Annals of Pure and Applied Logic 57 (1992) 1–26
North-Holland

Complete topoi representing models of set theory

Andreas Blass
Mathematics Department, University of Michigan, Ann Arbor, MI 48109, United States

Andre Scedrov
Mathematics Department, University of Pennsylvania, Philadelphia, PA 19104, United States

Communicated by T. Jech
Received 6 February 1991

Abstract

Blass, A. and A. Scedrov, Complete topoi representing models of set theory, Annals of Pure and Applied Logic 57 (1992) 1–26.

By a model of set theory we mean a Boolean-valued model of Zermelo–Fraenkel set theory allowing atoms (ZFA), which contains a copy of the ordinary universe of (two-valued, pure) sets as a transitive subclass; examples include Scott–Solovay Boolean-valued models and their symmetric submodels, as well as Fraenkel–Mostowski permutation models. Any such model M can be regarded as a topos. A logical subtopos $\mathscr{E}$ of M is said to *represent* M if it is complete and its cumulative hierarchy, as defined by Fourman and Hayashi, coincides with the usual cumulative hierarchy of M. We show that, although M need not be a complete topos, it has a smallest complete representing subtopos, and we describe this subtopos in terms of definability in M. We characterize, again in terms of definability, those models M whose smallest representing topos is a Grothendieck topos. Finally, we discuss the extent to which a model can be reconstructed when its smallest representing topos is given.

1. Introduction

We shall be concerned with models of set theory that are extensions of the ordinary universe V of (pure) sets. Among these extensions are the universes $V(A)$ built from a collection of atoms, the Fraenkel–Mostowski permutation submodels of such universes, the Boolean-valued models $V^{\mathscr{B}}$ of Scott and Solovay, and their symmetric submodels. Any model M of set theory gives rise to a topos, also called M, whose objects are the sets of M and whose morphisms $x \to y$ are those $f \in M$ that satisfy in M the formula 'f is a function from x to y'. The topoi obtained in this manner need not be complete, but Fourman [2] (see

0168-0072/92/$05.00

also Hayashi [3]) showed how to associate, to many models M of set theory, Grothendieck topoi $\mathscr{E}$ ('sheaf models') that are essentially equivalent to M. The concept of 'essentially equivalent' used in [2, 3] was that the first-order sentences in the language of set theory that are true in M are precisely those satisfied by a certain interpretation, which we call the Fourman–Hayashi interpretation, of the language of set theory given by $\mathscr{E}$. In fact, Fourman's and Hayashi's topoi $\mathscr{E}$ are equivalent to the corresponding models M in a stronger sense, first isolated in [1] and there expressed as '$\mathscr{E}$ represents M'.

The purpose of this paper is to clarify the connection between models of set theory and complete topoi by analyzing the relation '$\mathscr{E}$ represents M'. One of our main results exhibits, for each model M of set theory extending V, a canonical, smallest, complete topos representing M. We also discuss whether and how a model M can be recovered from a representing topos.

2. Models of set theory

The models of set theory that we shall work with are Boolean-valued models of ZFA that extend the universe V of all pure sets. We begin by spelling out what this means.

ZFA is the first-order theory, in the language with a binary predicate symbol $\in$ and a constant symbol A, that is like Zermelo–Fraenkel set theory (ZF) except for allowing a set A of atoms. A Boolean-valued model M of ZFA consists of a class M, a complete Boolean algebra $\mathscr{B}$, a distinguished element A (or A_M when there is danger of ambiguity), and two binary operations $\|x = y\|$ and $\|x \in y\|$ (also written with subscript M when necessary) on M with values in $\mathscr{B}$, such that the ZFA axioms have truth value 1, where truth values of formulas are defined inductively, starting with $\|x = y\|$ and $\|x \in y\|$ for atomic formulas and using the finitary (respectively, infinitary) Boolean operations of $\mathscr{B}$ for the propositional connectives (respectively, quantifiers). We also require that M satisfy the following 'patching principle'. If b_i ($i \in I$) are elements of $\mathscr{B}$ and x_i ($i \in I$) are elements of M with $\|x_i = x_j\| \geq b_i \wedge b_j$ for all $i, j \in I$, then there exists $x \in M$ with $\|x = x_i\| \geq b_i$ for all $i \in I$. Any M not satisfying this patching principle can be enlarged to one that does satisfy it, without altering any truth values; just adjoin all possible 'patched together' elements. We also identify any two elements $x, y \in M$ for which $\|x = y\| = 1$.

To say that M is an extension of V means that there is a mapping $V \to M$, written $x \mapsto \check{x}$, such that, for all $x \in V$ and $z \in M$,

$$\|z = \check{x}\| = \bigvee_{y \in x} \|z = \check{y}\|.$$

By the convention at the end of the preceding paragraph, this equation uniquely determines $\check{x}$ by recursion on x; we are demanding that such an $\check{x}$ exists for every

x. We also require that M contain no extra ordinals after those of V; that is, for all $z \in M$,

$$\|z \text{ is an ordinal}\| = \bigvee_{\alpha \text{ an ordinal of } V} \|z = \check{\alpha}\|.$$

(The requirement here is $\leqslant$; it is easy to check that $\geqslant$ follows from the definition of $\check{}$.)

Finally, for technical reasons, we impose some smallness requirements. Although M can (indeed must) be a proper class, we insist that the Boolean algebra $\mathscr{B}$ of truth values be a set and that there be only a set of $x \in M$ such that $\|x \in A \text{ or } x = \emptyset\| = 1$. It then follows, by induction on α, that there are only a set of $x \in M$ with $\|x \text{ has rank } \leqslant \check{\alpha}\| = 1$ and indeed that, for each $b \in \mathscr{B}$, there are only a set of $x \in M$ with $\|x \text{ has rank } \leqslant \check{\alpha}\| \geqslant b$ modulo the equivalence relation $\|x = x'\| \geqslant b$. Using the facts that $\|(\forall y)\, y \text{ has an ordinal rank}\| = 1$ and M has no extra ordinals, we infer that, for every $y \in M$ and every $b \in \mathscr{B}$, there are only a set of $x \in M$ with $\|x \in y\| \geqslant b$, modulo the equivalence $\|x = x'\| \geqslant b$.

Henceforth, when we refer to models of set theory, we shall mean models satisfying all the requirements set forth above. We observe that the Scott–Solovay Boolean-valued models $V^{\mathscr{B}}$ and their symmetric submodels (as defined in, e.g., [4, Chapter 5]) are models of set theory in this sense, with $A = \emptyset$. Also, the universes $V(A)$ built from a set A of atoms, as well as their permutation submodels (as in [4, Chapter 4]), are models of set theory with the two-element Boolean algebra as $\mathscr{B}$. Examples with both A and $\mathscr{B}$ nontrivial are obtainable as Boolean-valued extensions of $V(A)$ or submodels thereof.

Any model M of set theory gives rise to a category, also called M, as follows. Objects are the sets of M, those $x \in M$ which are definitely not atoms, i.e., for which $\|x \in A\| = 0$. Morphisms from x to y are those $f \in M$ for which $\|f \text{ is a function from } x \text{ to } y\| = 1$. The composite of f and g is the unique h such that $\|h = f \circ g\| = 1$, where $\circ$ refers to the usual definition of composition of functions in set theory. Note that the patching property of M is needed to ensure the existence of h as well as the existence of identity morphisms. It is straightforward to check (using patching repeatedly) that the category M is a topos. For example, the power object of x is the unique y such that $\|y \text{ is the set of all subsets of } x\| = 1$. In particular, the truth value object of this topos is $\check{2}$, whose global sections (elements $x \in M$ with $\|x = \check{0} \text{ or } x = \check{1}\| = 1$) are identified with the elements of $\mathscr{B}$ (x being identified with $\|x = \check{1}\|$). M has a natural numbers object, namely $\check{\omega}$. Our smallness requirements on M ensure, as indicated above, that each object x has only a set of points $1 \to x$; applying this with $x = a^b$, we find that there are only a set of morphisms $b \to a$, so the category M has small hom-sets.

In general, the topos M need not be complete; in fact, the coproduct of countably many copies of $\check{1}$ need not exist. To see this, note first that, if this coproduct exists, then it is the natural numbers object $\check{\omega}$. But then, given any sequence (in the real world) of elements $x_n \in M$, there would have to be an $f \in M$

with $\|f$ is a function on $\check{\omega}\| = 1$ and $\|f(\check{n}) = x_n\| = 1$ for all n. That is, M would be closed under ω-sequences. But there are plenty of models that are not closed under ω-sequences, for example, the basic Fraenkel model [4, p. 48] or the basic Cohen model [4, p. 66] or indeed any model violating the countable axiom of choice.

3. Representing topoi

Fourman [2] and Hayashi [3] observed that, if $\mathscr{E}$ is a complete topos and A is an object of $\mathscr{E}$, then one can define a cumulative hierarchy $V_\alpha(A)$ over A by

$$V_0(A) = A, \qquad V_{\alpha+1}(A) = A + \text{power set of } V_\alpha(A),$$

and, for limit ordinals λ,

$$V_\lambda(A) = \varinjlim_{\alpha<\lambda} V_\alpha(A),$$

where the colimit is taken with respect to transition maps defined by recursion simultaneously with the objects $V_\alpha(A)$. Fourman and Hayashi used this hierarchy to interpret the language of ZFA in $\mathscr{E}$ (using the object A of $\mathscr{E}$ to interpret the constant symbol A of ZFA). The essential idea is that quantifiers restricted to objects of $\mathscr{E}$ can be interpreted as in the internal logic of $\mathscr{E}$, and unrestricted quantifiers can be interpreted as if they were restricted to $V_\alpha(A)$ for large enough α. This works because hom-sets in $\mathscr{E}$ are small, so restricting a quantifier to $V_\alpha(A)$ produces a truth value that is independent of α once α is large enough.

As in [1], we say that a complete topos $\mathscr{E}$ *represents* a model M of set theory if $\mathscr{E}$ is a logical subtopos of M (i.e., a subcategory such that the inclusion functor $\mathscr{E} \hookrightarrow \mathrm{M}$ preserves the topos structure), $\mathscr{E}$ contains the object A (the set of atoms) of M, and the Fourman–Hayashi hierarchy $V_\alpha(A)$ over A in $\mathscr{E}$ coincides with the usual von Neumann hierarchy in M (i.e., $\|V_\alpha(A)$ is the set of all atoms and all sets of rank $<\check{\alpha}\| = 1$).

To clarify this concept, we consider a fairly typical example. Let M be the basic Fraenkel model, defined as follows. Let $V(A)$ be the universe built from a countable set A of atoms by iterating the power set operation through all the ordinals. The group $\mathscr{G}$ of all permutations of A acts on $V(A)$ as a group of automorphisms. An element of $V(A)$ is called *symmetric* if it is invariant under a subgroup of $\mathscr{G}$ of the form

$$\mathrm{Fix}(F) = \{\pi \in \mathscr{G} \mid \pi(a) = a \text{ for all } a \in F\},$$

for some finite $F \subseteq A$. Then M consists of the hereditarily symmetric elements of $V(A)$, i.e., those $x \in V(A)$ such that x, its members, their members, etc. are all symmetric. It is well known that M is a model of ZFA in which the countable axiom of choice fails. By the remarks at the end of the last section, the topos M is not complete.

Let $\mathscr{E}$ be the subcategory of M consisting of those objects and morphisms that are invariant under the whole group $\mathscr{G}$, not just some $\mathrm{Fix}(F)$. Then $\mathscr{E}$ is a logical subtopos of M, because the topos structure of M is preserved by the action of $\mathscr{G}$. Furthermore, unlike M, $\mathscr{E}$ is complete. The coproduct of any family $\mathscr{E}$-objects X_i, indexed by a pure set I, can be taken to be the disjoint union

$$\{(i, x) \mid i \in I \text{ and } x \in X_i\},$$

because, if $f_i : X_i \to Y$ are $\mathscr{G}$-invariant, then the induced function f on the disjoint union, given by $f(i, x) = f_i(x)$, is also $\mathscr{G}$-invariant. (Contrast this with the situation in M. If each f_i is $\mathrm{Fix}(F_i)$-invariant, then f need only be $\mathrm{Fix}(\bigcup_i F_i)$-invariant, so if $\bigcup_i F_i$ is infinite, then f need not be symmetric.) Products can be handled similarly in $\mathscr{E}$. Finally, it is straightforward to check that in $\mathscr{E}$ (unlike M) the set of atoms and sets of rank $<\lambda$ is, for limit λ, the colimit of the corresponding sets with λ replaced by $\alpha < \lambda$. This provides the limit stages in an inductive proof that the von Neumann hierarchy of M coincides with the Fourman–Hayashi hierarchy over A in $\mathscr{E}$; the successor stages of the induction are automatic since both hierarchies use power sets at successor stages.

These considerations show that $\mathscr{E}$ is a complete topos representing M. Other examples, with the same M, can be obtained by replacing $\mathscr{G}$ with a suitable subgroup. For example, choose one particular atom $a_0 \in A$, let $\mathscr{G}_0 = \mathrm{Fix}\{a_0\}$, and let $\mathscr{E}'$ consist of the objects and morphisms of M that are invariant under $\mathscr{G}_0$. Then $\mathscr{E}'$ is also a complete topos representing M.

As an abstract topos, $\mathscr{E}'$ is equivalent to $\mathscr{E}$. Indeed, $\mathscr{E}$ is the topos of continuous $\mathscr{G}$-sets (where $\mathscr{G}$ has the topology generated by the subgroups $\mathrm{Fix}(F)$ for finite $F \subseteq A$), $\mathscr{E}'$ is the topos of continuous $\mathscr{G}_0$-sets, and $\mathscr{G}$ is isomorphic to $\mathscr{G}_0$. As subtopoi of M, however, $\mathscr{E}$ and $\mathscr{E}'$ are quite different. For example, in $\mathscr{E}$ the object A has only two subobjects, namely $\emptyset$ and A, but in $\mathscr{E}'$ there are also $\{a_0\}$ and its complement.

As subtopoi of M, $\mathscr{E}$ and $\mathscr{E}'$ are related as follows. $\mathscr{E}'$ can be obtained from $\mathscr{E}$ by adjoining a point in A, namely the morphism $\check{1} \to A$ with value a_0. More precisely, the slice topos $\mathscr{E}/A$ is equivalent to $\mathscr{E}'$ via the functor sending $X \xrightarrow{f} A$ to $f^{-1}\{a_0\}$, and this equivalence fits into a commutative diagram

$$\begin{array}{ccc} \mathscr{E}/A & \cong & \mathscr{E}' \\ \uparrow & & \downarrow \\ \mathscr{E} & \hookrightarrow & M \end{array}$$

where the functor $\mathscr{E} \to \mathscr{E}/A$ is the standard embedding sending X to the projection $X \times A \to A$.

The idea that led to $\mathscr{E}'$ can be iterated. Choose n distinct atoms $a_0, \ldots, a_{n-1}$, and let $\mathscr{E}^{(n)}$ be the subtopos of M consisting of objects and morphisms invariant under $\mathrm{Fix}\{a_0, a_1, \ldots, a_{n-1}\}$. Then $\mathscr{E}^{(n)}$ is a complete topos representing M. M is the colimit of the system of logical inclusions

$$\mathscr{E} \hookrightarrow \mathscr{E}' \hookrightarrow \cdots \hookrightarrow \mathscr{E}^{(n)} \hookrightarrow {}^{(n+1)} \hookrightarrow \cdots,$$

in which each topos is equivalent to a slice topos of its predecessor,

$$\mathscr{E}^{(n+1)} \cong \mathscr{E}^{(n)}/(A - \{a_0, a_1, \ldots, a_{n-1}\}).$$

The preceding discussion of the basic Fraenkel model has an analog for the basic Cohen model. This model consists of those elements of the Boolean-valued model $V^{\mathscr{B}}$, with $\aleph_0$ Cohen reals, that are hereditarily symmetric with respect to the group $\mathscr{G}$ of automorphisms of $\mathscr{B}$ that permute the $\aleph_0$ Cohen reals. This model is not a complete topos, but for each $n \in \omega$ the subtopos $\mathscr{E}^{(n)}$ of objects and morphisms invariant under all permutations fixing the first n Cohen reals is a complete topos representing the model. The model is the colimit of the sequence of these subtopoi, and in this sequence each $\mathscr{E}^{(n+1)}$ is a slice topos of its predecessor $\mathscr{E}^{(n)}$. One difference between the basic Franekel model and the basic Cohen model is that for the former any representing subtopos must, by definition, contain the set A of atoms, but for the latter a representing subtopos need not contain the specific set of Cohen reals that were adjoined to form $V^{\mathscr{B}}$. For example, we could let $\mathscr{G}^{+}$ be the group of automorphisms of $\mathscr{B}$ generated by the $\mathscr{G}$ above plus the automorphism that replaces each of the Cohen reals (viewed as a subset of ω) with its complement. Then the $\mathscr{G}^{+}$-invariant objects and morphisms of M constitute a representing subtopos properly included in $\mathscr{E}^{(0)}$. (The crucial property of $\mathscr{G}^{+}$ that makes this work is that it normalizes the filter of subgroups of $\mathscr{G}$ used in defining symmetry and thus in defining M.)

As mentioned in the Introduction, 'representing', as a notion of 'equivalence' between a topos $\mathscr{E}$ and a model M, is stronger than the equivalence relation of 'matching truth values' used by Fourman and Hayashi [2, 3]. Indeed, if $\mathscr{E}$ represents M, then the inclusion $\mathscr{E} \hookrightarrow M$ preserves the interpretations of membership, equality, logical connectives and bounded quantifiers (as it is a logical morphism) as well as the hierarchies (of Fourman and Hayashi in $\mathscr{E}$ and of von Neumann in M). Furthermore, the interpretation of unbounded quantification as a limit of bounded quantification over $V_\alpha(A)$ as α increases, which is a matter of definition in the Fourman–Hayashi interpretation, is also correct in M because it is provable in ZFA that the $V_\alpha(A)$ exhaust the universe. Thus, the inclusion $\mathscr{E} \hookrightarrow M$ preserves all truth values.

4. Minimal representing topoi

In this section, we shall construct, in every model M of set theory, a smallest representing subtopos. We begin by considering what elements of M must be in every representing topos $\mathscr{E}$.

First, the definition of 'representing' explicitly requires that, for each ordinal α, the unique element $x \in M$ satisfying $\|x$ is the set of all atoms and all sets of rank $< \alpha\| = 1$ must be in $\mathscr{E}$; indeed, this x must be the object $V_\alpha(A)$ of the Fourman–Hayashi hierarchy over A in $\mathscr{E}$, and we henceforth refer to it as $V_\alpha(A)$, even when we are discussing M and not $\mathscr{E}$.

Second, we shall see that, for every $x \in V$, $\check{x}$ is in $\mathscr{E}$. Before proving this, however, we must sort out two possible meanings of '$\check{x}$ is in $\mathscr{E}$'. Suppose x has rank α, so $\check{x} \subseteq V_\alpha(A)$ and $\check{x} \in V_{\alpha+1}(A)$ with truth value 1. (The A's here are unnecessary but harmless, as x is a pure set.) We can ask that $\check{x}$ be 'in $\mathscr{E}$' as a subobject of $V_\alpha(A)$ or as a point $\check{1} \to V_{\alpha+1}(A)$. That these two meanings are equivalent follows from the fact that $V_{\alpha+1}(A)$ is A + power set of $V_\alpha(A)$. Note that it then also follows that $\check{x}$ is in $\mathscr{E}$ as a point $\check{1} \to V_\beta(A)$ for all $\beta > \alpha$.

Now to prove that $\check{x}$ is in $\mathscr{E}$ (in any or all of these senses), we proceed by induction on the rank α of x. Suppose then that every $\check{y}$, for $y \in x$, is in $\mathscr{E}$ as a morphism $\check{y}: \check{1} \to V_\alpha(A)$ (since y has rank $<\alpha$). Let C be the coproduct of $\check{1}$'s indexed by x (which exists as $\mathscr{E}$ is complete), and let $f: C \to V_\alpha(A)$ be the morphism that agrees with $\check{y}$ on the yth copy of $\check{1}$ in C. Then f is $\check{x}$ as a subobject of $V_\alpha(A)$. Thus $\check{x}$ is in $\mathscr{E}$ in one, hence all, of the senses described in the preceding paragraph.

Third, consider any element $z \in M$ that is definable over $V \cup \{A\}$ in the following sense. There is an α with $\|z \in V_\alpha(A)\| = 1$, and there is an $x \in V$ of rank $<\alpha$, and there is a formula $\varphi(v_0, v_1)$ in the language of ZFA such that $\|z$ is the unique element of $V_\alpha(A)$ satisfying $\varphi(z, \check{x})$ in $V_\alpha(A)\| = 1$. Then z, as a morphism $1 \to V_\alpha(A)$, is in $\mathscr{E}$ because it is definable in the internal logic of $\mathscr{E}$, using φ with all bound variables viewed as having sort $V_\alpha(A)$ and with v_1 substituted by $\check{x}: 1 \to V_\alpha(A)$. It follows, of course, that z is also in $\mathscr{E}$ as a subobject of $V_\alpha(A)$.

Finally, suppose $z \in M$ and z is definable in the following sense. There is $x \in V$ and there is a formula $\varphi(v_0, v_1)$ in the language of ZFA such that $\|z$ is the unique element satisfying $\varphi(z, \check{x})\| = 1$. Then, by the reflection principle of ZFA, z is also definable in the sense of the preceding paragraph, hence is in $\mathscr{E}$. The converse also holds; if z is definable in the sense of the preceding paragraph, then also in the sense of the present paragraph, since $V_\alpha(A)$ is definable from $\check{\alpha}$. We refer to these two equivalent notions of definability as *V-definability* since parameters of the form $\check{x}$ for $x \in V$ are allowed; notice that the set A of atoms is also allowed as a parameter, since the defining formulas are in the language of ZFA which has a constant symbol for A.

What we have done here closely resembles the familiar notion of ordinal definability (see [4, p. 42] and [6]) with additional parameters allowed, so a few words may be in order to clarify the connection and prevent possible confusion. We can define a unary predicate, or class, $\check{V}$ of 'standard pure sets' in M by

$$\|z \in \check{V}\| = \bigvee_{x \in V} \|z = \check{x}\|.$$

Then the formula 'z is ordinal definable with parameters from $\check{V} \cup \{A\}$' makes sense, i.e., has a truth value for each $z \in M$. If z is V-definable, then this truth value is clearly 1. But the converse can fail. The point is that z may be definable using a parameter p which equals various $\check{x}$'s with various truth values. For example, if $b \in \mathscr{B}$ is any truth value, then the unique z with $\|z = \check{0}\| = \neg b$ and $\|z = \check{1}\| = b$ is certainly ordinal definable over $\check{V} \cup \{A\}$ with truth value 1, being an ordinal with truth value 1. But this z will be V-definable only if $b = \|\varphi(\check{x})\|$ for some ZFA-formula $\varphi(v)$ and some $x \in V$, and this will not be the case if, for example, M is $V^{\mathscr{B}}$ and b is moved by some automorphism of $\mathscr{B}$. This example also shows, since $\check{0}$ and $\check{1}$ are V-definable and $\|z = \check{0}$ or $z = \check{1}\| = 1$, that V-definability does not coincide with any internal concept in M. (Of course, the distinction between V-definability and ordinal definability over $\check{V} \cup \{A\}$ depends upon the nontriviality of $\mathscr{B}$. If M is two-valued, then the distinction disappears.)

Theorem 1. *Let M be any model of set theory. Its V-definable objects and morphisms constitute the smallest (up to equivalence) complete subtopos representing M.*

Proof. We have already seen that, if $\mathscr{E}$ represents M and x is a V-definable element of M, then x is in $\mathscr{E}$ as a subobject of some $V_\alpha(A)$. If we assume, as we may if we care about categories only up to equivalence, that $\mathscr{E}$ is closed under isomorphisms, then x itself is an object of $\mathscr{E}$ and its inclusion map into $V_\alpha(A)$ is a morphism of $\mathscr{E}$. If $f: x \to y$ is V-definable, then, since f, viewed as a point $1 \to V_\alpha(A)$, is in $\mathscr{E}$ and f viewed as a morphism $x \to y$ is internally definable from this point, it follows that f is in $\mathscr{E}$ as a morphism $x \to y$. Thus, any representing subtopos $\mathscr{E}$ of M must include the category of V-definable objects and morphisms of M. It remains to prove that this category is itself a representing subtopos of M.

That it is a logical subtopos is easy to see. The product of two V-definable sets and the power set of a V-definable set are clearly also V-definable along with the relevant structure (projections for the products, membership relation for the power sets), and the adjointnesses clearly preserve V-definability.

To see that the topos of V-definable sets is complete (even if M is not), we need some preliminary considerations about V-definability. Specifically, we need a version, appropriate for the present context, of the theorem of [6] that there is a single formula that suffices for all the definitions involved in the notion of ordinal definability. Let $\delta(v_0, v_1)$ be a formula, in the language of ZFA, expressing 'v_1 is an ordered triple $\langle x, y, z\rangle$ such that x is a formula in the language of ZFA with just two free variables, y is an ordinal, and v_0 is the unique object such that the formula x becomes true when its free variables are interpreted as v_0 and z and its bound variables are interpreted as ranging over atoms and sets of rank $<y$'. Notice that, if u is V-definable in M as the unique solution of $\varphi(u, \check{p})$ in $V_\alpha(A)$, then the same u is also definable as the unique solution of $\delta(u, \langle\varphi, \alpha, p\rangle^\vee)$ in M or in $V_\beta(A)$ for any $\beta > \alpha$. Thus, in the concept of V-definability, we may assume that the defining formula is always δ and that only the parameters vary. Note also that the same definition works in all sufficiently large $V_\beta(A)$.

We now show that the topos of V-definable sets in M is complete by constructing a product for an arbitrary indexed family x_i ($i \in I$) of objects, where $I \in V$. (Cocompleteness can be proved similarly or deduced from completeness via [7].) By the preceding discussion (and the axiom of choice in the metatheory, i.e., in V), we can fix an indexed family of parameters $p = (p_i)_{i\in I} \in V$ such that, for each i, x_i is (with truth value 1 in M) the unique solution of $\delta(x_i, \check{p}_i)$. Now let Π be (with truth value 1 in M) the set of all functions f with domain $\check{I}$ such that, for every $i \in \check{I}$, $f(i)$ is a member of the unique set x satisfying $\delta(x$, the ith component of $\check{p}$). The idea here is simply that Π is the product, in M, of the family of x_i's. The somewhat awkward definition is needed because it is not clear a priori that there is such a family (or product) in M; although each x_i is in M, the family was given in V. The definition shows also that Π is V-definable, the parameters in its definition being $\check{I}$ and $\check{p}$. It is clear that, for each $i \in I$, the projection map $\Pi \to x_i$ sending f to (the element of M that is with truth value 1) $f(\check{i})$ is V-definable (using $\check{i}$ as an additional parameter). Finally, to check that Π with these projections is the product of the x_i's, let V-definable maps $g_i : y \to x_i$ be given. Again, use choice in V to fix a system $q = (q_i)_{i\in I} \in V$ such that g_i is defined by δ with parameter q_i. Then the unique map $g : y \to \Pi$ whose composite with the ith projection is g_i is given by the following definition: g is the function with domain y such that, for each $z \in y$, $g(z)$ is the element of Π whose value at any $i \in \check{I}$ is the same as the value at z of the unique function h satisfying $\delta(h$, the ith component of $\check{q}$). This definition shows that g is V-definable and thus completes the proof that the topos of V-definable sets is complete.

To finish the proof that the topos of V-definable sets represents M, we must show that it contains A and that its Fourman–Hayashi hierarchy over A coincides with the von Neumann hierarchy in M. That A is V-definable is obvious, since the defining formulas are allowed to mention A. That the Fourman–Hayashi and von Neumann hierarchies agree is proved by transfinite induction. Both hierarchies begin with A at level 0 and form A plus the power set of the previous stage at successor stages, so the crux of the proof concerns limit stages. Writing, as above, $V_\alpha(A)$ for the αth level of the von Neumann hierarchy, i.e., the (element of M that is with truth value 1) the set of atoms and sets of rank $<\check{\alpha}$, we must show that, for each limit ordinal λ, $V_\lambda(A)$ is the colimit of the system of $V_\alpha(A)$, for $\alpha < \lambda$, with the obvious inclusion maps.

The uniqueness clause in the definition of colimit is easy to check and has nothing to do with V-definability. Indeed, if g and h are distinct morphisms $V_\lambda(A) \to y$ in M, then the truth value

$$\begin{aligned}\|g \neq h\| &= \|\exists x\,(x \text{ an atom or a set of rank } <\check{\lambda}, \text{ and } g(x) \neq h(x))\| \\ &= \|\exists \alpha < \check{\lambda}\ \exists x\,(x \text{ an atom or a set of rank } <\alpha, \text{ and } g(x) \neq h(x))\| \\ &= \bigvee_{\alpha<\lambda} \|g \mid V_\alpha(A) \neq h \mid V_\alpha(A)\|\end{aligned}$$

is nonzero. Then, for some $\alpha<\lambda$, $\|g \mid V_\alpha(A) \neq h \mid V_\alpha(A)\| \neq 0$, so g and h have distinct restrictions to $V_\alpha(A)$.

It remains to verify the existence clause in the definition of colimit; this verification is similar to the earlier verification that Π is a product. Let V-definable maps $g_\alpha : V_\alpha(A) \to y$ be given for $\alpha<\lambda$, with $g = g_\beta \mid V_\alpha(A)$ whenever $\alpha<\beta<\lambda$. Using choice in V, fix $p=(p_\alpha)_{\alpha<\lambda} \in V$ such that g_α is defined by δ with parameter $\check{p}_\alpha$. Then we can define a map $g : V_\lambda(A) \to y$ having the g_α's as restrictions, as follows. For each $z \in V_\lambda(A)$ and each $\alpha<\check{\lambda}$, if z is an atom or has rank $<\alpha$, then $g(z)$ equals the value at z of the unique h such that $\delta(h$, the αth component of $\check{p}$). It is easily verified that this defines a unique g, with the correct restrictions, and it is clear that g is V-definable. This completes the verification that $V_\lambda(A)$ is the colimit of the previous $V_\alpha(A)$'s in the topos of V-definable sets and thus also the proof that this topos represents M. □

The topos $\mathscr{E}$ of V-definable objects and morphisms of M, being complete, admits a (unique up to isomorphism) geometric morphism to the topos V (see [5, p. 119]). We note for future reference that, for any set I, its image in $\mathscr{E}$ under the left-adjoint part of this geometric morphism, namely the coproduct of an I-indexed family of copies of $\check{1}$, can be taken to be $\check{I}$, with injections $\check{1} \xrightarrow{i} \check{I}$ given by the points $i : 1 \to I$.

5. Relative definability

To analyze further the minimal representing topoi obtained in the preceding section, we need some information about relative V-definability. The present section is devoted to this information; it will be applied to representing topoi in subsequent sections. Let M be a model of set theory, and let x and y be elements of M. We say that x is *V-definable from* y and we write $x \leqslant y$ if there is a V-definable function f such that $\|f(y)=x\|=1$. An equivalent statement is that x is the unique element of M satisfying (with truth value 1) some formula with y and some $\check{p}$ as parameters. As in our earlier discussion of V-definability, it does not matter whether the formula is interpreted in M or in some sufficiently large $V_\alpha(A)$, and a single formula can be used for all definitions (with only the parameter $\check{p}$ varying). Clearly, $\leqslant$ is a pre-ordering; the associated equivalence relation $x \leqslant y \leqslant x$ is written $x \equiv y$, and the equivalence classes $[x]$ are called *V-degrees*. (For technical reasons, we use Scott's trick and define $[x]$ to be the set of y of least rank satisfying $x \equiv y$; thus $[x]$ is a set rather than a proper class.) The pre-order $\leqslant$ induces a partial order, also written $\leqslant$, on the V-degrees. The V-degrees form an upper semilattice, the least upper bound of $[x]$ and $[y]$ being $[(x, y)]$, and the V-definable sets constitute the smallest V-degree.

For the further study of V-degrees, we use the following result, which seems to be of some independent interest.

Theorem 2. *For each* $x \in M$, *there is a smallest* V*-definable* d *such that* $\|x \in d\| = 1$; *here 'smallest' means that, for any other* V*-definable* d' *with* $\|x \in d'\| = 1$, $\|d \subseteq d'\| = 1$.

The proof will be simpler if we first adopt the following conventions. Recall that $\delta(v_0, v_1)$ is a fixed formula such that, as p ranges over V, $\delta(v_0, \check{p})$ defines all V-definable sets in M. We can easily arrange that $\|\forall v_1 \exists! v_0 \, \delta(v_0, v_1)\| = 1$, simply by replacing $\delta(v_0, v_1)$ with

$$[\delta(v_0, v_1) \wedge \exists! v_0 \, \delta(v_0, v_1)] \vee [v_0 = 0 \wedge \neg \exists! v_0 \, \delta(v_0, v_1)].$$

Then let Def be the operation defined in ZFA by $\forall v_1 \, \delta(\mathrm{Def}(v_1), v_1)$. So the V-definable elements of M are precisely the elements $\mathrm{Def}(\check{p})$ for $p \in V$.

Proof of Theorem 2. Let $x \in M$ be given. Working in V, temporarily fix a large ordinal α, and let P be the set of all p of rank $<\alpha$ such that

$$\|x \in \mathrm{Def}(\check{p})\| = 1.$$

Let d be the unique element of M such that

$$\|d \text{ is the intersection of the sets } \mathrm{Def}(u) \text{ for } u \in \check{P}\| = 1.$$

Clearly, d is V-definable (with $\check{P}$ as parameter). Also, since

$$\|u \in \check{P}\| = \bigvee_{p \in P} \|u = \check{p}\|,$$

it follows from the definition of P that $\|x \in d\| = 1$. Suppose d' is another V-definable set with $\|x \in d'\| = 1$. If $d' = \mathrm{Def}(\check{p})$ with p of rank $<\alpha$, then $p \in P$, so $\|d \subseteq d'\| = 1$. To remove the rank restriction on p, let α no longer be fixed and observe that as α increases, d decreases. Since a set in M cannot have a proper class (in the metatheory) of subsets, it follows that d is independent of α once α is large enough. For this stable value of d, the argument above, showing that $\|d \subseteq d'\| = 1$, applies regardless of the rank of p, since we can increase α to exceed this rank. □

The set d given by Theorem 2 will be called the *V-definable hull* of x and written $\mathrm{Hull}(x)$.

Corollary. *There are only a set of* V*-degrees below any given* V*-degree* $[x]$. *The ordering of these* V*-degrees is canonically isomorphic to the ordering of the* V*-definable partitions of* $\mathrm{Hull}(x)$ *by the relation 'coarser than'.*

Proof. Consider any $y \leqslant x$, and fix a V-definable f with $\|f(x) = y\| = 1$. Since f and $\mathrm{Hull}(y)$ are V-definable, so is $f^{-1}\,\mathrm{Hull}(y)$, which contains x (with truth value 1) and therefore includes $\mathrm{Hull}(x)$ (with truth value 1). Replacing f with its restriction to $\mathrm{Hull}(x)$, which is still V-definable and maps x to y, we can assume that $f:\mathrm{Hull}(x)\to\mathrm{Hull}(y)$ (with truth value 1). The range of f is V-definable and contains y, hence includes all of Hull(y). So $\|f$ maps Hull(x) onto Hull(y)$\| = 1$. Furthermore, f is uniquely determined by the information that it is V-definable, it sends x to y, and its domain is Hull(x). Indeed, if f' were another such function, then the set $\{z \in \mathrm{Hull}(x) \mid f(z)=f'(z)\}$ would be a V-definable set that contains x and would therefore equal all of Hull(x).

Associate to y the V-definable partition of Hull(x) consisting of the fibers of f (i.e., the pre-images of singletons). If $z \leq y$, say with $z = h(y)$ and h V-definable, then the partition associated to z consists of the fibers of $h \circ f$ (as $h \circ f$ sends x to z and is V-definable), hence is coarser than the partition associated to y.

Conversely, suppose $z \leq x$ and the partition associated to z is coarser than the partition associated to y. Let $g:\mathrm{Hull}(x)\to\mathrm{Hull}(z)$ be definable and send x to z. The coarseness assumption says that each fiber of f is included in a (unique) fiber of g. But then we can define a V-definable mapping $h:\mathrm{Hull}(y)\to\mathrm{Hull}(z)$ as the set (in M) of ordered pairs $(f(u), g(u))$ with $u \in \mathrm{Hull}(x)$. This h sends $y=f(x)$ to $z=g(x)$. So $z \leq y$.

Finally, we observe that every V-definable partition P of Hull(x) is associated to some $y \leq x$. Indeed, we can take y to be the element of P that contains x; the V-definable map that sends x to y is just the canonical projection $\mathrm{Hull}(x)\to P$ and the partition it induces is P.

This shows that, below x, the relative V-definability ordering corresponds to the 'coarser than' ordering of partitions of Hull(x). In particular, if $y, z \leq x$, then $y \equiv z$ if and only if the associated partitions coincide. This proves the second sentence of the corollary, and the first follows because a set of M has only a set (in the metatheory) of elements. □

6. Generating sets

Throughout this section, M will be a model of set theory and $\mathscr{E}$ will be its smallest representing topos, the topos of V-definable sets and functions of M. We consider the question whether $\mathscr{E}$, which we already know is a complete topos, is a Grothendieck topos. That is, does $\mathscr{E}$ have a set of generators? We begin by reducing the question to one about the V-degrees of M. Then we show that the answer is affirmative when M is a permutation model or a symmetric model. We do not know whether the conditions imposed on M in our definition of 'model of set theory' are sufficient by themselves to imply that $\mathscr{E}$ is a Grothendieck topos.

Theorem 3. *Let M be a model of set theory and $\mathscr{E}$ its smallest representing topos. $\mathscr{E}$ is a Grothendieck topos if and only if the class of V-degrees of M is a set.*

Proof. Suppose first that $\mathscr{E}$ is a Grothendieck topos. Then $\mathscr{E}$ contains an object G (namely the coproduct of any set of generators) such that every object of $\mathscr{E}$ is a quotient of a subobject of $G \times \check{I}$ for some $I \in V$. In V, let κ be a cardinal that is at least as large as any antichain (= set of pairwise disjoint elements) in the Boolean algebra $\mathscr{B}$ of truth values of M. We shall show that every element of M is V-definable from some y satisfying $\|y \in G \times \check{\kappa}\| = 1$. Since these y's constitute a set, it will follow, by the Corollary to Theorem 2, that the V-degrees also constitute a set, as desired.

So let any $x \in M$ be given. By our choice of G, there exists an $I \in V$ and a V-definable surjection f from a subset of $G \times \check{I}$ to Hull(x) in M. In particular, $\|(\exists u \in G)(\exists v \in \check{I})\, f(u, v) = x\| = 1$, so the patching property of M (along with the axiom of choice in the metatheory) provides elements u and v of M such that

$$\|u \in G\| = \|v \in \check{I}\| = \|f(u, v) = x\| = 1.$$

By definition of $\check{I}$, it follows that $\bigvee_{i \in I} \|v = \check{i}\| = 1$. The truth values $\|v = \check{i}\|$ form an antichain in $\mathscr{B}$, so at most κ of them are nonzero. Thus, there exists (in V) a function $g : \kappa \to I$ such that

$$1 = \bigvee_{\alpha < \kappa} \|v = g(\alpha)\check{}\,\| = \bigvee_{\alpha < \kappa} \|v = \check{g}(\check{\alpha})\| = \|(\exists \beta < \check{\kappa})\, v = \check{g}(\beta)\|.$$

By patching, find $\beta \in M$ with $\|\beta < \check{\kappa}\| = \|v = \check{g}(\beta)\| = 1$. Then $\|f(u, \check{g}(\beta)) = x\| = 1$. Since f and (trivially) $\check{g}$ are V-definable, this shows that x is V-definable from (u, β), an element (with truth value 1) of $G \times \check{\kappa}$. This completes the proof of one direction of the theorem.

The converse is easier. If the V-degrees constitute a set, fix an ordinal α such that every V-degree has a representative x with $\|x$ has rank $< \check{\alpha}\| = 1$. Then, for each such representative x, Hull(x) is a subobject in $\mathscr{E}$ of $V_\alpha(A)$. Given any $y \in M$, we have $y \leq x$ for one of the representatives x, and therefore, as in the proof of the Corollary to Theorem 2, we have a V-definable surjection $f : \text{Hull}(x) \twoheadrightarrow \text{Hull}(y)$ sending x to y. If Y is a V-definable set, then the restriction of f to $f^{-1}(Y \cap \text{Hull}(y))$ is a V-definable function, having y in its range to the extent that $y \in Y$, and having domain a subobject of $V_\alpha(A)$. This means that $V_\alpha(A)$ serves as an object of generators for $\mathscr{E}$. □

In the following theorem, we consider permutation models of ZFA. These are the two-valued models obtained, as in [4, Chapter 4], by starting with the full universe $V(A)$ built over a set A of atoms and then cutting down to the submodel of sets (and atoms) hereditarily symmetric with respect to a specified group $\mathscr{G}$ of permutations of A and a specified normal filter $\mathscr{F}$ of subgroups of $\mathscr{G}$.

Theorem 4. *If M is a permutation model, then its smallest representing topos is a Grothendieck topos.*

Proof. We shall verify the equivalent condition from Theorem 3 by finding a set $S \subseteq M$ such that every element of M is V-definable from an element of S; this will suffice, by the Corollary to Theorem 2.

We begin by constructing a copy of M inside the submodel V of pure sets. For the time being, we work in $V(A)$. Here, the axiom of choice holds, so there is a pure set A' with a bijection $b: A' \to A$. We can assume, e.g. by taking A' to be an ordinal, that no member of A' is an ordered pair. Then we can build a copy of $V(A)$ within V by using the pair $(0, x)$ instead of x; more precisely, our copy $V(A)'$ of $V(A)$ is defined recursively to consist of the elements of A' and all pairs $(0, x)$ where $x \subseteq V(A)'$. (The membership relation in $V(A)'$ puts y into $(0, x)$ if and only if $y \in x$ in the ordinary sense; the set of atoms in $V(A)'$ is given by $(0, A')$.) b induces an isomorphism, still called b, from $V(A)'$ onto $V(A)$; it sends $(0, x)$ to $\{b(y) \mid y \in x\}$. The group $\mathcal{G}$ and the filter $\mathcal{F}$ correspond, via b, to a group

$$\mathcal{G}' = \{b^{-1}gb \mid g \in \mathcal{G}\}$$

of permutations of A' (hence automorphisms of $V(A)'$) and a normal filter $\mathcal{F}'$ of subgroups of $\mathcal{G}'$. The hereditarily symmetric elements of $V(A)'$ form an isomorphic (via b) copy M' of M.

Bijections from A' to A other than b also induce isomorphisms from $V(A)'$ to $V(A)$. We shall have to deal in particular with bijections of the form gb with $g \in \mathcal{G}$. Since $\mathcal{F}$ is normal in $\mathcal{G}$, these bijections also send M' onto M.

Fix an ordinal α large enough so that every group $H \in \mathcal{F}$ that occurs as the symmetry group (= stabilizer) of some $x \in M$ actually occurs for some $x \in M$ of rank $<\alpha$. This is possible because $\mathcal{F}$ is a set. Define C to be the set of ordered quadruples of the form $(x, gb(x), y, gb(y))$ where x and y are elements of M' of rank (in the sense of $V(A)'$) smaller than α and where $g \in \mathcal{G}$. Notice that, since x and y are pure sets and $gb(x)$ and $gb(y)$ are in M, C is a subset of M. We claim that C is in fact an element of M. To prove this claim, it suffices (since $C \subseteq M$) to verify that C is symmetric. In fact, C is stabilized by the whole group $\mathcal{G}$ because, for any $h \in \mathcal{G}$,

$$h(x, gb(x), y, gb(y)) = (x, (hg)b(x), y, (hg)b(y)) \in C,$$

since x and y, being pure sets, are fixed by h.

We shall show that every element of M is V-definable from C and some ordered pair of the form $(x, b(x))$ with x of rank $<\alpha$ in M'. This will suffice to conclude the proof of Theorem 4.

For $x \in M'$, x' of rank $<\alpha$ in M', and $u \in M$, we define $Q(C, x, x', u)$ as follows. If $x \in A'$ and if there is a unique $a \in A$ with $(x, a, x', u) \in C$, then this a is $Q(C, x, x', u)$. If x is a set in the sense of M', i.e., if x has the form $(0, t)$, then $Q(C, x, x', u)$ is the set of all $Q(C, y, y', v)$ where $y \in t$ and $(x', u, y', v) \in C$ and $\mathrm{Sym}(y') \leq \mathrm{Sym}(y)$; here Sym means the stabilizer in $\mathcal{G}'$.

Lemma. *If $x \in M'$ and x' has rank $<\alpha$ in M' and* $\mathrm{Sym}(x') \leq \mathrm{Sym}(x)$ *and $g \in \mathcal{G}$, then $Q(C, x, x', gb(x')) = gb(x)$.*

Before proving the lemma, we observe that it will complete the proof of the theorem. Indeed, given any $z \in M$, we have $z = b(x)$ for some $x \in M'$, we can find x' of rank $<\alpha$ in M' with the same stabilizer as x (by our choice of α), and we then have, taking $g = 1$ in the lemma,

$$z = Q(C, x, x', b(x')).$$

So, since $x \in V$, z is V-definable from C,x' and $b(x')$, as desired.

Proof of the Lemma. We proceed by induction on x. Suppose first that $x \in A'$. According to the definition of Q in this case, we must check that $gb(x)$ is the unique atom a such that $(x, a, x', gb(x')) \in C$. The atoms a satisfying this requirement are, by definition of C, just those of the form $hb(x)$ with $h \in \mathcal{G}$ and $hb(x') = gb(x')$. Certainly $gb(x)$ is of this form; just take $h = g$. We must show that no other $a \in A$ is of this form, i.e., that $hb(x') = gb(x')$ implies $hb(x) = gb(x)$. But this follows from the assumption in the lemma that $\mathrm{Sym}(x') \subseteq \mathrm{Sym}(x)$; indeed, if $hb(x') = gb(x')$, then $b^{-1}h^{-1}gb \in \mathrm{Sym}(x') \subseteq \mathrm{Sym}(x)$, so $hb(x) = gb(x)$. This establishes the lemma when x is an atom.

Suppose now that x is a set in the sense of M', say $x = (0, t)$, and suppose that the lemma holds when x is replaced by any element of t. An element of $Q(C, x, x', gb(x'))$ is, by definition of Q, of the form $Q(C, y, y', v)$ with $y \in t$ and $(x', gb(x'), y', v) \in C$ and $\mathrm{Sym}(y') \subseteq \mathrm{Sym}(y)$. By definition of C, v must be $hb(y')$ for some $h \in \mathcal{G}$ such that $hb(x') = gb(x')$. So the elements of $Q(C, x, x', gb(x))$ are the objects $Q(C, y, y', hb(y')) = hb(y)$ (by induction hypothesis), where $y \in t$ and $hb(x') = gb(x')$. Among these, we have (with $h = g$) all the elements $gb(y)$ for $y \in t$; these are exactly the elements of $gb(x)$. To complete the proof, we must show that $Q(C, x, x', gb(x'))$ has no other elements. Consider any such element $hb(y)$, where $y \in t$ and $hb(x') = gb(x')$. Since $y \in t$, we have $hb(y) \in hb(x)$. But also, as in the previous paragraph, we deduce from $hb(x') = gb(x')$ and $\mathrm{Sym}(x') \subseteq \mathrm{Sym}(x)$ that $hb(x) = gb(x)$. So $hb(y) \in gb(x)$, as required. This completes the proof of the lemma and thus also the proof of Theorem 4. □

We turn next to the symmetric Boolean-valued models, as defined in [4, Chapter 5]. Such a model consists of the elements of a Boolean extension $V^{\mathcal{B}}$ that are hereditarily symmetric with respect to a group $\mathcal{G}$ of automorphisms of $\mathcal{B}$ and a normal filter $\mathcal{F}$ of subgroups of $\mathcal{G}$. Symmetric models satisfy ZF, without atoms.

Theorem 5. *If M is a symmetric model, then its smallest representing topos is a Grothendieck topos.*

The proof will be similar to the proof of Theorem 4. We shall have a copy of M within V, and generic filters will provide (as bijections from A' to A did in the previous proof) correspondences between this copy and the actual M. We shall define each element of M from an element of V (its code in this copy of M) and limited information about the correspondence (analogous to C and $(x', b(x'))$ in the previous proof).

A major notational difference between the proofs of Theorems 4 and 5 arises from the following circumstance. In the case of permutation models, V was really a submodel of $V(A)$, and we built a copy $V(A)'$ of $V(A)$ inside V. In the case of symmetric models, V is not actually included in $V^{\mathcal{B}}$ but is canonically embedded by the map sending each x to $\check{x}$. On the other hand, there is no need to construct a copy of $V^{\mathcal{B}}$ within V, for the definition of $V^{\mathcal{B}}$ makes it a subclass of V, a certain class of functions. Abstractly, the two situations are the same; both $V(A)$ and $V^{\mathcal{B}}$ include a copy of V and are themselves embedded in V. But the standard definitions use the former inclusion for $V(A)$ and the latter embedding for $V^{\mathcal{B}}$ as identifications.

To fix our notations, we remark that any x in $V^{\mathcal{B}}$ is also in V, so there exists $\check{x} \in V^{\mathcal{B}}$. The relation between x and $\check{x}$ is given by $\|x = \mathrm{Val}(G, \check{x})\| = 1$, where G is, in $V^{\mathcal{B}}$, the canonical $\check{V}$-generic ultrafilter in $\check{B}$ (containing $\check{b}$ with truth value b, for each $b \in \mathcal{B}$), and where Val is defined recursively by

$$\mathrm{Val}(G, z) = \{\mathrm{Val}(G, y) \mid y \in \mathrm{Domain}(z) \text{ and } z(y) \in G\}.$$

It is well known that we also have

$$\mathrm{Val}(G, z) = \{\mathrm{Val}(G, y) \mid \|y \in z\| \in G\}.$$

We shall often write simply $\mathrm{Val}(z)$ instead of $\mathrm{Val}(G, z)$. To avoid possible confusion, we remark that, although x can be obtained in $V^{\mathcal{B}}$ from $\check{x}$ (and G, via Val), $\check{x}$ cannot be obtained from x by any operation internal to $V^{\mathcal{B}}$; indeed, if $\|x = y\| < 1$, then $\|\check{x} = \check{y}\| = 0$ since $x \neq y$.

Proof of Theorem 5. For the first part of the proof, we work in V, remembering that $V^{\mathcal{B}}$ and M are subclasses of V. By the *symmetry type* of an $x \in M$ we mean the function $\mathcal{G} \to \mathcal{B}$ sending each $g \in \mathcal{G}$ to $\|x = g(x)\|$. Since there are only a set of functions $\mathcal{G} \to \mathcal{B}$, we can fix a sub*set* M_0 of M such that every element of M has the same symmetry type as some element of M_0. Replacing each element $x \in M_0$ with an ordered pair (in the sense of M) (x, ξ), using distinct ordinals ξ for distinct x's, we can arrange without loss of generality that, whenever x and y are distinct elements of M_0, then $\|g(x) = h(y)\| = 0$ for all $g, h \in \mathcal{G}$.

Let P be the element of $V^{\mathcal{B}}$ defined as follows. Its domain consists of all pairs, in the sense of M, $(\check{x}, g(x))$ where $x \in M_0$ and $g \in \mathcal{G}$. Its value at each point in its domain is 1. Since the pairs $(\check{x}, g(x))$ are in M (as $\mathcal{G}$ preserves M) and since P is stabilized by all of $\mathcal{G}$ (as $h(\check{x}, g(x)) = (\check{x}, hg(x))$ for all $h \in \mathcal{G}$), we have $P \in M$.

Observing that $g(x) = \mathrm{Val}(\check{g}(\check{x}))$, we easily check that the following statement has truth value 1 in $V^{\mathcal{B}}$: P is the set of all pairs of the form $(x, \mathrm{Val}(g(x)))$, where $x \in \check{M}_0$ and $g \in \check{\mathcal{G}}$.

Let Z be the set of all functions $f: \mathcal{G}^2 \to \mathcal{B}$ that are equivariant in the sense that, for all $g, h_1, h_2 \in \mathcal{G}$,

$$f(gh_1, gh_2) = g(f(h_1, h_2)).$$

We say that a function $f \in Z$ and a pair of elements $x, y \in M_0$ *cohere* if, for all $g_1, g_2, h_1, h_2 \in \mathcal{G}$,

$$\|g_1(x) = g_2(x)\| \wedge \|h_1(y) = h_2(y)\| \wedge f(g_1, h_1) \leq f(g_2, h_2).$$

Let E be the element of $V^{\mathcal{B}}$ defined as follows. Its domain consists of triples, in the sense of M, $(g(x), h(y), \check{f})$ where $f \in Z$ coheres with $x, y \in M_0$, and where $g, h \in \mathcal{G}$. Its values are given by

$$E(g(x), h(y), \check{f}) = f(g, h).$$

Coherence and our preliminary normalization of M_0 imply that E is well-defined. Indeed, if $(g(x), h(y), \check{f}) = (g'(x'), h'(y'), \check{f}')$, then $f = f'$, normalization gives $x = x'$ and $y = y'$, and coherence gives $f(g, h) \leq f(g', h')$; the converse inequality follows symmetrically. In fact, coherence and normalization also imply that, for f, x, y, g and h as in the definition of E,

$$\|(g(x), h(y), \check{f}) \in E\| = f(g, h).$$

Indeed $\geq$ follows from the definition of E, and for $\leq$ it suffices to prove

$$\|(g(x), h(y), \check{f}) = (g'(x'), h'(y'), \check{f}')\| \wedge f'(g', h') \leq f(g, h),$$

whenever $f' \in Z$ is coherent with $x', y' \in M_0$ and $g', h' \in \mathcal{G}$. Unless $f = f'$, $x = x'$ and $y = y'$, the required inequality is trivial since the truth value on the left is 0 (by normalization). And in the remaining case, the required inequality is given by coherence.

Each triple in the domain of E is clearly in M, and E is stabilized by all of $\mathcal{G}$, since, for $k \in \mathcal{G}$,

$$\begin{aligned} E(k(g(x), h(y), \check{f})) &= E(kg(x), kh(y), \check{f}) = f(kg, kh) \\ &= k(f(g, h)) = k(E(g(x), h(y), \check{f})). \end{aligned}$$

So $E \in M$.

The following observations about equivariance and coherence of specific functions will be needed later. Consider any $x, y \in M$ and define $f: \mathcal{G}^2 \to \mathcal{B}$ by $f(g, h) = \|g(y) \in h(x)\|$. Then f is equivariant because $\|k(g(y)) \in k(h(x))\| = k(\|g(y) \in h(x)\|)$. If x_0 and y_0 are in M_0 and have the same symmetry types as x and y, respectively, then f is coherent with y_0, x_0 (in this order!) because

$$\begin{aligned} &\|g_1(y_0) = g_2(y_0)\| \wedge \|h_1(x_0) = h_2(x_0)\| \wedge f(g_1, h_1) \\ &\quad = g_1(\|y_0 = g_1^{-1}g_2(y_0)\|) \wedge h_2(\|x_0 = h_1^{-1}h_2(x_0)\|) \wedge \|g_1(y) \in h_1(x)\| \\ &\quad = g_1(\|y = g_1^{-1}g_2(y)\|) \wedge h_1(\|x = h_1^{-1}h_2(x)\|) \wedge \|g_1(y) \in h_1(x)\| \\ &\quad = \|g_1(y) = g_2(y)\| \wedge \|h_1(x) = h_2(x)\| \wedge \|g_1(y) \in h_1(x)\| \\ &\quad \leq \|g_2(y) \in h_2(x)\| = f(g_2, h_2). \end{aligned}$$

We now leave V and begin working inside M. We define a five-place function Q as follows when the first two arguments are P and E (introduced above), the third argument is in $\check{M}$ (the copy of M inside the ground model $\check{V}$), the fourth argument is in $\check{M}_0$, and the fifth argument is arbitrary (in M, of course). It won't matter what Q does when its arguments are not of this form.

$Q(P, E, x, x_0, u)$ is the set of all $Q(P, E, y, y_0, v)$ such that

(i) $y \in \check{M}$, $y_0 \in \check{M}_0$, y and y_0 have the same symmetry type with respect to $\check{\mathscr{G}}$ and $\check{\mathscr{B}}$, and y has lower rank than x;

(ii) $(y_0, v) \in P$; and

(iii) if $f : \check{\mathscr{G}}^2 \to \check{\mathscr{B}}$ is the function given by

$$f(g, h) = \|g(y) \in h(x)\|_{\check{\mathscr{B}}},$$

then $(v, u, f) \in E$.

The computations we performed in V just before moving to M can be transferred to $\check{V}$, where they show that, if $x \in \check{M}$ and $x_0 \in \check{M}_0$ have the same symmetry type, then, in view of (i), the f defined in (iii) is equivariant and coheres with y_0, x_0.

Lemma. *The following statement has truth value* 1 *in* $V^{\mathscr{B}}$. *If* $x \in \check{M}$ *and* $x_0 \in \check{M}_0$ *have the same symmetry type, and if* $h \in \check{\mathscr{G}}$, *then*

$$Q(P, E, x, x_0, \mathrm{Val}(h(x_0))) = \mathrm{Val}(h(x)).$$

We recall that $\mathrm{Val}(z)$ means $\mathrm{Val}(G, z)$. The lemma is formulated for $V^{\mathscr{B}}$ rather than M because of the reference to G, which will normally not be in M.

Before proving the lemma, we show how it implies the theorem. Working in V, let any $x \in M$ be given. By choice of M_0, find an $x_0 \in M_0$ of the same symmetry type. Then, by the lemma with $h = 1$, we have with truth value 1,

$$Q(P, E, \check{x}, \check{x}_0, x_0) = Q(P, E, \check{x}, \check{x}_0, \mathrm{Val}(\check{x}_0)) = \mathrm{Val}(\check{x}) = x.$$

So x is V-definable in M from (P, E, x_0). Since P and E are fixed and x_0 ranges over only a set M_0, this shows that the ordering of V-degrees has a cofinal set. By the Corollary to Theorem 2, it follows that the whole ordering is a set, and by Theorem 3 it follows that the smallest representing topos is a Grothendieck topos.

Proof of the Lemma. We work in $V^{\mathscr{B}}$ and proceed by induction on the rank of x. Let x, x_0 and h be as in the lemma. The elements of $Q(P, E, x, x_0, \mathrm{Val}(h(x_0)))$ are, by definition of Q, all $Q(P, E, y, y_0, v)$ such that (i)–(iii) hold with u replaced by $\mathrm{Val}(h(x_0))$.

Clause (ii) says that $(y_0, v) \in P$. By our internal description of P in $V^{\mathscr{B}}$, this means that $y_0 \in \check{M}$ (which we also know by clause (i)) and $v = \mathrm{Val}(g(y_0))$ for some $g \in \check{\mathscr{G}}$. Since y has lower rank than x, the induction hypothesis implies that our typical element $Q(P, E, y, y_0, v)$ of $Q(P, E, x, x_0, \mathrm{Val}(h(x_0)))$ is

$$Q(P, E, y, y_0, v) = Q(P, E, y, y_0, \mathrm{Val}(g(y_0)) = \mathrm{Val}(g(y)),$$

with y, y_0 and g as above.

Furthermore, the assertion in clause (iii) that $(v, u, f) \in E$ can now be equivalently formulated as

$$(\mathrm{Val}(g(y_0)), \mathrm{Val}(h(x_0)), \mathrm{Val}(\check{f})) \in E = \mathrm{Val}(\check{E}).$$

hence as

$$\|(g(y_0), h(x_0), \check{f}) \in \check{E}\|_{\check{\mathscr{B}}} \in G.$$

But our earlier computation, in V, of the truth values in $\mathscr{B}$ of statements about membership in E can be transferred to $\check{V}$ to exhibit the truth values in $\check{\mathscr{B}}$ of statements about membership in $\check{E}$. As a result, our equivalent formulation of $(v, u, f) \in E$ simplifies to $f(g, h) \in G$. But this means, by definition of f, that

$$\|g(y) \in h(x)\|_{\check{\mathscr{B}}} \in G,$$

and therefore

$$\mathrm{Val}(g(y)) \in \mathrm{Val}(h(x)).$$

We have thus proved that an arbitrary element of $Q(P, E, x, x_0, \mathrm{Val}(h(x_0)))$ is also an element of $\mathrm{Val}(h(x))$.

Conversely, consider any element of $\mathrm{Val}(h(x))$. It has the form $\mathrm{Val}(y)$ for some $y \in \check{M}$ with $\|y \in h(x)\|_{\check{\mathscr{B}}} \in G$. In fact, we can take $y \in \mathrm{Domain}(h(x))$ with $h(x)(y) \in G$; in particular, we can take y to be of lower rank than x. Furthermore, by our choice of M_0, we can find $y_0 \in \check{M}_0$ with the same symmetry type as y. Let $v = \mathrm{Val}(y_0)$. Then clauses (i) and (ii) in the definition of $Q(P, E, x, x_0, \mathrm{Val}(h(x_0)))$ are clearly satisfied. As for clause (iii), we remarked right after the definition of Q that, when (as here) x and x_0 have the same symmetry type, then f is equivariant and coheres with y_0, x_0. To verify that $(\mathrm{Val}(y_0), \mathrm{Val}(h(x_0)), f) \in E$, as required by clause (iii), we proceed through the same chain of equivalences as in the preceding paragraph (now with $g = 1$) to reduce the problem to showing that $\|g(y) \in h(x)\|_{\check{\mathscr{B}}} \in G$. But this we already know, so the lemma and therefore also the theorem are proved. □

7. Reconstructing a model from its smallest representing topos

Let M be a model of set theory and $\mathscr{E}$ its smallest representing topos, the topos of V-definable sets and functions of M. Motivated by our earlier observation that the basic Fraenkel and Cohen models are, when considered as topoi, directed unions of slice topoi of Fourman's representing topoi, we would like to express M as a directed union of certain slice topoi of $\mathscr{E}$. Our first objective in this section is to show that M can always be expressed in this fashion.

We shall need to consider factorizations of the inclusion $\mathscr{E} \hookrightarrow \mathrm{M}$ through the canonical logical morphisms $H^* : \mathscr{E} \to \mathscr{E}/H$ (sending X to the projection $X \times H \to H$) for various objects H of $\mathscr{E}$. Such a factorization

$$\begin{array}{ccc} \mathscr{E} & \longrightarrow & M \\ {\scriptstyle H^*}\searrow & & \nearrow \\ & \mathscr{E}/H & \end{array}$$

is given by specifying an element z of H in M (viewed as a morphism $z : 1 \to H$ in M) to serve as the image of the canonical element of $H^*(H) = (\text{projection} : H \times H)$ given by the diagonal map Δ:

$$\begin{array}{ccc} H & \xrightarrow{\ \Delta\ } & H \times H \\ {\scriptstyle H^*(1)} \| & & \swarrow {\scriptstyle H^*(H)} \\ & H & \end{array}$$

Specifically, any $z \in M$ with $\|z \in H\| = 1$ induces a logical morphism $z^+ : \mathscr{E}/H \to M$ sending each object $p : X \to H$ of $\mathscr{E}/H$ to (the element of M that is with truth value 1) $p^{-1}\{z\}$ and sending each morphism

$$\begin{array}{ccc} X & \xrightarrow{\ f\ } & Y \\ {\scriptstyle p}\searrow & & \swarrow{\scriptstyle q} \\ & H & \end{array}$$

of $\mathscr{E}/H$ to the restriction $f \mid p^{-1}\{z\} : p^{-1}\{z\} \to q^{-1}\{z\}$.

It is easy to check that $z^+ \circ H^*$ is canonically isomorphic to the inclusion of $\mathscr{E}$ in M, so z^+ provides a factorization as above.

Unlike the inclusion $\mathscr{E} \hookrightarrow M$, z^+ need not be faithful. In order to express M as a directed union of slices $\mathscr{E}/H$, we shall need to choose the H's and z's so as to make z^+ faithful. Quite generally, a logical morphism between Boolean topoi is faithful if and only if its restriction to subobjects of 1, which is always a homomorphism of Boolean algebras, is one-to-one. (*Proof.* One direction is trivial, since a faithful functor is, in particular, one-to-one on the set of morphisms $1 \to 2$. For the other direction, notice that, if two distinct morphisms $X \to Y$ are sent to the same morphism, then their equalizer $E \hookrightarrow M$ is not an isomorphism but is sent to an isomorphism, so the complement $C = X - E$ is nonzero but sent to zero, so the truth value 'support of C', the image of $C \to 1$, is nonzero but sent to zero.) In the case of z^+, the subobjects of 1 in $\mathscr{E}/H$ are the subobjects S of H in $\mathscr{E}$, i.e., the V-definable subsets of H, and the subobjects of 1 in M are the members of $\mathscr{B}$. z^+ sends S to $\|z \in S\|$, so it is faithful if and only if $S = H$ is the only subobject of H with $\|z \in S\| = 1$, i.e., if and only if $H = \mathrm{Hull}(z)$.

Because of this result, we henceforth confine our attention to embeddings in M of slice topoi of $\mathscr{E}$ of the form

$$z^+ : \mathscr{E}/\mathrm{Hull}(z) \hookrightarrow M.$$

What elements of M are in the logical subtopos of $\mathscr{E}/\mathrm{Hull}(z)$ (via z^+), in any of the senses discussed earlier when we were identifying the smallest representing topos? Specifically, which elements of $V_\alpha(A)$ in M are in the image of z^+? Inspection of the definition of z^+ shows that the elements in question are exactly those of the form $f(z)$ where f is a V-definable function $H \to V_\alpha(A)$.

Thus, the image of z^+ consists (up to equivalence of categories) of the elements of M that are V-definable from z.

Clearly, M is the union of all the subtopoi obtained in this fashion as z is allowed to vary. And this is a directed union, since the subtopoi obtained from z_1 and z_2 are both included in the subtopos obtained from (z_1, z_2). It is also clear that, instead of using all elements $z \in M$, it would suffice to use a class of z's cofinal in the ordering of the V-degrees. The following theorem summarizes this discussion plus some easy observations whose verification we leave to the reader.

Theorem 6. *Let M be a model of set theory, and let $\mathscr{E}$ be its smallest representing topos. Let I be any subclass of M that is cofinal in the ordering of V-degrees. Then M is the directed union of the representing subtopoi*

$$\mathscr{E}(z) = \textit{topos of sets and functions } V\textit{-definable from } z,$$

for $z \in I$. Each $\mathscr{E}(z)$ is equivalent to $\mathscr{E}/\mathrm{Hull}(z)$ via z^+. If $z_1 \leq z_2$, then the inclusion $\mathscr{E}(z_1) \hookrightarrow \mathscr{E}(z_2)$ corresponds to the logical morphism $\mathscr{E}/\mathrm{Hull}(z_1) \to \mathscr{E}/\mathrm{Hull}(z_2)$ given by pulling back along the unique V-definable function $\mathrm{Hull}(z_2) \to \mathrm{Hull}(z_1)$ that sends z_2 to z_1.

According to Theorem 3, the index class I for the directed union in Theorem 6 can be taken to be a set precisely when $\mathscr{E}$ is a Grothendieck topos.

Theorem 6 is not entirely satisfactory because it does not enable us to recover M from just the topos $\mathscr{E}$. The index set I and the family of objects $(\mathrm{Hull}(z))_{z \in I}$ used to form the slice topoi are not determined by $\mathscr{E}$. We devote the rest of this section to showing that this unsatisfactory situation is unavoidable. It is possible for different models of set theory to have equivalent topoi of V-definable sets. We begin with a general criterion for such equivalence.

Theorem 7. *Let M and M' be models of set theory, and let $\mathscr{E}$ and $\mathscr{E}'$ be their respective smallest representing topoi. Then $\mathscr{E}$ and $\mathscr{E}'$ are equivalent with A corresponding to A' if and only if, for every formula $\varphi(x)$ in the language of* ZFA *and for every $p \in V$, if $\|\varphi(\check{p})\| = 1$ in M, then also $\|\varphi(\check{p})\| = 1$ in M' and vice versa.*

Proof. (Only if) Given an equivalence from $\mathscr{E}$ to $\mathscr{E}'$ sending the set A of atoms of M to the set A' of atoms of M', we can choose the objects $V_\alpha(A)$, $V_\alpha(A')$ of the Fourman–Hayashi hierarchy (which are defined up to isomorphism) so that they correspond to each other under the equivalence. It then follows, by induction on the rank of $p \in V$, that the equivalence sends the point $\check{p}: 1 \to V_\alpha(A)$ in $\mathscr{E}$ (for $\alpha > \mathrm{rank}(p)$) to the point $\check{p}: 1 \to V_\alpha(A)$ in $\mathscr{E}'$. Thus, the equivalence preserves all the ingredients in the definition of the Fourman–Hayashi interpretation of formulas of the form $\varphi(\check{p})$. Since such a formula is true in M (respectively, M') if and only if it is true in the Fourman–Hayashi interpretation in $\mathscr{E}$ (respectively $\mathscr{E}'$), the conclusion of the theorem follows.

(If) The hypothesis, that truth (with value 1) of formulas $\varphi(\check{p})$ is preserved between M and M', remains true if there are several parameters $\check{p}_1, \ldots, \check{p}_n$ instead of only one; just take $p = (p_1, \ldots, p_n)$ and modify φ accordingly. To show that $\mathscr{E}$ and $\mathscr{E}'$ are equivalent, we describe them in terms of the truth of such formulas.

An object of $\mathscr{E}$, i.e., a V-definable set of M, is determined by a formula $\gamma(x, \check{p})$ with one free variable, such that the sentence

$$(\exists!\, x)\, \gamma(x, \check{p}) \wedge \neg(\exists x \in A)\, \gamma(x, \check{p})$$

is true in M; another such formula $\gamma'(x, \check{p}')$ determines the same object if and only if the sentence

$$\exists x\, (\gamma(x, \check{p}) \wedge \gamma'(x, \check{p}'))$$

is true in M. The morphisms and the category structure of $\mathscr{E}$ can be defined similarly in terms of truth in M of sentences of the form $\varphi(\check{p}_1, \ldots, \check{p}_n)$. Since, for such sentences, truth is the same in M as in M', it follows that $\mathscr{E}$ is equivalent to $\mathscr{E}'$. $\square$

We shall employ the theorem just proved to show that two specific models have equivalent representing topoi. The first of the two models is the basic Fraenkel model M, the permutation model built from a countably infinite set A of atoms using the group $\mathscr{G}$ of all permutations of A and the normal filter $\mathscr{F}$ generated by the subgroups

$$\mathrm{Fix}(F) = \{g \in \mathscr{G} \mid g \text{ fixes each } a \in F\},$$

for finite $F \subseteq A$. The other model M' is built similarly (i.e., using the group $\mathscr{G}'$ of all permutations and the filter $\mathscr{F}'$ generated by subgroups fixing finite F) but starting with an uncountable set A' of atoms. The two models are clearly different as A and A' have different cardinalities (in the metatheory), but we shall show that they are equivalent in the sense described by Theorem 7. In fact, we shall define an elementary embedding $j: M \to M'$ such that $j(p) = p$ for all pure sets p. (Note that, in these models, $\check{p}$ is simply p.)

To simplify notation, we assume that $A \subseteq A'$. (Notice that we do not have $M \subseteq M'$. For example, $A \in M$ but $A \notin M'$.) As is customary, we say that a finite set F of atoms *supports* (or is *a support of*) an element x of M or M' if $\mathrm{Fix}(F)$ stabilizes x; by definition of permutation models, every x has a finite support. (In fact, every x in M or M' has a smallest support, but we shall avoid using this fact.) We define the elementary embedding $j: M \to M'$ by recursion on rank as follows. If $a \in A$, then $j(a) = a$. If x is a set of M with support F, then

$$j(x) = \{g(j(y)) \mid y \in x \text{ and } g \in \mathrm{Fix}'(F)\},$$

where the prime in $\mathrm{Fix}'(F)$ indicates that we use all permutations of A' that pointwise fix F. Since a set x can be supported by different F's, we need to check that $j(x)$ is well-defined, independently of the choice of F. Until we do so, we shall write $j(x)_F$ when necessary to avoid ambiguity.

It is clear from the definition that, if F supports x, then F also supports $j(x)_F$. Thus, j maps M into M'.

Lemma. (a) *$j(x)$ is well-defined.*

(b) *If $h \in \mathscr{G}'$ and h maps A onto A, then $h(j(x)) = j(\bar{h}(x))$, where $\bar{h} \in \mathscr{G}$ is the restriction of h to A.*

Proof. We prove both parts of the lemma by simultaneous induction on the rank of x. Induction hypothesis (a) ensures that $j(y)$ is well-defined for all $y \in x$.

(a) If F and F' are supports of x, then so is $F \cup F'$, so it suffices to prove $j(x)_F = j(x)_{F'}$ when $F \subseteq F'$. In fact, since F' is finite, it suffices to prove that $j(x)_F = j(x)_{F \cup \{a\}}$. The $\supseteq$ half of this equation is obvious, as $\mathrm{Fix}'(F \cup \{a\}) \subseteq \mathrm{Fix}'(F)$. To prove the $\subseteq$ half, consider an arbitrary element of $j(x)_F$, say $g(j(y))$ where $y \in x$ and g fixes F pointwise. Let $h \in \mathscr{G}'$ be a permutation of A' mapping A onto A, agreeing with g on F (i.e., fixing F pointwise), and agreeing with g on some support E of y except where this conflicts with the requirement that A be mapped to itself. Let the exceptional points, the atoms $e \in E$ with $g(e) \in A' - A$, be mapped by h to points in $A - (F \cup \{a\})$. It is clear that such an h exists. Let $k \in \mathscr{G}'$ be the identity except that it interchanges $g(e)$ with $h(e)$ for each of the exceptional atoms $e \in E$ where h disagreed with g. (Note that all these $h(e)$'s are in A while the $g(e)$'s are outside A, so k is well-defined.) Then g agrees with kh on E, which is a support of y and therefore of $j(y)$. So the element of $j(x)_F$ under consideration, $g(j(y))$, is equal to $kh(j(y))$ and therefore, by induction hypothesis (b), to $kj(\bar{h}(y))$. Here $\bar{h}$, the restriction of h to A, fixes F pointwise, hence fixes x, so from $y \in x$ we infer $\bar{h}(y) \in \bar{h}(x) = x$. Also, it is clear from the definition that k fixes $F \cup \{a\}$ pointwise. So $g(j(y)) = k(j(\bar{h}(y))) \in j(x)_{F \cup \{a\}}$, as desired.

(b) Let F support x. Then $h(F)$ supports $\bar{h}(x)$. Therefore,

$$j(\bar{h}(x)) = \{g'(j(y')) \mid y' \in \bar{h}(x),\ g' \in \mathrm{Fix}'(h(F))\}.$$

Every $y' \in \bar{h}(x)$ is of the form $\bar{h}(y)$ with $y \in x$ (namely, let $y = \bar{h}^{-1}(y')$). Every $g' \in \mathrm{Fix}'(h(F))$ is hgh^{-1} for some $g \in \mathrm{Fix}'(F)$ (namely, let $g = h^{-1}g'h$). So

$$j(\bar{h}(x)) = \{hgh^{-1}(j(\bar{h}(y))) \mid y \in x,\ g \in \mathrm{Fix}'(F)\}.$$

By induction hypothesis,

$$hgh^{-1}(j(\bar{h}(y))) = hgh^{-1}h(j(y)) = hg(j(y)).$$

Therefore,

$$j(\bar{h}(x)) = \{hg(j(y)) \mid y \in x,\ g \in \mathrm{Fix}'(F)\} = \{hz \mid z \in j(x)\}$$
$$= h(j(x)). \quad \square$$

The next lemma says that $j : M \to M'$ is an embedding of ZFA-models.

Lemma. *(a) $j(x) = j(y)$ if and only if $x = y$.*
(b) *$j(x) \in j(y)$ if and only if $x \in y$.*
(c) *$j(A) = A'$.*

Proof. The 'if' part of (a) is clear as j is well-defined, and the 'if' part of (b) is immediate from the definition of j (by taking $g = 1$). We prove the 'only if' parts of (a) and (b) by simultaneous induction on rank.

(a) If $x \neq y$, there is $z \in x$ with $z \notin y$ (or vice versa). Then $j(z) \in j(x)$ and, by induction hypothesis (b), $j(z) \notin j(y)$ (or vice versa). So $j(x) \neq j(y)$.

(b) Suppose $j(x) \in j(y)$. This means, by definition of j, that $j(x) = gj(z)$, where $z \in y$ and $g \in \mathrm{Fix}'(F)$ and F supports y. As in the proof of part (a) of the previous lemma, we can find h and k in $\mathscr{G}'$ such that h maps A onto itself and fixes F pointwise, k pointwise fixes a support S of x, and g agrees with kh on a support E of z. (First choose the supports S and E. Then choose h to agree with g on F and on E except at those $e \in E$ that g maps out of A; h maps these exceptional e's into $A - S$. Then let k interchange $g(e)$ with $h(e)$ for the exceptional e's.) We have

$$\begin{aligned} j(x) &= g(j(z)) \\ &= kh(j(z)), \quad \text{as } g \text{ and } kh \text{ agree on } E \text{ which supports } j(z) \\ &= k(j(\bar{h}(z))), \quad \text{by the preceding lemma.} \end{aligned}$$

So

$$\begin{aligned} j(\bar{h}(z)) &= k^{-1}(j(x)) \\ &= j(x), \quad \text{as } k^{-1} \text{ pointwise fixes } S \text{ which supports } j(x). \end{aligned}$$

By induction hypothesis (a), we have $x = \bar{h}(z)$. But $\bar{h}$ pointwise fixes a support F of y, so it fixes y. Since $z \in y$, we have

$$x = \bar{h}(z) \in \bar{h}(y) = y.$$

This completes the proof of (a) and (b).

(c) Since A is supported by the empty set, the definition of j gives

$$\begin{aligned} j(A) &= \{gj(a) \mid a \in A,\ g \in \mathrm{Fix}'(\emptyset) = \mathscr{G}'\} \\ &= \{ga \mid a \in A,\ g \in \mathscr{G}'\} = A'. \quad \square \end{aligned}$$

Since pure sets are fixed by all permutations of atoms, the definition of j immediately implies, by induction on rank, that $j(p) = p$ for all pure sets p.

Finally, we show that the embedding $j: M \to M'$ is elementary, by using the well-known Tarski criterion. That is, we assume that M' satisfies $\varphi(j(x), y)$ for a certain $x \in M$ and $y \in M'$, and we find a $z \in M$ such that M' satisfies $\varphi(j(x), j(z))$. Let $F \subseteq A$ support x, and let $E \subseteq A'$ support y. Let g be a permutation that fixes F pointwise and maps E into A. Since F supports $j(x)$ and since g is an automorphism of M', we find that M' satisfies $\varphi(g(j(x)), g(y))$, which is $\varphi(j(x), g(y))$. Also, $g(y)$ is supported by $g(E) \subseteq A$.

To complete the proof, all we need is the following lemma, which ensures that $g(y) = j(z)$ for some $z \in M$.

Lemma. *If $x \in M'$ has a support included in A, then $x = j(z)$ for some $z \in M$.*

Proof. We proceed by induction on the rank of x. Let $F \subseteq A$ support x, and let

$$z = \{y \in M \mid j(y) \in x\}.$$

We first check that $z \in M$. z is a set, because j preserves ranks, and clearly $z \subseteq M$, so we need only check that z has a finite support. In fact, F supports z. Indeed, any permutation in $\mathrm{Fix}(F)$ is $\bar{h}^{-1}$ for some $h \in \mathrm{Fix}'(F)$ (namely, g^{-1} extended arbitrarily to a permutation of A'), and we have

$$\begin{aligned}\bar{h}^{-1}(z) &= \{\bar{h}^{-1}(y) \mid y \in M \text{ and } j(y) \in x\} \\ &= \{u \mid u \in M \text{ and } j(\bar{h}(u)) \in x\}.\end{aligned}$$

By an earlier lemma, $j(\bar{h}(u)) = h(j(u))$. Also, as h fixes a support F of x, $h(x) = x$. Thus,

$$\begin{aligned}\bar{h}^{-1}(z) &= \{u \in M \mid h(j(u)) \in h(x)\}. \\ &= \{u \in M \mid j(u) \in x\} = z.\end{aligned}$$

Thus, $z \in M$.

$j(z)$ consists of all $g(j(y))$ such that $j(y) \in x$ and $g \in \mathrm{Fix}'(F)$. As F supports x, all these elements of $j(z)$ satisfy $g(j(y)) \in g(x) = x$, so $j(z) \subseteq x$. For the converse inclusion, consider an arbitrary $u \in x$ and let $E \subseteq A'$ support u. Let h be a permutation of A' that fixes F pointwise and maps E into A. Then $h(u)$ is supported by $h(E) \subseteq A$, so by induction hypothesis $h(u) = j(y)$ for some $y \in M$. Since h fixes the support F of x pointwise, and since $u \in x$, we have

$$j(y) = h(u) \in h(x) = x.$$

Therefore, using again that h and therefore h^{-1} fix F pointwise, we obtain

$$u = h^{-1}(j(y)) \in j(z).$$

This completes the proof that $x = j(z)$. □

The lemma completes the proof that j is an elementary embedding and therefore the topoi of V-definable sets and functions of M and M' are equivalent.

8. Open problems

(1) Develop a similar theory without our smallness requirements on M. That is, allow a proper class A of atoms or a proper class $\mathscr{B}$ of truth values or both.

(2) Do our smallness requirements imply that the smallest representing topos is a Grothendieck topos?

(3) Although a model M is not uniquely determined by its smallest representing topos $\mathscr{E}$, is there a way to (canonically) recover one of the models that $\mathscr{E}$ represents?

References

[1] A. Blass and A. Scedrov, Freyd's models for the independence of the axiom of choice, Mem. Amer. Math. Soc. 404 (1989).
[2] M.P. Fourman, Sheaf models for set theory, J. Pure Appl. Algebra 19 (1980) 91–101.
[3] S. Hayashi, On set theories in toposes, in: G.H. Müller, G. Takeuti, and T. Tugué, eds., Logic Symposia Hakone 1979, 1980, Lecture Notes in Math. 891 (Springer, Berlin, 1981) 23–29.
[4] T. Jech, The Axiom of Choice, Stud. Logic Found. Math. 75 (North-Holland, Amsterdam, 1973).
[5] P.T. Johnstone, Topos Theory, London Math. Soc. Monographs 10 (Academic Press, London, 1977).
[6] J. Myhill and D. Scott, Ordinal definability, in: D. Scott, ed., Axiomatic Set Theory, Proc. Sympos. Pure Math. 13, Part I (Amer. Math. Soc., Providence, RI, 1971) 271–278.
[7] R. Paré, Colimits in topoi, Bull. Amer. Math. Soc. 80 (1974) 556–561.