula $Fin(x^{i+1})$ of L_{TST}. This is defined in terms of the operation S defined on each A_{i+2} as follows (we drop the type superscripts for readability and write $\in$ instead of ε):

$$S(z) = \{y : (\exists y_1 \in z)(\exists x \notin y_1)(y = y_1 \cup \{x\})\}.$$

Intuitively, S sends the class of all sets (of some level) with n elements, to the class of sets (of the same level) with $n+1$ elements. For example, if $z = \{\emptyset\}$, then $S(z)$ is the class of singletons, and so on. Now a set is finite if it belongs to the intersection of all sets which contain $\emptyset$ and whenever they include y, they include also $S(y)$. Formally (dropping superscripts again)

$$Fin(x) \iff \forall z[\emptyset \in z \ \wedge (\forall y)(y \subseteq z \Rightarrow S(y) \subseteq z) \ \Rightarrow x \in z].$$

In contrast to $Fin(x)$, we have also the external notion of finiteness (i.e., with respect to the ground model M), which we express by writing simply "x is finite". For every model $\mathcal{A}$ of TST and every $x \in A_{i+1}$, one can easily see by induction on $|x|$ that

$$x \text{ is finite } \Rightarrow \mathcal{A} \models Fin(x). \tag{1}$$

The converse of (1) is false in general[3]. However there are models $\mathcal{A}$ for which the converse of (1) is also true, e.g. the full models $\langle\langle X \rangle\rangle$. Such models, for which finiteness is absolute, are considered in section 3, where they are called "regular".

1.2 The system NF

The language $L_{\rm NF}$ of NF consists of the predicate ε and untyped variables $x, y, \ldots$. A formula of $L_{\rm NF}$ is called *stratified* if it results from a formula of $L_{\rm TST}$ if we erase all type superscripts from its variables. The axioms of NF are stratified comprehension and extensionality:

(StCo) $(\exists x)(\forall y)(y \varepsilon x \Leftrightarrow \phi(y))$, for every stratified $\phi(y) \in L_{\rm NF}$, possibly with extra free variables.

(Ex) $(\forall x)(x \varepsilon y \Leftrightarrow x \varepsilon z) \Rightarrow \ y = z$.

Given a model $\mathcal{A} = (A_0, A_1, \ldots)$ of TST, let $\mathcal{A}^+ = (A_1, A_2, \ldots)$. $\mathcal{A}^+$ is still a model of TST.

Given a formula $\phi \in L_{\rm TST}$ let ϕ^+ denote the formula resulting from ϕ if we raise the type of each variable of ϕ by one. If TST $\vdash \phi$ then TST $\vdash \phi^+$. But the converse is false. The axiom scheme:

(Amb) $\phi \Leftrightarrow \phi^+$, $\phi \in L_{\rm TST}$,

is called *Typical Ambiguity.*

[3] See e.g. the model that Grishin constructs in [2], in which A_0 is externally infinite and yet $\mathcal{A} \models Fin(x)$ holds for every $x \in A_{i+1}$.

Definition 1.4 Let $\mathcal{A} = (A_0, A_1, \ldots, R)$ be a model of TST (not necessarily s.t.) A *type-shifting automorphism* (or just an *automorphism*) of $\mathcal{A}$ is a sequence of mappings $f = (f_0, f_1, \ldots)$ such that:

(a) each f_i is 1-1,

(b) $dom(f_i) = A_i$,

(c) $rng(f_i) = A_{i+1}$,

(d) $xRy \iff f_i(x)Rf_{i+1}(y)$ for every $x \in A_i$ and $y \in A_{i+1}$.

We often denote this automorphism by $A_0 \xrightarrow{f_0} A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} \cdots$

The following fundamental result of E. Specker relates TST, typical ambiguity, and NF.

Theorem 1.5 (Specker [4])

The following are equivalent:

(i) NF *is consistent.*

(ii) There is a model $\mathcal{A}$ *of* TST *possessing an automorphism* $f : \mathcal{A} \to \mathcal{A}^+$.

(iii) There is a model $\mathcal{A}$ *of* TST *such that* $\mathcal{A} \equiv \mathcal{A}^+$.

(iv) TST + Amb *is consistent.*

The difficult and most important implication is (iv)$\Rightarrow$ (ii). For a proof see [4], or [1], p. 58.

Lemma 1.6 *For every model* $\mathcal{A}$ *of* TST*, there exists an automorphism on* $\mathcal{A}$ *iff there exists an automorphism on* $lc(\mathcal{A})$.

Proof. Let $\mathcal{A} = (A_0, A_1, \ldots, R)$ be a model of TST and let

$$A_0 \xrightarrow{f_0} A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} \cdots$$

be a an automorphism. Let $lc(\mathcal{A}) = (B_0, B_1, \ldots)$ and let $\sigma : \mathcal{A} \to lc(\mathcal{A})$ be the level collapsing mapping. Put $g_i = \sigma \circ f_i \circ \sigma^{-1}$, for every i. Since σ is 1-1 on each level A_i, g_i is well-defined. And clearly $g_i : B_i \to B_{i+1}$ is a bijection. Moreover, in view of lemma 1.2, for every $x \in B_i$ and every $y \in B_{i+1}$,

$$x \in y \iff \sigma^{-1}(x)R\sigma^{-1}(y) \iff f_i\sigma^{-1}(x)Rf_{i+1}\sigma^{-1}(y) \iff$$

$$\sigma f_i\sigma^{-1}(x) \in \sigma f_{i+1}\sigma^{-1}(y) \iff g_i(x) \in g_{i+1}(y).$$

Therefore

$$B_0 \xrightarrow{g_0} B_1 \xrightarrow{g_1} B_2 \xrightarrow{g_2} \cdots$$

is an automorphism. The converse is similar. Given the automorphism $(g_0, g_1, \ldots)$ of $lc(\mathcal{A})$, it suffices to define $f_i = \sigma^{-1} \circ g_i \circ \sigma$. $\dashv$

In view of lemma 1.6, theorem 1.5 remains true if we replace the word "model" by "s.t. model" everywhere.

If $\mathcal{A}$ is s.t. and $f : \mathcal{A} \to \mathcal{A}^+$ is an a automorphism, then condition (d) of definition 1.4 becomes

$$x \in y \iff f_i(x) \in f_{i+1}(y)$$

for every $x \in A_i$ and $y \in A_{i+1}$. Since f_i are onto, this is equivalent to the property that for every $x \in A_{i+1}$,

$$f_{i+1}(x) = f_i''x, \tag{2}$$

i.e., that $f_{i+1} = f_i''$.

1.3 Fragments of NF

For $n > 0$, a formula ϕ of L_{TST} is an *n-formula*, if every variable of ϕ is of type $< n$. Let TST_n be the subtheory of TST whose axioms are those of TST restricted to n-formulas. A model of TST_n is an n-sequence $\mathcal{A} = (A_0, A_1, \ldots, A_{n-1})$.

Similarly, a formula ϕ of L_{NF} is *n-stratified* if it results from an n-formula of TST by erasing the types from the variables. NF_n is the subtheory of NF in which the scheme of Stratified Comprehension is restricted to n-stratified formulas.

An easy consequence of theorem 1.5 is the following.

Corollary 1.7 *For every n, NF_n is consistent iff there is a model $(A_0, A_1, \ldots, A_{n-1})$ of TST_n possessing an automorphism $f : (A_0, A_1 \ldots, A_{n-2}) \to (A_1, A_2 \ldots, A_{n-1})$.*

Grishin [2] has shown that NF_3 is consistent, and that NF is equivalent to NF_4.

Theorem 1.8 (Grishin) *(a)* NF_3 *is consistent.*
(b) $\mathrm{NF} = \mathrm{NF}_4 = \mathrm{NF}_3 + \exists z\ z = \{\{\{x\}, y\} : x \varepsilon y\}$.

The following result is an adaptation of an idea of Grishin's (used in [2] to prove 1.8 (a)), to the fragment NF_4.

Lemma 1.9 NF *is consistent iff there is a model* $\mathcal{A} = (A_0, A_1, \ldots)$ *of* TST *and a pair of bijections* $A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3$ *such that for all* $x, x_1, x_2 \in A_1$ *and all* $y \in A_2$

$$x \in y \iff f_1(x) \in f_2(y), \tag{3}$$

$$x_1 \subseteq x_2 \iff f_1(x_1) \subseteq f_1(x_2). \tag{4}$$

Proof. If NF is consistent then, by 1.5, there is a model $\mathcal{A} = (A_0, A_1, \ldots)$ of TST with an automorphism $f = (f_0, f_1, f_2, \ldots)$. Then clearly the pair (f_1, f_2) satisfies (3) and (4).

Conversely, suppose there is a model $\mathcal{A} = (A_0, A_1, \ldots)$ of TST and a pair of bijections $A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3$ satisfying (3) and (4). Then (A_0, A_1, A_2, A_3) is a model of TST_4. As Grishin observed in [2] (for the case of NF_3 and TST_3, respectively), if we define $f_0 : A_0 \to A_1$ by setting

$$f_0(a) = x \iff f_1(\{a\}) = \{x\},$$

then f_0 is a bijection because, by (4), f_1 sends atoms to atoms. Moreover (f_0, f_1, f_2) is an automorphism from (A_0, A_1, A_2) onto (A_1, A_2, A_3). Indeed, by the definition of f_0 we have that for every a, $f_1(\{a\}) = \{f_0(a)\}$. So for every $a \in A_0$ and $x \in A_1$

$$f_0(a) \in f_1(x) \iff \{f_0(a)\} \subseteq f_1(x) \iff f_1(\{a\}) \subseteq f_1(x).$$

Now by (4),

$$f_1(\{a\}) \subseteq f_1(x) \iff \{a\} \subseteq x \iff a \in x.$$

Therefore combining the above equivalences we get

$$a \in x \iff f_0(a) \in f_1(x).$$

The last equivalence and (3) show that (f_0, f_1, f_2) is an automorphism $(A_0, A_1, A_2,) \to (A_1, A_2, A_3)$. By 1.7, this implies that there is a model of NF_4, and hence, by 1.8 (b), there is a model of NF. $\dashv$

In the sequel, when a pair of bijections (f_1, f_2) satisfies conditions (3) and (4) above, we shall say that f_1 and f_2 are $\in$- and $\subseteq$-*isomorphisms.*

2 Attacking the NF consistency problem with forcing

Let $\mathcal{A} = (A_0, A_1, A_2, A_3)$ be a model of TST_4 (or more generally a model $\mathcal{A} = (A_0, A_1, A_2, \ldots)$ of TST). By lemma 1.9, in order to turn $\mathcal{A}$ into a model of NF_4, and hence of NF, it suffices to construct a pair of bijections $A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3$ satisfying conditions (3) and (4). Now it is natural to attempt to construct (f_1, f_2) by forcing, i.e., via its finite (and not only finite) parts.

The idea is the usual one: If we start with a model M of ZFC containing $\mathcal{A} = (A_0, A_1, A_2, A_3)$, and we are able to find an appropriate set of forcing conditions $(P, \leq)$ consisted of parts of the pair (f_1, f_2), then (f_1, f_2) will emerge as a generic subset of P, and hence a model of NF will appear in the extension $M[G]$ of M.

However, the preceding idea comprises two main steps: (a) To define an appropriate set of forcing conditions, and (b) to show that the generic set does the job it was designed to, i.e., provides an automorphism for the underlying model.

Concerning step (a), the forcing conditions are going to be not all, but *some* (if any) of the finite approximations of the sought pair (f_1, f_2). So we shall define first certain pairs of functions $p = (p_1, p_2)$, that we shall call "coherent pairs". We do not yet call (p_1, p_2) "forcing conditions" because they lack in general the key property that forcing conditions ought to have, namely *extendibility*. So in section 2.1 we introduce coherent pairs. In fact in order to speak about them formally a certain extension of the theory TST and its language is needed.

Step (b) concerns the crucial property of extendibility of coherent pairs. We deal with this in section 2.2. A pair p is extendible if for every element t of the model, there is a pair q that extends p and captures t. Stronger notions of n-extendibility, for $n \geq 1$, and ω-extendibility are introduced. A pair p is $(n+1)$-extendible if for every element t, there is an n-extendible pair q that extends p and captures t. p is ω-extendible, if it is n-extendible for every $n \geq 1$. The reduction of NF consistency problem stated in the title, consists in replacing it with the problem of whether, for every $n \geq 1$, there is a model $\mathcal{A}$ of TST containing an n-extendible coherent pair. If this is the case, then there is a model $\mathcal{B}$ of TST containing ω-extendible coherent pairs. Then the set P_ω of ω-extendible coherent pairs can be used as the set of forcing conditions, and any generic $G \subseteq P_\omega$ provides a triple (f_0, f_1, f_2) of isomorphisms for the structure (B_0, B_1, B_2, B_3). (Actually G provides the pair (f_1, f_2). f_0 is trivially defined through f_1.) So a model of NF_4 (and hence of NF) exists in the generic extension $M[G]$ of the ground model M.

To summarize: If for every $n \in \mathbb{N}$, n-extendible coherent pairs exist in a model M of ZFC, then there is a model of NF in a generic extension $M[G]$ of M. The converse is also true rather trivially: If M contains a model of NF, then it contains a model $\mathcal{A}$ of TST in which there are n-extendible coherent pairs, for every $n \geq 1$. This is briefly the content of the main theorem 2.8.

2.1 Coherent pairs

Throughout M will be a fixed countable s.t. model of ZFC and $\mathcal{A} = (A_0, A_1, \ldots) \in M$ will be a countable (in the sense of M) s.t. model of TST. In view of lemma 1.9, in order for $\mathcal{A}$ to be turned into a model of NF, it suffices that a pair of bijections $A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3$ be constructed satisfying conditions (3) and (4). Then the parts[4] of (f_1, f_2) should also satisfy conditions (3) and (4). The parts of (f_1, f_2) will be called "coherent pairs" and will be denoted $p = (p_1, p_2)$. Thus, a coherent pair of $\mathcal{A}$) is a pair of functions $p = (p_1, p_2)$, such that $dom(p_1) \subseteq A_1$, $dom(p_2) = rng(p_2) \subseteq A_2$, $rng(p_2) \subseteq A_3$, and for all $x \in dom(p_1)$ and $y \in dom(p_2)$

$$x \in y \iff p_1(x) \in p_2(y),$$

and for all $x_1, x_2 \in dom(p_1)$

$$x_1 \subseteq x_2 \iff p_1(x_1) \subseteq p_1(x_2).$$

However these conditions are not enough. For example, whenever $x \in A_1$, $y \in A_2$ and $f_1(x) = y$, then $f_0''x = y$, hence $|x| = |y|$ and moreover $|A_0 - x| = |A_1 - y|$ since f_0 is a bijection. Similarly for y, z such that $f_2(y) = z$. So we need first the following definition:

Definition 2.1 Let V, W be sets such that $|V| = |W|$, and let $X \subseteq \mathcal{P}(V)$ and $Y \subseteq \mathcal{P}(W)$. We say that X and Y are *similar* and we write $(X)_V \sim$

[4]We do not mean necessarily "finite parts", although in many forcing constructions one constructs a generic object from its finite parts, whenever this possible. Here, on the contrary, finiteness would be an obstacle because, as we saw in section 1.1, this property is not absolute for models of TST.

$(Y)_W$, or just $X \sim Y$, if there is a bijection $h : V \to W$ such that $Y = \{h''x : x \in X\}$. Further, if g is a 1-1 mapping such that $dom(g) \subseteq X$ and $rng(g) \subseteq Y$, we say that X and Y are *similar modulo* g and we write $X \sim_g Y$, if there is a bijection $h : V \to W$ such that $Y = \{h''x : x \in X\}$ and $g \subseteq h$.

For finite X, Y, an equivalent description of the relation $X \sim Y$, can be given as follows. Each finite set $X \subseteq \mathcal{P}(V)$ generates a partition of V into sets which are the atoms of the least Boolean algebra that contains X. Let $B(X)$ denote this Boolean algebra and let $B_0(X)$ denote the set of atoms of $B(X)$. Then the following holds.

Lemma 2.2 *Let V, W be sets such that $|V| = |W|$, and let $X \subseteq \mathcal{P}(V)$ and $Y \subseteq \mathcal{P}(W)$ be finite. The following are equivalent:*
(a) $X \sim Y$,
(b) $B_0(X) \sim B_0(Y)$,
(c) There is a bijection $g : B_0(X) \to B_0(Y)$ such that $|x| = |g(x)|$ for every $x \in B_0(X)$. Of course this bijection can be extended to the whole $B(X)$ with the same property.

Proof. (a)⇒(b): Let $X \sim Y$, let $X = \{x_1, \ldots, x_n\}$, let $h : V \to W$ be a bijection such $Y = \{h''x : x \in X\}$, and let $h''x_k = y_k$, for $k = 1, \ldots, n$. Therefore $Y = \{y_1, \ldots y_n\}$. The elements of $B_0(X)$ have the form $\bigcap_{k=1}^n x_k^{\sigma(k)}$, where σ is a mapping $\sigma : \{1, \ldots, n\} \to \{0, 1\}$ and $x_k^{\sigma(k)} = x_k$ if $\sigma(k) = 1$, while $x_k^{\sigma(k)} = \overline{x_k} = V - x_k$ if $\sigma(k) = 0$. And similarly for $B_0(Y)$. Then for every atom $\bigcap_{k=1}^n x_k^{\sigma(k)}$ of $B_0(X)$ we clearly have $h''(\bigcap_{k=1}^n x_k^{\sigma(k)}) = \bigcap_{k=1}^n h''x_k^{\sigma(k)} = \bigcap_{k=1}^n y_k^{\sigma(k)}$, which is the corresponding atom of $B_0(Y)$. Therefore $X \sim Y$ implies $B_0(X) \sim B_0(Y)$.

(b)⇒(a) is trivial.

(b)⇒(c): Let $B_0(X) \sim B_0(Y)$, and let $h : V \to W$ be the bijection such that $B_0(Y) = \{h''x : x \in B_0(X)\}$. If we set $g(x) = h''x$ for every $x \in B_0(X)$ then $g : B_0(X) \to B_0(Y)$ is a bijection such that $|g(x)| = |x|$.

(c)⇒(b): Let $g : B_0(X) \to B_0(Y)$ be a bijection such that $|g(x)| = |x|$. For every $x \in B_0(X)$ pick a bijection $h_x : x \to g(x)$. If we set $h = \bigcup_{x \in B_0(X)} h_x$, h is a bijection between V and W (since the elements of $B_0(X)$ form a partition of V) such that $B_0(Y) = \{h''(x) : x \in B_0(X)\}$ and hence $B_0(X) \sim B_0(Y)$. ⊣

We are now in a position to define coherent pairs.

Definition 2.3 Let $\mathcal{A} = (A_0, A_1, \ldots)$ be a model TST. A *coherent pair* of $\mathcal{A}$ is a pair $p = (p_1, p_2)$ of 1-1 mappings with the following properties:

(1) $dom(p_1) \in A_2$, $rng(p_1) = dom(p_2) \in A_3$, and $rng(p_2) \in A_4$. We set $u_1 = dom(p_1)$, $u_2 = rng(p_1) = dom(p_2)$ and $u_3 = rng(p_2)$.

(2) p_1, p_2 are $\in$-isomorphisms, i.e., for every $x \in u_1$ and $y \in u_2$,

$$x \in y \iff p_1(x) \in p_2(y).$$

(3) $u_1 \sim u_2$ and $u_2 \sim_{p_1} u_3$, i.e., there are bijections $g : A_0 \to A_1$ and $h : A_1 \to A_2$ such that $u_2 = \{g''x : x \in u_1\}$, $u_3 = \{h''y : y \in u_2\}$ and $p_1 \subseteq h$.

For simplicity we often say just "pair" instead of "coherent pair". And instead of $p = (p_1, p_2)$ we often write more suggestively

$$u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3.$$

Remark 2.4 1) Let

$$A_0 \xrightarrow{f_0} A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3 \cdots$$

be an automorphism. It is easy to see that every restriction of f to a *finite* subset of A_1 is a coherent pair. Indeed, let u_1 be a finite subset of A_1, and let $p_1 = f_1 \restriction u_1$, $u_2 = p_1'' u_1$, $p_2 = f_2 \restriction u_2$ and $u_3 = p_2'' u_2$. Consider the pair $u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3$. Since $f_{i+1} = f_i''$, we have $u_2 = f_1'' u_1 = \{f_1(x) : x \in u_1\} = \{f_0'' x : x \in u_1\}$, hence $u_1 \sim u_2$. Also $u_3 = f_2'' u_2 = \{f_2(y) : y \in u_2\} = \{f_1'' y : y \in u_2\}$ and moreover $p_1 \subseteq f_1$, therefore $u_2 \sim_{p_1} u_3$.

2) Condition (3) of the above definition clearly implies that the mappings p_1, p_2 can be extended to isomorphisms between the corresponding Boolean algebras $B(u_i)$ generated by the sets u_i. So without loss of generality in the above definitions we could take u_i to be finite Boolean algebras. (However the converse is false. If u_i are Boolean algebras such that $u_1 \cong u_2$ and $u_2 \cong u_3$, then (3) need not be true.)

Notational convention. To facilitate reading, we shall henceforth follow the following convention: The variable x (possibly with subscripts) will range exclusively over the level A_1 of $\mathcal{A}$, the variable y will range exclusively over A_2 and the variable z will range exclusively over A_3.

We define a partial ordering $\leq$ between pairs as follows: If $p = (p_1, p_2)$, and $q = (q_1, q_2)$, $p \leq q$ iff $p_1 \supseteq q_1$ and $p_2 \supseteq q_2$. In particular, if $u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2}$

u_3, and $v_1 \xrightarrow{q_1} v_2 \xrightarrow{q_2} v_3$ are two pairs and $p \le q$, then for every $i = 1, 2, 3$ $u_i \supseteq v_i$.

EXAMPLE 1. The simplest example of a coherent pair is that in which u_1, u_2, u_3 are the trivial Boolean subalgebras of A_1, A_2, A_3 respectively and p_1, p_2 are the trivial isomorphisms between them. Namely let: $u_1 = \{\emptyset, A_0\}$, $u_2 = \{\emptyset, A_1\}$, $u_3 = \{\emptyset, A_2\}$, $p_i(\emptyset) = \emptyset$, for $i = 1, 2$, $p_1(A_0) = A_1$, and $p_2(A_1) = A_2$. So properties (1) and (2) of definition 2.3 are trivially satisfied. Further, since $\mathcal{A}$ is countable, $|A_0| = |A_1| = |A_2|$ in the ground model. Take any bijection $g : A_0 \to A_1$. Then $g''A_0 = A_1$ and $g''\emptyset = \emptyset$, that is $u_2 = \{g''x : x \in u_1\}$. Hence $u_1 \sim u_2$. To show that $u_2 \sim_{p_1} u_3$, it suffices to find a bijection $h : A_1 \to A_2$ that extends p_1, i.e., such that $h(\emptyset) = \emptyset$ and $h(A_0) = A_1$. But such a bijection obviously exists: Simply put $h(\emptyset) = \emptyset$, $h(A_0) = A_1$ and then extend it on A_1 arbitrarily. $\dashv$

EXAMPLE 2. Let $a \in A_0$, let $g : A_0 \to A_1$ be a bijection and let $h : A_1 \to A_2$ be a bijection such that $h(\emptyset) = \emptyset$, $h(A_0) = A_1$, $h(\{a\}) = \{g(a)\}$ and $h(A_0 - \{a\}) = A_1 - \{g(a)\}$. Let also

$u_1 = \{\emptyset, A_0, \{a\}, A_0 - \{a\}\}$,
$u_2 = \{\emptyset, A_1, \{g(a)\}, A_1 - \{g(a)\}\}$,
$u_3 = \{\emptyset, A_2, \{hg(a)\}, A_2 - \{hg(a)\}\}$,

Let $p_1 : u_1 \to u_2$ and $p_2 : u_2 \to u_3$ be the bijections that preserve the orderings of the elements of u_i as cited above. Then $u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3$ is a coherent pair. Indeed $u_2 = \{g''x : x \in u_1\}$, $u_3 = \{h''y : y \in u_2\}$ and $p_1 \subseteq h$. Hence $u_1 \sim u_2$ and $u_2 \sim_{p_1} u_3$. Finally it is easy to check that $p = (p_1, p_2)$ is an $\in$-automorphism. $\dashv$

What kind of object is a coherent pair $p = (p_1, p_2)$ with respect to the model $\mathcal{A}$? By definition $u_1 \in A_2$, $u_2 \in A_3$ and $u_3 \in A_4$. But p_1, p_2 need not be coded in $\mathcal{A}$, when they are infinite. Moreover, the property of coherence requires the existence of bijections $g : A_0 \to A_1$ or $h : A_1 \to A_2$. By lemma 1.3, *no* bijection $g : A_0 \to A_1$ or $h : A_1 \to A_2$ can be definable in $\mathcal{A}$ (and hence be coded by an object in $\mathcal{A}$). Therefore g, h necessarily, and p_1, p_2 possibly, will be external objects for $\mathcal{A}$, existing only for an observer in the ground model M, so we can't speak about them in L_{TST}.

In order to be able to speak about coherent pairs formally, we need to extend the language L_{TST} by adding new untyped set variables, denoted by lower case Greek letters $\alpha, \beta, \ldots$, intended to range over unstratified objects of models of TST. These unstratified objects will include infinite mappings between distinct levels of $\mathcal{A}$.

To motivate the definitions, given a model $\mathcal{A} = (A_0, A_1, \dots)$ of TST, define (in the ground model M) the countable cumulative hierarchy over A_0, denoted by $V_\omega(A_0)$, as usual by:

$V_0(A_0) = A_0$.
$V_{n+1}(A_0) = V_n(A_0) \cup \mathcal{P}(V_n(A_0))$.
$V_\omega(A_0) = \bigcup_{n\in\mathbb{N}} V_n(A_0)$.

Since for every n, $A_{n+1} \subseteq \mathcal{P}(A_n)$, it is easy to check by induction that $A_n \subseteq V_n(A_0)$ for all $n \in \mathbb{N}$. Let us call the elements of $V_\omega(A_0)$ *complex*, in contrast to the elements of $\bigcup_n A_n$ which are the *stratified* ones. For example a bijection between A_i and A_{i+1} is a complex object.

Note that according to the above definition each stratified set is already complex. So the relation of complex sets to stratified ones is similar to the relation between classes and sets in the standard theory of classes. Each set is a class but not the other way around. So as we speak of proper classes, we shall speak here of *proper complex* sets. Also as the variable X in class theory ranges over classes in general (including sets), while x ranges over sets only, similarly here α ranges over complex sets in general (including stratified ones), while x^i ranges over stratified sets only.

Let L_{TST^c} be the new language containing the variables α, besides x^i. The formulas of L_{TST^c} are those of L_{TST} plus formulas of the form $\alpha\varepsilon\beta$, $x^i\varepsilon\alpha$, $\alpha = \beta$. A formula of L_{TST^c} is said to be stratified if it is a formula of L_{TST}.

Let TST^c be the theory in the language L_{TST^c} with axioms[5]:

(I) Comprehension for stratified formulas. (This is just the corresponding axiom (Co) of TST.)

(II) Every stratified set is complex:

$$\forall x^i \exists \alpha (x^i = \alpha).$$

[5]The theory TST^c, as we use it here, is only a temporary tool for the proof of the main theorem, with no further significance. That is why we do not specify any particular kinds of complex objects that could exist in the theory. I.e., we do not include any comprehension scheme for complex objects. Perhaps a more specific and refined extension of TST, with an interest in itself, might be considered and studied.

(III) Elements of stratified sets are stratified:

$$\forall\alpha(\alpha\varepsilon x^{i+1} \Rightarrow \exists x^i(x^i = \alpha)).$$

(IV) Extensionality:

$$\forall\gamma(\gamma\varepsilon\alpha \Leftrightarrow \gamma\varepsilon\beta) \;\Rightarrow\; \alpha = \beta.$$

In view of axioms (II) and (III), it is clear that (IV) extends the axiom (Ex) of TST, so TST^c is an extension of the theory TST.

Standard models of TST^c have the form $(\mathcal{A}, C) = (A_0, A_1, \ldots, C)$, where A_i contain the stratified sets as usual, and C contains the proper complex objects. Natural such models are e.g. those of the form

$$(X, \mathcal{P}(X), \mathcal{P}^2(X), \ldots, C),$$

where $C = V_\omega(X) - \bigcup_n \mathcal{P}^n(X))$.

In TST^c, ordered pairs, relations, functions etc. are defined as usual.

Lemma 2.5 *The property "α is a coherent pair"is definable in* TST^c*.*

Proof. The property "α is a coherent pair" is the conjunction of the following statements:

(1) α is a pair of objects (α_1, α_2) each of which is a 1-1 mapping, such that for some x_1^2, x_2^3, x_3^4, $dom(\alpha_1) = x_1^2$, $rng(\alpha_1) = dom(\alpha_2) = x_2^3$ and $rng(\alpha_2) = x_3^4$.

(2) α_1 and α_2 are ε- and $\subseteq$-preserving mappings.

(3) There are bijections $\gamma_1 : A_1 \to A_2$ and $\gamma_2 : A_2 \to A_3$ which entail that $x_1^2 \sim x_2^3$ and $x_2^3 \sim_{\alpha_1} x_3^4$.

The statements "$\gamma_1 : A_1 \to A_2$ is a bijection" and "$\gamma_2 : A_2 \to A_3$ is a bijection" are clearly expressible in L_{TST^c}. $\dashv$

2.2 Extendibility

Recall that if P is the set of finite 1-1 mappings from a set A into a set B, ordered by (reverse) inclusion, and G is a generic subset of P, in order to prove that $f = \cup\{p : p \in G\}$ is a bijection between A and B, we need to show that for every $p \in P$ and every $a \in A$ we can *extend* p to a condition q such that $a \in dom(q)$ and/or $a \in rng(q)$. This is the *extendibility property* of P.

Extendibility guarantees that the sets of conditions q such that $a \in dom(q)$ or $a \in rng(g)$ are all dense in P, and hence meet G. For the P of the example just mentioned the property holds trivially. However for the set of coherent pairs extendibility is not a simple matter to hold. As follows from Example 3 below, a lot of coherent pairs are not extendible.

Definition 2.6 Let $p = u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3$ be a pair. We say that p is *everywhere extendible* or just *extendible* if for every $t \in A_1 \cup A_2 \cup A_3$, there is a pair $v_1 \xrightarrow{q_1} v_2 \xrightarrow{q_2} v_3$ such that $q \leq p$ and $t \in v_1 \cup v_2 \cup v_3$. When such a pair $q = (q_1, q_2)$ exists, we say for simplicity that q *captures* t.

Definition 2.7 Let $p = u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3$ be a pair. p is said to be *1-extendible* if it is extendible. p is said to be $(n+1)$-*extendible* if for every $t \in A_1 \cup A_2 \cup A_3$ there is a pair $q = (q_1, q_2)$ such that q captures t, $q \leq p$ and q is n-extendible. p is said to be ω-*extendible* if it is n-extendible for all $n \geq 1$.

An easy induction shows that if (p_1, p_2) is n-extendible and $m < n$, then (p_1, p_2) is m-extendible.

Example 3. Consider the coherent pair of Example 2, where
$u_1 = \{\emptyset, A_0, \{a\}, A_0 - \{a\}\}$,
$u_2 = \{\emptyset, A_1, \{g(a)\}, A_1 - \{g(a)\}\}$,
$u_3 = \{\emptyset, A_2, \{hg(a)\}, A_2 - \{hg(a)\}\}$.
If we choose the bijections g, h so that $|g(a)| \neq |hg(a)|$, then we can easily see that the pair (p_1, p_2) is not extendible. Indeed, first note that $hg(a) \notin u_2$. For if $hg(a) \in u_2$, then, because of the coherence, $g(a) \in u_1$ and $p_1(g(a)) = hg(a)$, hence $|g(a)| = |hg(a)|$, a contradiction. Therefore $hg(a) \notin u_2$. By extendibility, there must be a pair $q = (q_1, q_2)$ such that $q \leq p$ and q captures $hg(a)$. But $hg(a)$ is the unique element of an element of u_3, namely $\{hg(a)\}$. Therefore, in order for q to be coherent, it must contain the corresponding unique element $g(a)$ of $\{g(a)\}$ of u_2 and moreover $q_1(g(a)) = hg(a)$. But then also we must have $|g(a)| = |hg(a)|$, which is not the case. Thus p is not extendible. $\dashv$

Theorem 2.8 *(*Main Theorem) *Let M be a countable model of* ZFC *in which for every $n \in \mathbb{N}$, there is a s.t. model $\mathcal{A}$ of* TST *that contains an n-extendible coherent pair. Then there is a generic extension $M[G]$ of M*

that contains a model of NF. *Conversely, if M contains a model of* NF*, then in M there is a s.t. model $\mathcal{A}$ of* TST *that contains an n-extendible pair, for every $n \geq 1$.*

Proof. Let us show first the converse. Let M contain a model of NF. Then M contains also a model $\mathcal{A} = (A_0, A_1, A_2, \ldots, R)$ of TST with an automorphism

$$A_0 \xrightarrow{f_0} A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3 \cdots$$

By lemma 1.6 we may assume that $\mathcal{A}$ is s.t. As we saw in remark 2.4 (1), if $u_1 \subseteq A_1$ is finite, and we set $p_1 = f_1 \restriction u_1$, $u_2 = p_1'' u_1$, $p_2 = f_2 \restriction u_2$ and $u_3 = p_2'' u_2$, then

$$p = u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3$$

is a coherent pair. Moreover p is n-extendible for every $n \geq 1$ (that is, ω-extendible). This can be shown inductively for all such restrictions p of f. Indeed, obviously all such restrictions p are extendible, i.e., 1-extendible. Suppose all such p are n-extendible and pick some specific p. If $t \in A_1 \cup A_2 \cup A_2$, then we can choose a part q of f which extends p and captures t. Since by assumption q is n-extendible, p is $(n+1)$-extendible.

We come to the converse. Let M be a countable model of ZFC such that for every $n \in \mathbb{N}$, there is a s.t. model $\mathcal{A}$ of TST in M that contains an n-extendible coherent pair. We turn to the language of TST^c. By lemma 2.5, the property "α is a coherent pair" is definable in TST^c. Let $\theta_0(\alpha)$ be the formula that defines it. It is easy to see by induction on n, that for every $n \geq 1$ there is a formula $\theta_n(\alpha)$ of L_{TST^c} expressing the fact that "α is an n-extendible coherent pair". Recall that a coherent pair is of the form $p = u_1 \xrightarrow{p_1} u_2 \xrightarrow{p_2} u_3$ with $u_i \in A_{i+1}$, i.e., u_i is of type $i+1$. So using typed variables in the place of u_i and untyped variables in the place of p_i and p, a coherent pair is more formally written as $\alpha = x_1^2 \xrightarrow{\alpha_1} x_2^3 \xrightarrow{\alpha_2} x_3^4$.

Now suppose that $\theta_n(\alpha)$ is defined for $n \geq 0$, where,

$$\alpha = x_1^2 \xrightarrow{\alpha_1} x_2^3 \xrightarrow{\alpha_2} x_3^4.$$

Then $\theta_{n+1}(\alpha)$ is defined to be the formula:

$$\bigwedge_{i=1}^{3} [\forall z^i \exists y_1^2 y_2^3 y_3^4 \exists \beta_1 \beta_2 \exists \beta (\beta = y_1^2 \xrightarrow{\beta_1} y_2^3 \xrightarrow{\beta_2} y_3^4 \ \&\ \theta_n(\beta) \ \&\ \beta \leq \alpha \ \&\ z^i \varepsilon y_i^{i+1})].$$

It is easy to check that $\theta_{n+1}(\alpha)$ expresses the fact that α is an $(n+1)$-extendible coherent pair. Moreover, by induction we easily check that for $m < n$, every n-extendible coherent pair is m-extendible, therefore

$$m < n \ \&\ \theta_n(\alpha) \ \Rightarrow\ \theta_m(\alpha). \tag{5}$$

Consider the theory

$$T = \mathrm{TST}^c + \{\theta_n(b) : n \in \mathbb{N}\}$$

in the language $L_{\mathrm{TST}^c}(b)$ with a new constant b. By our assumption, for every $n \in \mathbb{N}$, $\mathrm{TST}^c + \theta_n(b)$ has a model, so by (5) and compactness T has a model. Let $(\mathcal{A}, C)$ be a countable model of T in M. Already $(\mathcal{A}, C)$ contains an ω-extendible element (the interpretation of the constant b in $(\mathcal{A}, C)$). Let $(\mathcal{B}, D) \in M$ be a saturated model such that $(\mathcal{A}, C) \preceq (\mathcal{B}, D)$. (As to the kind of saturation, ω-saturation and even recursive saturation suffices.) $(\mathcal{B}, D) \models T$, so $(\mathcal{B}, D)$ also contains ω-extendible pairs. Let P_ω be the set of ω-extendible coherent pairs of $(\mathcal{B}, D)$. Clearly $P_\omega \in M$ and we use $(P_\omega, \leq)$ as a set of forcing conditions in M.

Claim 1: Let $\mathcal{B} = (B_0, B_1, B_2, \ldots)$. If $p \in P_\omega$ and $t \in B_1 \cup B_2 \cup B_3$, then there is a $q \in P_\omega$ such that $q \leq p$ and q captures t.

Proof. By assumption, p is ω-extendible, i.e., n-extendible for every $n \geq 1$. Therefore, given $t \in B_1 \cup B_2 \cup B_3$ for each n, there is an n-extendible pair q_n such that $q_n \leq p$ and q_n captures t. Consider the type

$$s(\alpha) = \{\alpha \leq p \ \ \&\ \alpha \text{ captures } t \ \ \&\ \theta_n(\alpha) : n \in \mathbb{N}\}$$

in the language L_{TST^c}. Then $s(\alpha)$ is finitely satisfiable in $(\mathcal{B}, D)$ (moreover $s(\alpha)$ is a recursive type, if one wants to use recursive saturation). Therefore it is satisfiable in $(\mathcal{B}, D)$, and hence there is a $q \in P_\omega$ such that $q \leq p$ and q captures t. This proves Claim 1.

Claim 2. Let G be a generic subset of P_ω. Then G defines a pair of bijections $B_1 \xrightarrow{f_1} B_2 \xrightarrow{f_2} B_3$ that satisfies conditions (3) and (4).

Proof. Obviously if $f_1 = \bigcup\{p_1 : p \in G\}$ and $f_2 = \bigcup\{p_2 : p \in G\}$, then (f_1, f_2) is a pair of $\in$- and $\subseteq$-preserving mappings. So it suffices to see that $dom(f_1) = B_1$, $dom(f_2) = rng(f_1) = B_2$ and $rng(f_2) = B_3$. But this follows immediately from Claim 1 (and genericity). This proves Claim 2.

Now in the generic extension $M[G]$ of M, $\mathcal{B}$ is still a model of TST, since the sentences "$\mathcal{B} \models \phi$", for $\phi \in L_{\mathrm{TST}}$, are Δ_0 (since $\mathcal{B} \in M$) and hence absolute between M and $M[G]$. Therefore in $M[G]$, $\mathcal{B}$ is a model of TST, whose part (B_1, B_2, B_3), by Claim 2, has a pair of bijections (f_1, f_2) satisfying (3) and (4). Thus, by lemma 1.9, $\mathcal{B}$ yields a model of NF in $M[G]$. $\dashv$

After theorem 2.8, the first task is to find models of TST containing extendible (i.e., 1-extendible) pairs. The most natural candidate pair to be extendible would be the pair of Example 1. Given the model $\mathcal{A}$, let us denote it by $o^{\mathcal{A}}$, i.e.,

$$o^{\mathcal{A}} := \mathbf{2}_1 \xrightarrow{o_1} \mathbf{2}_2 \xrightarrow{o_2} \mathbf{2}_3,$$

where $\mathbf{2}_i$ denotes the trivial Boolean subalgebra of A_i, $i = 1, 2, 3$, and o_1, o_2 are the mappings such that $o_i(\emptyset) = \emptyset$ and $o_i(A_{i-1}) = A_i$. So the question becomes: Are there $\mathcal{A}$ such that $o^{\mathcal{A}}$ is extendible in $\mathcal{A}$? Of course crucial for the extendibility of $o^{\mathcal{A}}$ will be the properties of the underlying model $\mathcal{A}$. We shall show in the next section that for $\mathcal{A}$ satisfying some rather mild conditions, $o^{\mathcal{A}}$ is 1-extendible. Moreover, the main theorem above can be equivalently formulated as follows:

Theorem 2.9 *(*Main Theorem*) Let M be a countable model of* ZFC *in which for every $n \in \mathbb{N}$, there is a s.t. model $\mathcal{A}$ of* TST *such that $o^{\mathcal{A}}$ is n-extendible. Then there is a generic extension $M[G]$ of M that contains a model of* NF*. Conversely, if M contains a model of* NF*, then there is a s.t. model $\mathcal{A} \in M$ such that $o^{\mathcal{A}}$ is n-extendible, for every $n \geq 1$.*

3 Existence of 1-extendible pairs

The property of 1-extendibility (or just extendibility) of a pair (p_1, p_2) in $\mathcal{A}$ splits into three subproperties, namely A_i-extendibility, for $i = 1, 2, 3$. Specifically $p = (p_1, p_2)$ is said to be A_i*-extendible*, for $i = 1, 2, 3$, if for every $t \in A_i$ there is a pair $q = (q_1, q_2)$ such that $q \leq p$ and q captures t. So

$$p \text{ is extendible } \iff \bigwedge_{i=1}^{3} (p \text{ is } A_i\text{-extendible}). \tag{6}$$

Recall from section 1.1 that $Fin(x)$ denotes the property of internal finiteness, while "x is finite" means that x is finite in the ground model.

Definition 3.1 A model $\mathcal{A}$ of TST is called *regular* if for every $x \in A$,

$$x \text{ is finite } \iff \mathcal{A} \models Fin(x). \tag{7}$$

Another property we shall need concerning the model $\mathcal{A}$ is the property of "richness" used already by Grishin in [2]. (T. Forster [1], p. 60, calls the property "saturation in the sense of Grishin". Since saturation has several other meanings, we prefer a more neutral name.)

Definition 3.2 The Boolean algebra A_{i+1} is said to be *rich* if for every infinite (with respect to the ground model) $x \in A_{i+1}$, there is a $x_1 \in A_{i+1}$ such that $x_1 \subseteq x$ and both x_1 and $x - x_1$ are infinite. The structure $\mathcal{A}$ is said to be *rich* if every level A_{i+1}, for $i \geq 0$, is rich.

If $\mathcal{A}$ is regular, then the property of richness is definable in $\mathcal{A}$. Moreover the following holds:

Lemma 3.3 *Let $\langle\langle X\rangle\rangle$ be a full model of* TST *and let $\mathcal{A}$ be the level collapse of an elementary submodel of $\langle\langle X\rangle\rangle$. Then $\mathcal{A}$ is (a) regular and (b) rich.*

Proof. Let $\mathcal{B} \preceq \langle\langle X\rangle\rangle$ and let $\mathcal{A} = lc(\mathcal{B})$. Let $\sigma : \mathcal{B} \to \mathcal{A}$ be the level collapsing function.

(a) Let A_k be the levels of $\mathcal{A}$ and B_k the levels of $\mathcal{B}$. Let $x \in A_{i+1}$. We have to show that "x is finite" $\iff \mathcal{A} \models Fin(x)$.

By (1) of section 1.1, "x is finite" $\Rightarrow \mathcal{A} \models Fin(x)$ is always true.

So suppose $\mathcal{A} \models Fin(x)$. Let $x = \sigma(y)$, for some $y \in B_{i+1}$. Now $Fin(\cdot)$ is a formula of L_{TST}, so by lemma 1.2,

$$\mathcal{A} \models Fin(x) \iff lc(\mathcal{B}) \models Fin(\sigma(y)) \iff \mathcal{B} \models Fin(y).$$

Consequently, since $\mathcal{B} \preceq \langle\langle X\rangle\rangle$, $\langle\langle X\rangle\rangle \models Fin(y)$. The latter clearly implies that "y is finite". Now $\mathcal{B}$ is standard and $\sigma(y) = \{\sigma(u) : u \in y \cap B_i\}$. Therefore $x = \sigma(y)$ is finite, since $y \cap B_i$ is so.

(b) In view of regularity and the absoluteness of the property "x is finite", the property of richness is expressed by the sentence of L_{TST} (written without type indicators):

$$\phi : \forall x[\neg Fin(x) \Rightarrow \exists x_1 x_2(\neg Fin(x_1) \wedge \neg Fin(x_2) \wedge x = x_1 \cup x_2 \wedge x_1 \cap x_2 = \emptyset)].$$

Since obviously $\langle\langle X\rangle\rangle \models \phi$, it follows that $\mathcal{B} \models \phi$ and, by 1.2, $\mathcal{A} \models \phi$. $\dashv$

3.1 A_1- and A_2-extendibility

For every $x \in A_1$ (resp. $y \in A_2, z \in A_3$) we shall write for convenience $-x$ (resp. $-y$, $-z$) instead of $A_0 - x$ (resp. $A_1 - y$, $A_2 - z$).

Lemma 3.4 *Let $\mathcal{A}$ be a countable rich model of* TST. *Then the pair $o^{\mathcal{A}}$ is A_1- and A_2-extendible in $\mathcal{A}$.*

Proof. We have $o^{\mathcal{A}} = \mathbf{2}_1 \xrightarrow{o_1} \mathbf{2}_2 \xrightarrow{o_2} \mathbf{2}_3$, where $\mathbf{2}_1 = \{\emptyset, A_0\}$, $\mathbf{2}_2 = \{\emptyset, A_1\}$, $\mathbf{2}_3 = \{\emptyset, A_2\}$, and $o_i(\emptyset) = \emptyset$, for $i = 1, 2$, $o_1(A_0) = A_1$, $o_2(A_1) = A_2$.

A_1-Extendibility:

Pick first any $x_0 \in A_1$, hence $x_0 \subseteq A_0$. We have to find $y_0 \subseteq A_2$ and $z_0 \subseteq A_3$ such that, if
$v_1 = \{A_0, \emptyset, x_0, -x_0\}$,
$v_2 = \{A_0, \emptyset, y_0, -y_0\}$,
$v_3 = \{A_0, \emptyset, z_0, -z_0\}$,
and if $q_1 : v_1 \to v_2$, $q_2 : v_2 \to v_3$ are the bijections that preserve the orderings of the elements of v_i as cited above, then (q_1, q_2) is a coherent pair.

First, because of the richness of $\mathcal{A}$, we can choose $y_0 \in A_2$ such that
(i) $|y_0| = |x_0|$ and $|-y_0| = |-x_0|$.
This guarantees that $v_1 \sim v_2$. Then, we need a $z_0 \in A_3$ such that
(ii) $|z_0| = |y_0|$ and $|-z_0| = |-y_0|$
and moreover
(a) $\emptyset \in y_0 \iff \emptyset \in z_0$,
(b) $A_0 \in y_0 \iff A_1 \in z_0$,
(c) $x_0 \in y_0 \iff y_0 \in z_0$,
(d) $-x_0 \in y_0 \iff -y_0 \in z_0$.

The above conditions (i), (ii) and (a)-(d) suffice for (q_1, q_2) to be a coherent pair. Indeed, because of (a)-(d), (q_1, q_2) is an $\in$-automorphism. It remains to show that $v_2 \sim_{q_1} v_3$. For that it suffices to construct a bijection $h : A_1 \to A_2$ such that $h''y_0 = z_0$, $h'' - y_0 = -z_0$, and $h(\emptyset) = \emptyset$, $h(A_1) = A_2$, $h(x_0) = y_0$ and $h(-x_0) = -y_0$. But this is obviously possible because of (ii) and (a)-(d).

Now, the existence of a $z_0 \in A_3$ such that (ii) and (a)-(d) hold follows easily from the fact that $\mathcal{A}$ is rich. We just find z_0 so that (ii) holds and in addition z_0 skips or contains each of the four elements $\emptyset, A_1, y_0, -y_0$, accordingly when y_0 does so with respect to the corresponding element of v_1.

A_2-Extendibility:

Let $y_0 \in A_2$, i.e, $y_0 \subseteq A_1$. Using the richness of $\mathcal{A}$, we can choose a set $x_0 \in A_1$ such that $|x_0| = |y_0|$ and $|-x_0| = |-y_0|$. Now we are as in the first step of A_1-extendibility. We continue and find z_0 precisely as we did there. $\dashv$

3.2 A_3-Extendibility

The case of A_3-extendibility of $o^{\mathcal{A}}$ is harder.

Lemma 3.5 *Let $\langle\langle D\rangle\rangle$ be a full model of* TST *(with infinite D) and let $\mathcal{A} = lc(\mathcal{B})$ for some countable $\mathcal{B} \preceq \langle\langle D\rangle\rangle$. Then the pair $o^{\mathcal{A}}$ is A_3-extendible in $\mathcal{A}$.*

Proof. Given $z_0 \in A_3$ we have to show that there are $x \in A_1$ and $y \in A_2$ such that the following conditions, isolated above, in the case of A_1-extendibility, hold simultaneously:

(i) $|y| = |x|$ and $|-y| = |-x|$,
(ii) $|z_0| = |y|$ and $|-z_0| = |-y|$,
(a) $\emptyset \in y \iff \emptyset \in z_0$,
(b) $A_0 \in y \iff A_1 \in z_0$,
(c) $x \in y \iff y \in z_0$,
(d) $-x \in y \iff -y \in z_0$.

For $x \in A_1$, $y \in A_2$, $z \in A_3$ let us write $Sim(x, y)$ if $|x| = |y|$ and $|-x| = |-y|$, and similarly for $Sim(y, z)$. We shall first argue by contradiction, assuming that $o^{\mathcal{A}}$ is not extendible at z_0, and hence there are no $x \in A_1$ and $y \in A_2$ such that (i), (ii) and (a)-(d) hold simultaneously. Then the following is true:

$$(\forall x \in A_1)(\forall y \in A_2)[Sim(x, y) \ \&\ Sim(y, z_0) \ \Rightarrow (\emptyset \in y \not\Leftrightarrow \emptyset \in z_0)\vee \quad (8)$$
$$(A_0 \in y \not\Leftrightarrow A_1 \in z_0)\vee$$
$$(x \in y \not\Leftrightarrow y \in z_0)\vee$$
$$(-x \in y \not\Leftrightarrow -y \in z_0)].$$

Now since z_0 is given, one of the following is the case:
$\emptyset \in z_0$ & $A_1 \in z_0$,

$\emptyset \notin z_0$ & $A_1 \in z_0$,
$\emptyset \in z_0$ & $A_1 \notin z_0$,
$\emptyset \notin z_0$ & $A_1 \notin z_0$.
Without serious loss of generality (as will be evident below), we shall assume that $\emptyset \in z_0$ and $A_1 \in z_0$, or $\{\emptyset, A_1\} \subseteq z_0$. Then we have to search for $y \in A_2$ such that $\{\emptyset, A_0\} \subseteq y$. In view of this restriction (8) transforms to the following:

$$(\forall x \in A_1)(\forall y \in A_2)[Sim(x,y) \ \&\ Sim(y,z_0) \ \&\ \{\emptyset, A_0\} \subseteq y \ \Rightarrow \qquad (9)$$

$$(x \in y \not\Leftrightarrow y \in z_0) \vee$$

$$(-x \in y \not\Leftrightarrow -y \in z_0)].$$

Case 1. Suppose that z_0 is finite. Pick some $y_0 \in A_2$ such that $Sim(y_0, z_0)$, $\{\emptyset, A_0\} \subseteq y_0$ and $y_0 \notin z_0$ and $-y_0 \notin z_0$. Since z_0 is finite this choice of y_0 is obviously possible. Then it follows from (9) that for every x such that $Sim(x, y_0)$, it must be $x \in y_0$ or $-x \in y_0$. But clearly there are infinitely many such x, so y_0 must be infinite, contrary to the assumption that $|y_0| = |z_0|$=finite.

Case 2. Suppose that z_0 is cofinite. Pick a $x_0 \in A_1$ such that $Sim(x_0, z_0)$. Let

$$Y = \{y \in A_2 : Sim(y, z_0) \wedge \{\emptyset, A_0\} \subseteq y \ \wedge x_0 \in y \ \wedge \ -x_0 \in y\}.$$

Then it follows from (9) that $\forall y \in Y(y \notin z_0 \vee -y \notin z_0)$. That is, $\forall y \in Y(y \in -z_0 \vee -y \in -z_0)$. But Y is infinite, while $-z_0$ is finite. This is a contradiction.

Case 3. Suppose finally that z_0 and $-z_0$ are infinite. Since z_0 and $-z_0$ are countable this amounts to $|z_0| = |-z_0|$. Let us call such sets "uniform". Let $A_1^u = \{x \in A_1 : |x| = |-x|\}$ and $A_2^u = \{y \in A_2 : |y| = |-y|\}$ be the sets of "uniform" sets in A_1 and A_2 respectively. Since $\mathcal{A} = \pi''\mathcal{A}' \preceq \langle\langle D \rangle\rangle$, by lemma 3.3, $\mathcal{A}$ is regular. So "x is infinite" $\iff$ $\mathcal{A} \models \neg Fin(x)$. By countability and regularity,

$$|x| = |-x| \iff x \text{ is infinite } \wedge \ -x \text{ is infinite} \iff$$

$$\mathcal{A} \models \neg Fin(x) \wedge \neg Fin(-x).$$

Therefore the sets A_1^u and A_2^u are definable in $\mathcal{A}$. Moreover, since $|z_0| = |-z_0|$, the condition $Sim(x, y) \;\wedge\; Sim(y, z_0)$ in the formula (9) is equivalent to $x \in A_1^u \wedge y \in A_2^u$. So (9) is written more simply

$$\begin{aligned}(\forall x \in A_1^u)(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \;\Rightarrow& \\ &(x \in y \not\Leftrightarrow y \in z_0)\vee \\ &(-x \in y \not\Leftrightarrow -y \in z_0)].\end{aligned} \tag{10}$$

But if $x \in A_1^u$, then $-x \in A_1^u$, so (10) implies also

$$\begin{aligned}(\forall x \in A_1^u)(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \;\Rightarrow& \\ &(-x \in y \not\Leftrightarrow y \in z_0)\vee \\ &(x \in y \not\Leftrightarrow -y \in z_0)].\end{aligned} \tag{11}$$

Taking the conjunction of (10) and (11) and applying some logic calculus we get the following:

$$\begin{aligned}(\forall x \in A_1^u)(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \;\Rightarrow& \\ (x \in y \Leftrightarrow y \notin z_0) \wedge (-x \in y \Leftrightarrow y \notin z_0)&\vee \\ (x \in y \Leftrightarrow y \notin z_0) \wedge (x \in y \Leftrightarrow -y \notin z_0)&\vee \\ (-x \in y \Leftrightarrow -y \notin z_0) \wedge (-x \in y \Leftrightarrow y \notin z_0)&\vee \\ (-x \in y \Leftrightarrow -y \notin z_0) \wedge (x \in y \Leftrightarrow -y \notin z_0)&].\end{aligned} \tag{12}$$

By logic again (12) is written

$$\begin{aligned}(\forall x \in A_1^u)(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \;\Rightarrow& \\ (x \in y \Leftrightarrow -x \in y \Leftrightarrow y \notin z_0)&\vee \\ (y \notin z_0 \Leftrightarrow x \in y \Leftrightarrow -y \notin z_0)&\vee \\ (-y \notin z_0 \Leftrightarrow -x \in y \Leftrightarrow y \notin z_0)&\vee \\ (-x \in y \Leftrightarrow -y \notin z_0 \Leftrightarrow x \in y)&].\end{aligned} \tag{13}$$

Now each of the first and fourth disjuncts of (13) implies $(x \in y \Leftrightarrow -x \in y)$, while each of the second and third implies $(y \in z_0 \Leftrightarrow -y \in z_0)$. So (13) yields

$$(\forall x \in A_1^u)(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \ \Rightarrow \tag{14}$$
$$(x \in y \Leftrightarrow -x \in y) \vee$$
$$(y \in z_0 \Leftrightarrow -y \in z_0)].$$

(14) is equivalently written as follows:

$$(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \ \Rightarrow (\forall x \in A_1^u)(x \in y \Leftrightarrow -x \in y) \vee \tag{15}$$
$$(y \in z_0 \Leftrightarrow -y \in z_0)].$$

In order to simplify (15), we give a definition. Given sets $z_1, z_2 \in A_3$, let us call z_1 *closed with respect to* z_2, if

$$\forall y \in z_2 (y \in z_1 \iff -y \in z_1).$$

Let us denote this property by $C(z_1; z_2)$. Obviously $C(z_1, z_2)$ is a formula of L_{TST}. Similarly is defined the property $C(y_1; y_2)$ for $y_1, y_2 \in A_2$. Note that the subformula $(\forall x \in A_1^u)(x \in y \Leftrightarrow -x \in y)$ of (15) says that y is closed w.r.t. A_1^u, i.e., $C(y; A_1^u)$. So (15) is written as follows:

$$(\forall y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \ \wedge \ \neg C(y; A_1^u) \Rightarrow (y \in z_0 \Leftrightarrow -y \in z_0)]. \tag{16}$$

Let

$$X = \{y \in A_2^u : \{\emptyset, A_0\} \subseteq y \ \wedge \ \neg C(y; A_1^u)\}.$$

Clearly X is definable in $\mathcal{A}$ and $X \in A_3$. Moreover (16) is written

$$(\forall y \in X)(y \in z_0 \Leftrightarrow -y \in z_0)], \tag{17}$$

or, equivalently, "z_0 is closed with respect to X", i.e.

$$C(z_0; X). \tag{18}$$

Now since (8)$\Rightarrow$(18), if z_0 is not closed with respect to X, i.e., if $\neg C(z_0; X)$, then (8) is false, and hence $o^{\mathcal{A}}$ is extendible at z_0. So it remains to show extendibility when $C(z_0; X)$ is true.

Suppose $C(z_0; X)$. We have to show that

$$(\exists x \in A_1^u)(\exists y \in A_2^u)[\{\emptyset, A_0\} \subseteq y \wedge \tag{19}$$

$$(x \in y \Leftrightarrow y \in z_0) \wedge$$
$$(-x \in y \Leftrightarrow -y \in z_0)].$$

In view of the definition of X it suffices to show that

$$(\exists y \in X)(\exists x \in A_1^u)(y \in z_0 \ \wedge \ x \in y \ \wedge \ -x \in y) \tag{20}$$

or

$$(\exists y \in X)(\exists x \in A_1^u)(y \notin z_0 \ \wedge \ x \notin y \ \wedge \ -x \notin y). \tag{21}$$

Indeed if x_0, y_0 satisfy (20), then $-y_0 \in z_0$ holds also because z_0 is closed with respect to X, hence

$$(x_0 \in y_0 \Leftrightarrow y_0 \in z_0) \wedge (-x_0 \in y_0 \Leftrightarrow -y_0 \in z_0)$$

is true. Moreover $\{\emptyset, A_0\} \subseteq y_0$, because $y_0 \in X$. Therefore x_0, y_0 satisfy (19). Similarly if x_0, y_0 satisfy (21), then by closeness again $-y_0 \notin z_0$, hence

$$(x_0 \in y_0 \Leftrightarrow y_0 \in z_0) \wedge (-x_0 \in y_0 \Leftrightarrow -y_0 \in z_0)$$

is true. Therefore (19) holds again.

Claim. (20)$\vee$(21) is true.

Proof. It suffices to show that the negation of (20)$\vee$(21) leads to a contradiction. The negation of (20)$\vee$(21) is the conjunction of

$$(\forall y \in X)(\forall x \in A_1^u)(y \notin z_0 \ \vee \ x \notin y \ \vee \ -x \notin y) \tag{22}$$

and

$$(\forall y \in X)(\forall x \in A_1^u)(y \in z_0 \ \vee \ x \in y \ \vee \ -x \in y). \tag{23}$$

We can rewrite (22) and (23) as follows:

$$(\forall y \in X \cap z_0)(\forall x \in A_1^u)(x \in y \ \Rightarrow \ -x \notin y) \tag{24}$$

and

$$(\forall y \in X \cap -z_0)(\forall x \in A_1^u)(x \notin y \ \Rightarrow \ -x \in y). \tag{25}$$

(24) and (25) say that z_0 partitions X into a part $X \cap z_0$ consisting of "consistent" y only (i.e., not containing both x and $-x$ for any x), and a part $X \cap -z_0$ consisting of "complete" y only, (i.e., containing at least one of the x and $-x$ for every x). But this is false, since X must contain also y which are neither

consistent nor complete. To show this, we make use of the assumption that $\mathcal{A}$ is the level collapse of an elementary submodel of a full model. Actually this is the only point in the proof where we use this strong hypothesis. (For all other purposes, regularity and richness suffice.) Indeed let $\mathcal{A} = lc(\mathcal{B})$, for a countable $\mathcal{B} \preceq \langle\langle D\rangle\rangle$. In $\langle\langle D\rangle\rangle$, the corresponding set X contains sets that are neither consistent nor complete. Namely in $\langle\langle D\rangle\rangle$, using possibly the choice axiom of the ground model, we can find a $y^* \in \mathcal{P}^2(D)$ such that (a) $\{\emptyset, D\} \subseteq y^*$, (b) $|y^*| = |-y^*|$, (c) $\exists x_1(x_1 \in y^* \wedge -x_1 \notin y^*)$ (non-closed), (d) $\exists x_2(x_2 \in y^* \wedge -x_2 \in y^*)$ (non-consistent) and (e) $\exists x_3(x_3 \notin y^* \wedge -x_3 \notin y^*)$ (non-complete). Since $\mathcal{A} = lc(\mathcal{B})$ and $\mathcal{B}$ is countable, by lemma 1.2 there is a y in $\mathcal{A}$ with properties (a)-(e) above. Therefore (24) and (25) lead to a contradiction. This contradiction shows that (22)$\wedge$(23) is false, and hence (20)$\vee$(21) is true. This completes the proof of the Claim.

Thus $o^{\mathcal{A}}$ extends at z_0 in all cases, and this completes the proof of the lemma. $\dashv$

By lemma 3.3, the level collapse of an elementary submodel of a full model is rich. So from lemmas 3.4 and 3.5 we derive immediately the following:

Theorem 3.6 *Let $\langle\langle D\rangle\rangle$ be a full model of* TST *(with infinite D) and let $\mathcal{A}$ be the level collapse of a countable elementary submodel of $\langle\langle D\rangle\rangle$. Then the pair $o^{\mathcal{A}}$ is extendible in $\mathcal{A}$.*

4 2-Extendibility.

What about 2- (or higher) extendible coherent pairs? Again the simplest question is: Is there a model $\mathcal{A}$ of TST such that $o^{\mathcal{A}}$ is 2-extendible?

2-extendibility (as well as n-extendibility in general) splits again to A_i-2-extendibility, for $i = 1, 2, 3$. Given $\mathcal{A}$, $o^{\mathcal{A}}$ is A_1-2-extendible if

$$(\forall x \in A_1)(\exists q)(q \leq o^{\mathcal{A}} \wedge x \in dom(q_1) \wedge q \text{ is extendible}).$$

The last formula when decompressed becomes:

ψ_1: $\{\forall x \in A_1\ \exists y \in A_2\ \exists z \in A_3$ such that, if $v_1 = \{\emptyset, A_0, x, -x\}$, $v_2 = \{\emptyset, A_1, y, -y\}$, $v_3 = \{\emptyset, A_2, z, -z\}$, then the pair $v_1 \xrightarrow{q_1} v_2 \xrightarrow{q_2} v_3$, where q_1 and q_2 are the mappings which preserve the orderings of v_1, v_2, v_3 as exhibited above, is not only a coherent pair but also an extendible one.$\}$

For A_2-2- and A_3-2-extendibility of $o^{\mathcal{A}}$ the formulation is similar, except of the string of initial quantifiers. That is, A_2-2-extendibility is written:

ψ_2: $\{\forall y \in A_2\ \exists x \in A_1\ \exists z \in A_3$ such that, if $v_1 = \{\emptyset, A_0, x, -x\}$, $v_2 = \{\emptyset, A_1, y, -y\}$, $v_3 = \{\emptyset, A_2, z, -z\}$, then $v_1 \xrightarrow{q_1} v_2 \xrightarrow{q_2} v_3$ is an extendible pair.$\}$

And A_3-2-extendibility is written:

ψ_3: $\{\forall z \in A_3\ \exists x \in A_1\ \exists y \in A_2$ such that, if $v_1 = \{\emptyset, A_0, x, -x\}$, $v_2 = \{\emptyset, A_1, y, -y\}$, $v_3 = \{\emptyset, A_2, z, -z\}$, then $v_1 \xrightarrow{q_1} v_2 \xrightarrow{q_2} v_3$ is an extendible pair.$\}$

Thus 2-extendibility of $o^{\mathcal{A}}$ amounts to the truth of the formula $\psi_1 \wedge \psi_2 \wedge \psi_3$. At first sight it looks overwhelmingly involved, but if we find conditions deciding when an arbitrary pair $v_1 \xrightarrow{q_1} v_2 \xrightarrow{q_2} v_3$ is extendible, things may become easier.

Acknowledgement. I would like to thank M. Randall Holmes for reading and commenting on a draft of this paper.

References

[1] T.E. Forster, *Set theory with a universal set*, Oxford Logic Guides Vol. 20, Oxford U.P. 1992.

[2] V.N. Grishin, Consistency of a fragment of Quine's NF system, *Soviet Mathematics Doklady* **10**, No 6 (1969), 1387-1390.

[3] V.N. Grishin, *The method of stratification in set theory* (in Russian), Ph.D. Thesis, Moscow University 1972.

[4] E. Specker, Typical ambiguity, *Logic Methodology and Philosophy of Science*, Proceedings of the International Congress, Stanford, 1960, Stanford University Press 1962, pp. 116-124.