Words, free algebras, and coequalizers

by

Andreas Blass (Ann Arbor, Mich.)

Abstract. For a variety of infinitary algebras subject to no identities, the classical construction of free algebras as algebras of words is shown to work in any topos. Although coequalizers in such varieties cannot be constructed in the usual way, as algebras of equivalence classes, unless the internal axiom of choice holds, they do exist and can be obtained as algebras of words in any Boolean topos. When the definition of the variety includes identities, however, neither free algebras nor coequalizers need exist in the absence of internal choice, even in models of Zermelo–Fraenkel set theory (provided a certain large cardinal axiom is consistent).

Consider a variety of algebras, defined by specifying a set of operation symbols, the number of argument places or "arity" of each symbol, and a set of identities, i.e. equations between expressions built up in the usual fashion from the operation symbols and variables. An algebra in such a variety is a set together with interpretations of the operation symbols as operations (with the right number of arguments) on this set, such that the identities are true for all values of the variables. A homomorphism of algebras is a function that preserves the operations. One of the basic theorems of universal algebra asserts that every variety contains a free algebra on any set X of generators, i.e. an algebra A together with a function $\eta\colon X \to A$ such that, for any algebra A' in the variety and any function $\eta'\colon X \to A'$, there is a unique homomorphism $\alpha\colon A \to A'$ with $\eta' = \alpha\eta$.

This theorem has two standard proofs. The classical proof [2] begins with the set of well-formed expressions, called words, built up from the operation symbols and (names for) the members of X. There is a natural interpretation of the operation symbols as operations on words. If there were no identities to be satisfied, this algebra of words would be the desired free algebra. In general, however, this algebra is not in the variety, so one has to identify words as required by the identities. Formally, this means that the free algebra in the variety is obtained as the quotient of the "absolutely free" algebra of words by an appropriate congruence relation. The modern proof in [24] (see also [31] for a version of this proof that does not mention categories) applies the adjoint functor theorem to produce a left adjoint for the underlying-set functor from the variety to the category of sets. The value of this adjoint at X is the desired free algebra, η being the unit of the adjunction. The least trivial hypothesis of the adjoint functor theorem to verify is the solution set condition, but this reduces to finding a bound on the sizes of algebras generated by X, which is a straightforward calculation in cardinal arithmetic.

Even if some of the operation symbols are infinitary, everything in the previous paragraph remains true provided we stretch the meaning of "expression" to permit infinite length (and perhaps stretch the meaning of "straightforward calculation in cardinal arithmetic" a bit). However, the classical proof now uses the axiom of choice in forming the quotient algebra A of the word algebra W, for to define the value of an infinitary operation at a set of arguments in A one chooses representatives in W for these arguments and applies the operation to the representatives. The modern proof also uses the axiom of choice (even in the finitary case) to justify the computations involving infinite cardinals.

Paré and Schumacher [30] have developed a version of the adjoint functor theorem for indexed categories and used it to carry out the modern proof of the existence of absolutely free algebras in arbitrary topoi with natural numbers objects. Thus, not only the axiom of choice but even classical logic are dispensable. Rosebrugh [33] has used the indexed adjoint functor theorem to obtain free algebras for varieties involving identities, under the additional assumption that the topos is Grothendieck over a topos satisfying the axiom of choice.

We shall show in this paper that the classical construction of absolutely free algebras as algebras of words can also be carried out in any topos with a natural numbers object. We shall also show that some additional choice-related assumption (as in Rosebrugh's theorem) is needed to obtain free algebras when identities are present. Specifically, we shall prove the consistency (relative to a certain large cardinal axiom) of Zermelo–Fraenkel set theory ZF (without choice) plus the existence of a variety of infinitary algebras with no initial member, i.e., no algebra freely generated by the empty set.

We shall also consider some related questions about the existence and nature of coequalizers (or, equivalently, quotients by congruences) in varieties of infinitary algebras. We shall show that, although these coequalizers cannot in general be constructed as algebras of equivalence classes unless the axiom of choice holds, they do exist in Boolean topoi with natural numbers objects provided no identities are involved in the definition of the variety. When identities are present, however, coequalizers need not exist at all, even in models of ZF.

This paper consists of two parts that are almost completely independent of each other. The first part, Sections 1 through 7, is about varieties defined by operations only, without any identities. The main result here asserts that the classical construction of free algebras of words makes sense in an arbitrary topos with natural numbers object. In fact, we show (in Sections 4 through 6) that one can essentially copy the classical argument as an argument entirely within the internal logic of the topos. Sections 2 and 3 contain the needed information on how to handle and interpret this internal logic. Neither the general discussion of internal logic in Sections 2 and 3, nor the discussion of free algebras of words in the next three sections contain any essential innovations (compared to [29] and [31] respectively). The main point of these sections is simply that, when appropriately set up, the internal logic provides a very powerful tool, essentially eliminating the need for any innovation in passing from a classical set-theoretic argument (even a moderately complex one) to a proof in topos theory. This part ends with a section in which we apply word constructions to the study of coequalizers in varieties without identities. One cannot always construct coequalizers simply by identifying appropriate elements in a given algebra; indeed, the feasibility of such a construction implies the (internal) axiom of choice. However, we can construct coequalizers as algebras of words, assuming only that the topos is Boolean.

The second part, Sections 8 through 10, is about varieties whose definitions involve identities. Section 8 contains positive results about the existence of free algebras and coequalizers either assuming the internal axiom of choice or, in the case of models of ZF rather than general topoi, assuming a rather weak consequence of the axiom of choice. The main results of this part, however, are negative ones to the effect that free algebras (Section 9) and coequalizers (Section 10) need not exist, even in models of ZF (provided a certain large cardinal assumption is consistent with set theory). Except for the first proposition in Section 8, the second part of the paper makes no use of topos theory, although one could, of course, view the models of ZF as topoi.

1. Definitions and preliminary remarks. In Sections 1 through 7, we work with a topos $\mathscr{E}$, assuming from Section 5 on that $\mathscr{E}$ has a natural numbers object N and in Section 7 that $\mathscr{E}$ is Boolean. We are concerned with varieties of algebras in $\mathscr{E}$ not subject to any identities, and our objective is to construct free algebras and coequalizers in such a variety as algebras of words.

Our first task is to explain what varieties and algebras are in $\mathscr{E}$. Of course, if $\mathscr{E}$ were the topos of sets, our explanation should reduce to the classical definition. A variety (without identities) would be specified by giving its similarity type, namely a set J of operation symbols and, for each $j \in J$, an arity I_j, i.e. a set that indexes the argument places of j. An algebra would consist of a set A and, for each $j \in J$, an operation j_A: $A^{I_j} \to A$. (In the theory of finitary varieties [13], it is customary to use only natural numbers as arities, but, if we want to permit infinite arities and do not want to assume the axiom of choice, there is no convenient substitute for natural numbers preferable to arbitrary sets; see [32].) The definition of a variety is easily translated into arbitrary topoi, using the standard device [17, 22] of representing an indexed family $\{I_j | \ j \in J\}$ by the map τ from the disjoint union of the I_j to J, sending all elements of I_j to j. Thus, a variety in $\mathscr{E}$ is specified by giving a similarity type which is simply a morphism τ: $I \to J$ of $\mathscr{E}$. We think of J as the object of operation symbols and the fibers of τ as arities.

Before translating the definition of algebras, it will be helpful to recall some basic facts about localization of topoi; see [17] for details. For any object X of a topos $\mathscr{E}$, the category $\mathscr{E}/X$ of $\mathscr{E}$-morphisms into X (with commutative triangles as morphisms) is also a topos. If $f\colon X\to Y$, then composition with f defines a functor $\Sigma_f\colon \mathscr{E}/X\to\mathscr{E}/Y$. This functor has a right adjoint $\Delta_f\colon \mathscr{E}/Y\to\mathscr{E}/X$ defined by pullback along f, and Δ_f in turn has a right adjoint $\Pi_f\colon \mathscr{E}/X\to\mathscr{E}/Y$. $\mathscr{E}/1$ is isomorphic to (and identified with) $\mathscr{E}$, and if f is the unique morphism from X to 1 we write Σ_X, Δ_X, Π_X for Σ_f, Δ_f, Π_f. If $\mathscr{E}$ is the category of sets, we think of $\mathscr{E}/X$ as the category of X-indexed families of sets. Applied to such a family $\{A_x|\ x\in X\}$, Σ_f and Π_f (where $f\colon X\to Y$) yield the Y-indexed families $\{S_y|\ y\in Y\}$ and $\{P_y|\ y\in Y\}$, where S_y is the disjoint union of the A_x for $x\in f^{-1}\{y\}$ and where P_y is the set of all functions that assign to each $x\in f^{-1}\{y\}$ an element of A_x. Also, $\Delta_f\{B_y|\ y\in Y\}=\{B_{f(x)}|\ x\in X\}$.

In terms of these constructions, we can express the definition of algebras for the veriety specified by $\tau\colon I\to J$ (henceforth called τ-algebras) in a simple form that generalizes to arbitrary topoi. To see this, note first that, if $\mathscr{E}$ is the category of sets, $\Delta_I A$ is the I-indexed family all of whose members are A, so $\Pi_\tau\Delta_I A$ is the J-indexed family $\{A^{\tau^{-1}\{j\}}|\ j\in J\}$, and $\Sigma_J\Pi_\tau\Delta_I A$ is the disjoint union of this family. But a τ-algebra structure on A consists of, for each $j\in J$, a map $A^{\tau^{-1}\{j\}}\to A$; all these maps j_A may be combined into a single map $\mu\colon \Sigma_J\Pi_\tau\Delta_I A\to A$. In this form, the concept generalizes to arbitrary topoi.

DEFINITION. Let $\tau\colon I\to J$ be a morphism in a topos $\mathscr{E}$. A *τ-algebra* in $\mathscr{E}$ is a pair (A,μ), where A is an object of $\mathscr{E}$ and μ is a morphism from $\Sigma_J\Pi_\tau\Delta_I A$ to A. A *homomorphism* from (A,μ) to (B,ν) is a morphism $\alpha\colon A\to B$ such that $\nu\cdot\Sigma_J\Pi_\tau\Delta_I(\alpha)=\alpha\cdot\mu$.

It is trivial to check that τ-algebras and homomorphisms form a category with a forgetful functor to $\mathscr{E}$, sending (A,μ) to A. We shall define a left adjoint for this functor by constructing, for each object X of $\mathscr{E}$, an algebra of τ-words over X and showing that it is the free τ-algebra on X, i.e., that every morphism from X to the underlying object of any τ-algebra extends uniquely to a homomorphism on the algebra of words.

As a preliminary step, we show that it suffices to consider the case where X is the initial (i.e. empty) object 0 of $\mathscr{E}$. For suppose we could construct the free τ-algebra on 0 (i.e. the initial τ-algebra) for every τ, and suppose we are given $\tau\colon I\to J$ and X. Let $J'=J+X$, and let $\tau'\colon I\to J'$ be the composite of τ and the inclusion $i\colon J\hookrightarrow J+X$. (Intuitively, τ' has, in addition to the operation symbols of τ, all the elements of X as 0-ary operation symbols, i.e., constant symbols.) We claim that the free τ-algebra on X can be obtained as the free τ'-algebra on 0.

To see this, first check that $\Pi_i\colon \mathscr{E}/J\to\mathscr{E}/J'$ sends any object $\zeta\colon Z\to J$ to $\zeta+\mathrm{id}_X\colon Z+X\to J+X=J'$; this is easy using the definition of Π_i as the right adjoint of Δ_i and the fact that every object of $\mathscr{E}/J'$ has the form $\alpha+\beta\colon A+B\to J+X$ for some $\alpha\colon A\to J$ and $\beta\colon B\to X$ (with $\alpha=\Delta_i(\alpha+\beta)$, etc.). It follows that $\Sigma_{J'}\Pi_i\zeta=(\Sigma_J\zeta)+X$. Thus,

$$\Sigma_{J'}\Pi_{\tau'}\Delta_I A=\Sigma_{J'}\Pi_i\Pi_\tau\Delta_I A=(\Sigma_J\Pi_\tau\Delta_I A)+X.$$

A τ'-algebra structure μ' on A is therefore given by specifying two morphisms, $\mu\colon \Sigma_J\Pi_\tau\Delta_I A\to A$ and $\xi\colon X\to A$. In other words, a τ'-algebra is the same as a τ-algebra with a specified morphism of X into its underlying object, and a homomorphism of τ'-algebras is a homomorphism of τ-algebras that commutes with the specified morphisms. It is now clear that a τ-algebra (A,μ) is free on X, with inclusion of generators $\eta\colon X\to A$, if and only if the τ'-algebra $(A,\langle\mu,\eta\rangle)$ is initial.

Sections 2 through 6 will contain a construction of the initial τ-algebra as an algebra of words, for any τ in any topos $\mathscr{E}$ with natural numbers object. Before embarking on that construction, we make a few remarks to motivate some aspects of it.

Words built up from the operation symbols in J are customarily viewed as well-ordered sequences of symbols, but it is clear that, in the absence of the axiom of choice, we cannot expect the arities to be well-ordered sets, so the symbols in a word need not be well-ordered. Indeed, they need not even be linearly ordered. Fortunately, the ordering of the symbols is entirely irrelevant and may be dispensed with. But what, then, is a word? There are at least two approaches to answering this question. One involves iterated pairing, as in the customary formalization of the syntax of infinitary logic [1]; it seems to be unsuitable for direct translation into topos theory. The approach adopted here is to analyze the essential structure of words in the familiar situation where $\mathscr{E}$ is the category of sets. Suppose, for example, that J contains two binary operation symbols, $+$ and $\cdot$, one unary symbol exp, and three constant (i.e. 0-ary) symbols, 1, i, and π, so that $1+\exp(i\cdot\pi)$ is a word. (For the sake of uniformity, one ordinarily defines words so that operation symbols precede their arguments, so this word would properly be $+(1,\exp(\cdot(i,\pi)))$ or simply $+1\exp\cdot i\pi$, but we have already seen that the order of symbols is not part of the essential structure we are seeking.) The essential structure of this word can be analyzed as follows. First, its principal (i.e. outermost) operator is the binary symbol $+$, and this operator is applied to the words 1 and $\exp(i\cdot\pi)$. There is a strong temptation to say "1 and $\exp(i\cdot\pi)$ in that order", since there is no commutativity here, but this is reasonable only because the arity of the symbol $+$, which might be written {left, right} (since a sum has a left and a right summand), is commonly thought of as ordered, with left preceding right. What is essential is not the order but the fact that 1 corresponds to left and $\exp(i\cdot\pi)$ to right. Thus, we should say that $+$ is applied, not to the words 1 and $\exp(i\cdot\pi)$ in that order, but rather to the function, from its arity to words, left $\mapsto 1$, right $\mapsto \exp(i\cdot\pi)$. Continuing the analysis, we say that $\exp(i\cdot\pi)$ has principal operator exp applied to (the function that sends the unique member, say $*$, of the arity of exp to) the word $i\cdot\pi$, which in turn has principal operator $\cdot$ applied to the function left $\mapsto i$, right $\mapsto \pi$. (It seems too pedantic to distinguish this left and right from those for $+$.) Finally, 1 has principal operator 1 applied to the unique function from its empty arity to words, and likewise for i and π. Thus, the structure of the word $1+\exp(i\cdot\pi)$ is described by the following list of all the symbols occurring in it together with the location of the subword having any given symbol as principal operator:

Symbol	Position
$+$	$\langle\ \rangle$
1	$\langle \text{left} \rangle$
exp	$\langle \text{right} \rangle$
$\cdot$	$\langle \text{right}, * \rangle$
π	$\langle \text{right}, *, \text{right} \rangle$
i	$\langle \text{right}, *, \text{left} \rangle$,

where, for example, the last line means that i is the principal connective of the word assigned to left by the function on which the symbol at position $\langle \text{right}, * \rangle$ operates. It is easy to see that any word (even with repetitions of symbols) can be described by such a table, in which

(a) the positions are certain finite sequences of elements of arities,

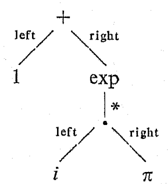
(b) there is a function assigning to each position a symbol,

(c) the empty sequence is a position,

(d) a nonempty sequence is a position if and only if, when it is decomposed as the concatenation $s^\frown\langle i \rangle$ where i is the last term of the sequence and s is the rest, s is a position, and i is in the arity of the symbol associated to s by the function in (b),

(e) there is no infinite sequence all of whose finite initial segments are positions.

A word can thus be represented by a set of positions and a function from this set into the set J of symbols; indeed, it suffices to specify the function since the set of positions is then determined as the domain of the function. A useful way to visualize a word is as a downward-branching labelled tree. Each node



in the tree corresponds to the position given by the labels of the edges leading from the top of the tree to that node, and the node is labelled with the operation symbol assigned to that position.

Condition (e) excludes trees with an infinite descending path. Such trees would correspond to "words" like $\exp(\exp(\exp(\ldots)))$; we could define an algebra that contained such words, but it would not be the initial algebra since the value of a homomorphism at such a word would not be well-defined. The condition that there be no infinite descending path is equivalent, in set theory with the axiom of choice, to well-foundedness, which means that every nonempty set of nodes has a minimal element. In set theory without the axiom of choice, the equivalence no longer holds, and well-foundedness turns out to be the correct condition to use; in non-Boolean topoi, even well-foundedness must be rephrased in a classically equivalent but intuitionistically inequivalent way (see Section 4). Our definition of words (in Section 5) will be a formulation, appropriate for general topoi, of the notion of a well-founded tree of finite sequences of elements of arities, together with a function assigning to each node of the tree an operation symbol, in such a way that conditions (c) and (d) above are fulfilled.

To find appropriate topos-theoretic formulations of set-theoretic concepts as in the preceding sentence, we make systematic use of the internal logic of topoi, introduced by Mitchell [26] and Bénabou [5]. Indeed, the main constructions and proofs (Sections 4, 5, and 6) will be done in this logic; only at the end of Section 6 will we return to the external world to show that we have an initial algebra as an actual object (and morphism) in $\mathscr{E}$ rather than as a term in the internal language.

Section 2 is a brief explanation of the internal logic of a topos $\mathscr{E}$, and Section 3 deals with the process of obtaining actual objects and morphisms from descriptions of them in the internal logic. The material in these two sections is not new and is included only for the convenience of the reader and to fix notational conventions. In Section 4 we introduce the concept of well-foundedness (as a term in the internal logic) and study some of its properties, the most important property being a version of Mikkelsen's theorem [25] on existence of inductively defined maps. Section 5 introduces trees of finite sequences (positions) and words. Finally, in Section 6, we complete the construction of an initial algebra by putting an appropriate algebra structure on the set of words and showing that the resulting algebra is initial.

2. Internal logic. The use of a many-sorted intuitionistic logic to efficiently express and prove internal properties of a topos was pioneered by Mitchell [26] and Bénabou [5] and developed in detail by Osius [29] and Fourman [10]. Unfortunately, different authors (including the present one) use different (though essentially equivalent) formulations of this internal logic, so we devote this section to presenting the particular formalization that we intend to use.

Throughout this section, let $\mathscr{E}$ be a fixed topos. The internal language of $\mathscr{E}$ has the following symbols. First, for each object A of $\mathscr{E}$, there are denumerably many variables of sort A; we assume that no variable is of two different sorts, and we write sort (v) for the (unique) sort of the variable v. Second, for each morphism $f\colon \prod_{i<n} A_i \to B$ with a specified representation of its domain as a product of a finite number n of factors, there is an n-ary operation symbol, also written f although, to be completely accurate, the notation for this operation symbol ought to include not only the morphism f but also the product representation of its domain. Finally, there is the symbol $\mapsto$, which will play the role of Church's λ-operator of function abstraction. Note that, in the definition of the operation symbols, $n = 0$ is allowed, so every morphism $1 \to B$ serves as a 0-ary operation symbol.

The syntax of the internal language is given by the definition of its well-formed expressions or *terms*. For all sets V of variables, we define the terms on V (also known as terms with all their free variables in V) and the sorts of these terms by the following induction, in which "$t\varepsilon_V A$" means that t is a term on V of sort A. (The symbol ε is intended to be different from, although reminiscent of, the membership symbol $\in$.)

(a) If v is a variable in V, then $v\varepsilon_V \operatorname{sort}(v)$.

(b) If $f\colon \prod_{i<n} A_i \to B$ and if $t_i \varepsilon_V A_i$ for each $i<n$, then $ft_0 \dots t_{n-1} \varepsilon_V B$. (We often write $f(t_0, \dots, t_{n-1})$ for greater legibility.)

(c) If x is any variable and $t\varepsilon_{V\cup\{x\}} A$, then $x \mapsto t\varepsilon_V A^{\operatorname{sort}(x)}$.

A term on the empty set of variables is said to be closed. We omit the subscript on ε when it is clear from the context or irrelevant.

It will be useful later to have the notion of substitution of one term for all free occurrences of a variable in another term and the notion of such a substitution being legitimate in the sense that no free variable becomes bound as a result of the substitution. These concepts are defined by the following induction on terms. Let y be a variable and $a\varepsilon \operatorname{sort}(y)$.

(a) $y[a/y] = a$, and the substitution is legitimate. $v[a/y] = v$ if v is a variable other than y, and the substitution is legitimate.

(b) $f(t_0, \dots, t_{n-1})[a/y] = f(t_0[a/y], \dots, t_{n-1}[a/y])$, and the substitution is legitimate if and only if all the substitutions $t_i[a/y]$ are legitimate.

(c) $(y \mapsto t)[a/y] = y \mapsto t$, and the substitution is legitimate. If x is a variable other than y, then $(x \mapsto t)[a/y]$ is $x \mapsto (t[a/y])$, and the substitution is legitimate if and only if the substitution $t[a/y]$ is legitimate and either x is not free in a (i.e. $a\varepsilon_W \operatorname{sort}(y)$ with $x \notin W$) or y is not free in t.

The usual elementary properties of substitution hold; we omit their easy proofs.

The semantics of the internal language of $\mathscr{E}$ assigns to each term $t\varepsilon_V A$ a denotation $|t|_V\colon \prod_{v\in V} \operatorname{sort}(v) \to A$, a morphism defined as follows, with $\operatorname{sort}(V)$ meaning $\prod_{v\in V} \operatorname{sort}(v)$.

(a) If $v \in V$ then $|v|_V\colon \operatorname{sort}(V) \to \operatorname{sort}(v)$ is the vth projection.

(b) If $f\colon \prod_{i<n} A_i \to B$ and $t_i \varepsilon_V A_i$, then $|ft_0 \dots t_{n-1}|_V$ is the composite

$$\operatorname{sort}(V) \xrightarrow[\langle |t_i|_V\rangle_{i<n}]{} \prod_{i<n} A_i \xrightarrow[f]{} B .$$

(c) If $t\varepsilon_{V\cup\{x\}} A$ then $|x \mapsto t|_V\colon \operatorname{sort}(V) \to A^{\operatorname{sort}(x)}$ is the exponential adjoint of the composite

$$\operatorname{sort}(V) \times \operatorname{sort}(x) \xrightarrow[\pi\times \mathrm{id}]{} \operatorname{sort}(V-\{x\}) \times \operatorname{sort}(x) \cong \operatorname{sort}(V \cup \{x\}) \xrightarrow[|t|_{V\cup\{x\}}]{} A ,$$

where $\pi\colon \operatorname{sort}(V) \to \operatorname{sort}(V-\{x\})$ is the projection if $x \in V$ and the identity otherwise. We write "$t_1 \equiv_V t_2$" to mean that $|t_1|_V = |t_2|_V$, it being understood that t_1 and t_2 are terms on V of the same sort. We write $t_1 \equiv t_2$ without the subscript when the intended V is clear from the context or irrelevant.

It should be emphasized that the internal language contains no meaningful expressions other than the terms. In particular, ε, $|\ |$, and $\equiv$ are not part of the internal language but rather part of the ordinary language of mathematics in which

we make statements about algebras, topoi, terms, etc. By contrast, note that nothing in the internal language has yet been interpreted as a statement. When it is being contrasted with a formal language or theory, ordinary mathematics (in which one can talk about the formal language as well as other things) is sometimes called the meta-theory.

It is easy to show, by induction on t, that if t is a term on V and $V \subseteq W$ then t is a term on W and $|t|_W$ is the composite of $|t|_V$ and the projection $\operatorname{sort}(W) \to \operatorname{sort}(V)$. Therefore, if $V \subseteq W$, then $t_1 \equiv_V t_2$ implies $t_1 \equiv_W t_2$, though the converse can fail if the projection is not an epimorphism.

Although rather poor in basic concepts, the internal logic of a topos turns out to be rich in expressive power. To indicate how this happens, we now introduce a number of common mathematical concepts as terms in this logic. (The situation is similar to that of classical set theory, where one begins with a very small supply of basic concepts but eventually introduces enough definitions to express almost all mathematical concepts.)

We begin by recalling that, when $n = 0$, clause (b) of the definition of terms says that every morphism $1 \to A$ is a term (on any V) of sort A. We write $*$ for the term id_1 of sort 1. Clearly, if $f\colon 1 \to A$, then $f(*) \equiv f$, where the f on the left is the unary operation symbol obtained by viewing the domain 1 of f as the product of a single factor 1, while the f on the right is obtained by viewing 1 as the product of no factors. (We could also view 1 as 1×1 and obtain $f(*, *) = f$, etc.) Note that, for any $t\,\varepsilon\, 1$, we have $t \equiv *$ because 1 is terminal.

The evaluation map $B^A \times A \to B$, i.e., the counit of the adjunction between product and exponential, treated as a binary operation symbol, will be denoted by ev (or $\mathrm{ev}_{A,B}$ if such precision is needed) and usually written between its arguments. Thus, if $t_1 \,\varepsilon\, B^A$ and $t_2 \,\varepsilon\, A$ then $(t_1 \mathrm{ev} t_2)\,\varepsilon\, B$. (We omit the subscript on ε since it is irrelevant, but of course it should be the same on all three ε's.) If $f\colon A \to B$ corresponds, via the adjunction, to $\ulcorner f \urcorner\colon 1 \to B^A$, then, for any $t\,\varepsilon\, A$, $(\ulcorner f \urcorner \mathrm{ev}\, t) \equiv f(t)$, as the reader may easily verify. Also, if $t \,\varepsilon_V\, B^A$ and if x is a variable of sort A that is not in V, then $x \mapsto (t \mathrm{ev} x) \equiv_V t$. Finally, there is the "λ-conversion" rule, $(x \mapsto t)\mathrm{ev}\, a \equiv t[a/x]$, where $t\,\varepsilon\, B$, $a \,\varepsilon\, A$, x is a variable of sort A, and the substitution is assumed to be legitimate. In particular, $(x \mapsto t)\mathrm{ev}\, x \equiv t$.

The identity map of $A\times B$, with its domain represented as the product of A and B, gives a binary operation symbol, written "$\langle\ ,\ \rangle$" around its arguments. Thus, if $a\,\varepsilon\, A$ and $b\,\varepsilon\, B$ then $\langle a, b\rangle\,\varepsilon\, A\times B$. We write 1^{st} and 2^{nd} for the projections of $A\times B$ to A and B. So $1^{\mathrm{st}}(\langle a, b\rangle) \equiv a$, $2^{\mathrm{nd}}(\langle a, b\rangle) \equiv b$, and, if $c\,\varepsilon\, A\times B$, $\langle 1^{\mathrm{st}}(c), 2^{\mathrm{nd}}(c)\rangle \equiv c$. Products of more than two factors are treated similarly.

Everything we have said about the internal language until now would make sense for any cartesian closed category. The additional structure available because $\mathscr{E}$ is a topos, namely the subobject classifier Ω, leads both to a deeper mathematical structure and to a different way of viewing certain terms of the internal logic. It is worthwhile to introduce appropriate notation to reflect this new viewpoint before going on to the new mathematical content. A term of sort Ω will be called a *formula*;

a closed formula is a *sentence*. Note that the denotation of a sentence φ is a truth value of $\mathscr{E}$, i.e. a global section of Ω, $|\varphi|_{\emptyset}\colon 1 \to \Omega$. If φ is a formula, we often write $x \mapsto \varphi$ as $\{x|\varphi\}$; this is intuitively reasonable since this term denotes a morphism into $\Omega^A = \mathscr{P}(A)$, the object of subobjects of A, where $A = \mathrm{sort}(x)$. For an even more suggestive notation, we may indicate the sort of the variable by writing $\{x \varepsilon A|\varphi\}$. The map $A \times \Omega^A \to \Omega$ obtained by interchanging the factors in the domain of ev will be written $\in$ (or $\in_A$) between its arguments. Note carefully the distinction between $\in$, which denotes a morphism of $\mathscr{E}$ and is therefore a symbol of the internal language, and ε, which is used in the meta-theory to indicate the sorts of terms. Results about ev mentioned earlier can be rewritten in terms of $\in$ to give the familiar-looking equivalences $\{x|x \in s\} \equiv_V s$, where $s \varepsilon_V \mathscr{P}(A)$ and x is a variable of sort A not in V, and $a \in \{x|\varphi\} \equiv \varphi[a/x]$, with the usual assumption that the substitution is legitimate.

Using, for the first time, the fact that Ω is the subobject classifier, we introduce the notation $=_A$, a binary operation symbol written between its arguments, for the morphism $A \times A \to \Omega$ that classifies the diagonal subobject $A \underset{\langle \mathrm{id},\mathrm{id}\rangle}{\rightarrowtail} A \times A$. When no confusion seems likely, we shall omit the subscript A and just write $=$. As with $\in$ and ε, note carefully that $=$ is in the internal language while $\equiv$ is in the meta-theory.

Using the equality symbol, one can define all the other basic concepts of intuitionistic logic as terms in the internal language. It is well-known that, for example, universal quantification over an object A is represented by a morphism $\Omega^A \to \Omega$. Either this morphism, as a unary operation symbol, or its exponential adjoint $1 \to \mathscr{P}\mathscr{P}(A)$, as a constant term, could serve as the universal quantifier of our internal logic. The latter is, in fact, equivalent to the term $\forall_A$ introduced below. We have chosen, instead, to define $\forall_A$ and the other logical concepts as shorthand for certain terms involving equality. The reason is primarily a matter of aesthetics and economy, but the formulas below also provide a way of constructing the morphisms of $\mathscr{E}$ that internalize these logical concepts and might thus be useful in presenting the fundamentals of topos theory. Before presenting these definitions which are part of the folklore of logic and may also be found in [21], it may be worthwhile to comment on the fact that we are apparently defining the logical concepts in any cartesian closed category with an object Ω equipped with equality morphisms $A \times A \to \Omega$ for all A; nothing more about the topos structure of $\mathscr{E}$ has been used. It is indeed possible to define the logical concepts in this generality, but the proof that they work properly, i.e. satisfy the expected equivalences, depends on having a topos or at least something fairly close to a topos, such as a logical category [21, 36, 37]. We now turn to the definitions; the symbol $:\equiv$ is used for definitional identity, since we never really care that two terms are identical, only that they are equivalent.

Recall that $*$ is the constant term given by the identity map of 1

(1) $\quad \mathrm{true} :\equiv * = * \ (\text{so } \mathrm{true}\ \varepsilon_{\emptyset}\ \Omega)$,

(2) $\quad \{\ \}_A :\equiv x \mapsto \{y|y = x\}, \quad \text{where} \quad x, y\ \varepsilon\ A \ (\text{so } \{\ \}_A\ \varepsilon_{\emptyset}\ \mathscr{P}(A)^A)$.

The reader should verify that the denotation $|\{\ \}_A|_{\emptyset}\colon 1 \to \mathscr{P}(A)^A$ is the adjoint of the usual singleton map $A \to \mathscr{P}(A)$. Similar verifications justify the other definitions given here. We usually abbreviate $\{\ \}_A$ ev a as $\{a\}_A$, and we omit the subscript A when possible

(3) $\quad \wedge :\equiv \{\langle \mathrm{true}, \mathrm{true}\rangle\} \quad (\text{so } \wedge\ \varepsilon_{\emptyset}\ \mathscr{P}(\Omega \times \Omega) = \Omega^{\Omega\times\Omega})$.

$|\wedge|_{\emptyset}\colon 1 \to \Omega^{\Omega\times\Omega}$ is the adjoint of the usual conjunction map $\Omega \times \Omega \to \Omega$. We abbreviate $\wedge\,\mathrm{ev}\langle\varphi, \psi\rangle$ as $\varphi \wedge \psi$, and similarly for the other binary connectives defined below ($\vee$, $\to$, $\leftrightarrow$). Thus,

$$\varphi \wedge \psi \equiv \langle\varphi, \psi\rangle \in \{\langle \mathrm{true}, \mathrm{true}\rangle\} \equiv \langle\varphi, \psi\rangle = \langle \mathrm{true}, \mathrm{true}\rangle,$$

(4) $\quad \mathrm{all}_A :\equiv \{x \in A|\mathrm{true}\} \quad (\text{so } \mathrm{all}_A\ \varepsilon_{\emptyset}\ \mathscr{P}(A))$,

(5) $\quad \forall_A :\equiv \{\mathrm{all}_A\}_{\mathscr{P}(A)} \quad (\text{so } \forall_A\ \varepsilon_{\emptyset}\ \mathscr{P}\mathscr{P}(A))$.

We abbreviate $\forall_A \mathrm{ev}\{x|\varphi\}$ as $\forall x\varphi$ or $\forall x \varepsilon A\varphi$, and similarly for the other quantifiers defined below ($\exists$, $\exists!$). Thus, $\forall x \varepsilon A \varphi \equiv \{x \varepsilon A|\varphi\} = \mathrm{all}_A$.

(6) $\quad \to :\equiv \{z \varepsilon \Omega \times \Omega | 1^{\mathrm{st}}(z) = \wedge\,\mathrm{ev}\,z\}$.

Thus, $\varphi \to \psi \equiv \varphi = (\varphi \wedge \psi)$. In the future, when writing terms of type $\mathscr{P}(A \times B)$, we shall frequently write, not the formally correct $\{z \varepsilon A \times B|\theta\}$, but the more legible $\{\langle a, b\rangle \varepsilon A \times B|\theta[\langle a, b\rangle/z]\}$. Thus, for example, the right side of (6) would be, after simplification, $\{\langle\varphi, \psi\rangle \varepsilon \Omega \times \Omega|\varphi = (\varphi \wedge \psi)\}$,

(7) $\quad \leftrightarrow :\equiv \{\langle\varphi, \psi\rangle \varepsilon \Omega \times \Omega|\varphi = \psi\}$,

(8) $\quad \vee :\equiv \{\langle\varphi, \psi\rangle \varepsilon \Omega \times \Omega|\forall u \varepsilon \Omega[((\varphi \to u) \wedge (\psi \to u)) \to u]\}$,

(9) $\quad \mathrm{false} :\equiv \forall u \varepsilon \Omega u$,

(10) $\quad \neg :\equiv \{\mathrm{false}\}_{\Omega}$.

We abbreviate $\neg\mathrm{ev}\varphi$ as $\neg\varphi$, so $\neg\varphi \equiv \varphi = \mathrm{false}$,

(11) $\quad \exists_A :\equiv \{z \varepsilon \mathscr{P}(A)|\forall u \varepsilon \Omega[\forall y \varepsilon A(y \in z \to u) \to u]\}$,

(12) $\quad \exists!_A :\equiv \{z \varepsilon \mathscr{P}(A)|\exists y \varepsilon A\ z = \{y\}\}$.

This completes the definition of the logical connectives and quantifiers in the internal language. Using them, along with $=$ and $\in$, we can interpret many statements of ordinary set theory as formulas in the internal language. This is the first main step in translating arguments in set theory into proofs about topoi.

We say that a formula φ on V is *valid* (on V), and we write $\vDash_V \varphi$ (omitting the subscript when V is clear or irrelevant) to mean that $\varphi \equiv_V \mathrm{true}$, i.e. that $|\varphi|_V\colon \mathrm{sort}(V) \to \Omega$ factors through $|\mathrm{true}|_{\emptyset}\colon 1 \to \Omega$. The second main ingredient in translating set theory into topos theory is the theorem [10, 29] that many-sorted intuitionistic logic is sound for this notion of validity. More precisely, if ψ is intuitionistically deducible from $\{\varphi_i|i<n\}$ and if $\vDash_V \varphi_i$ for all $i<n$, then $\vDash_V \psi$. A word of caution is in order concerning the frequently used definition of $\vDash \varphi$ as $\vDash_V \varphi$ for

the smallest V such that $\varphi\,\varepsilon_V\,\Omega$, i.e. for V the set of free variables of φ. The usual formulations of intuitionistic logic are not sound for this definition of $\models$, because the V's involved in determining the validity of the premises φ_i and the conclusion ψ may be different. There is no difficulty, however, if all variables free in any premise are also free in the conclusion; see [29] for details.

We now turn from logic to set theory. Again, we shall introduce the basic concepts as terms in the internal logic. Since some of the definitions would be rather cumbersome if written out in full, we adopt some stylistic conventions to make them easier to read. Rather than explain these conventions in abstract terms, we indicate how they are used in some examples.

Consider, for example, the notion of the domain of a relation between A and B. The object of such relations is $\mathscr{P}(A\times B)$, while the object of subsets of A (possible domains) is $\mathscr{P}(A)$, so domain ought to be of sort $\mathscr{P}(A)^{\mathscr{P}(A\times B)}$. A straightforward definition of it would be

(12) $$\mathrm{domain}_{A,B}:\equiv z\mapsto\{x\,\varepsilon A\,|\,\exists y\,\varepsilon B\,\langle x,y\rangle\in z\},$$

and this is, in fact, the definition we shall adopt. We should add to it the convention (similar to the conventions adopted in logic above) that $\mathrm{domain}_{A,B}\mathrm{ev}\,r$ is to be abbreviated $\mathrm{domain}_{A,B}(r)$, with the subscripts omitted when possible. We can succinctly express both the definition (12) and this convention by writing

(12′) $$\mathrm{domain}_{A,B}(z):\equiv\{x\,\varepsilon A\,|\,\exists y\,\varepsilon B\,\langle x,y\rangle\in z\}$$

or

(12″) $$x\in\mathrm{domain}_{A,B}(z):\equiv\exists y\,\varepsilon B\,\langle x,y\rangle\in z.$$

Both (12′) and (12″) should be viewed as nothing more than shorthand ways of writing (12) and the notational convention for $\mathrm{domain}_{A,B}\mathrm{ev}\,r$. In more complicated situations, it would be useful to indicate the sorts of the variable by writing the left sides of (12′) and (12″) as $\mathrm{domain}_{A,B}(z\,\varepsilon\mathscr{P}(A\times B))$ and $x\,\varepsilon A\in\mathrm{domain}_{A,B}(z\,\varepsilon\mathscr{P}(A\times B))$. We may also omit the subscripts of domain in (12′) and (12″); they are to be tacitly understood.

Similarly, we write

(13) $z\,\varepsilon\mathscr{P}(A\times B)$ is a function
$$:\equiv\forall x\,\varepsilon A\,\forall y\,\varepsilon B\,\forall y'\,\varepsilon B\,[(\langle x,y\rangle\in z\wedge\langle x,y'\rangle\in z)\to y=y'].$$

This is to be understood as abbreviating a definition, $\mathrm{function}_{A,B}:\equiv\{z|\forall x\,\varepsilon A\,\ldots\}$, of a closed term $\mathrm{function}_{A,B}\,\varepsilon\mathscr{P}\mathscr{P}(A\times B)$, together with the convention that $\mathrm{function}_{A,B}\mathrm{ev}\,r$ is to be written "r is a function." The definitions that follow are to be understood in the same way. Since this way of writing definitions makes them practically identical with the definitions in any standard set theory text, we give only some representative examples, assuming that the reader can easily supply any others that may be needed. (In particular, having defined domain, we refuse to define range.)

$$(x\,\varepsilon\mathscr{P}(A))\subseteq(y\,\varepsilon\mathscr{P}(A)):\equiv\forall v\,\varepsilon A\,[v\in x\to v\in y],$$
$$(x\,\varepsilon\mathscr{P}(A))\cap(y\,\varepsilon\mathscr{P}(A)):\equiv\{z\,\varepsilon A\,|\,z\in x\wedge z\in y\},$$
$$\bigcap(x\,\varepsilon\mathscr{P}\mathscr{P}(A))=\{z\,\varepsilon A\,|\,\forall y\,\varepsilon\mathscr{P}(A)\,[y\in x\to z\in y]\}.$$

Unions are defined analogously.

$$(x\,\varepsilon\mathscr{P}(A))\times(y\,\varepsilon\mathscr{P}(B)):\equiv\{\langle a,b\rangle\,\varepsilon A\times B\,|\,a\in x\wedge b\in y\},$$
$$(z\,\varepsilon\mathscr{P}(A\times B))\restriction(x\,\varepsilon\mathscr{P}(A)):\equiv z\cap(x\times\mathrm{all}_B),$$
$$\mathrm{Converse}\;(z\,\varepsilon\mathscr{P}(A\times B)):\equiv\{\langle b,a\rangle\,\varepsilon B\times A\,|\,\langle a,b\rangle\in z\}$$
$$(z\,\varepsilon\mathscr{P}(A\times B))\text{ is one to one}:\equiv\mathrm{Converse}\;(z)\text{ is a function},$$
$$(x\,\varepsilon\mathscr{P}(B\times C))\circ(y\,\varepsilon\mathscr{P}(A\times B))$$
$$=\{\langle a,c\rangle\,\varepsilon A\times C\,|\,\exists b\,\varepsilon B\,[\langle a,b\rangle\in y\wedge\langle b,c\rangle\in x]\}.$$

Definitions like these make translating statements of ordinary set theory into the internal language essentially just a matter of copying. To translate proofs, we use the soundness of intuitionistic logic for $\models_V$ along with the fact that the basic principles of set theory are valid (or, more precisely, have valid translations) in the internal interpretation. The most important of these are the comprehension principle

$$a\in\{x|\varphi\}\leftrightarrow\varphi[a/x]\text{ when the substitution is legitimate,}$$

and the extensionality principle

$$\forall x\,\varepsilon A\,(x\in z\leftrightarrow x\in z')\to z=_{\mathscr{P}(A)}z'.$$

The validity of the first of these follows immediately from $a\in\{x|\varphi\}\equiv\varphi[a/x]$; for the second, which is a special case of the extensionality principle for maps

$$\forall x\,\varepsilon A\,((z\,\mathrm{ev}\,x)=_B(z'\,\mathrm{ev}\,x))\to z=_{B^A}z',$$

see [29]. Among other familiar set theoretic principles whose internal validity we shall use, we mention Leibniz's law

$$x=_A y\leftrightarrow(\forall z\,\varepsilon\mathscr{P}(A))[x\in z\leftrightarrow y\in z],$$

and the characterization of products

$$\forall x\,\varepsilon A\times B\,\exists a\,\varepsilon A\,\exists b\,\varepsilon B\;x=\langle a,b\rangle\wedge$$
$$\wedge\forall a,a'\,\varepsilon A\,\forall b,b'\,\varepsilon B\,[\langle a,b\rangle=\langle a',b'\rangle\to a=a'\wedge b=b']$$

and similarly for products of more factors (or fewer: $\forall x\,\varepsilon 1\;x=*$).

On the basis of such principles, we can translate into the internal language those proofs of ordinary set theory which do not use the law of the excluded middle or the axioms of choice, infinity, and replacement (which can fail in topoi). If the topos has a natural numbers object, proofs using the axiom of infinity can be internalized, while in Boolean topoi the law of the excluded middle

$$\forall x \varepsilon \Omega [x \vee \neg x]$$

is valid. Proofs that use the excluded middle, replacement, or choice can often be reformulated so that they do not. Indeed, much of Sections 4 through 6 of this paper can be viewed as just such a reformulation of the classical construction of absolutely free algebras.

The preceding discussion may be summarized by saying that, to prove $\models \theta$ for some given sentence θ in the internal language, it suffices to give an intuitionistic proof of θ on the basis of set theoretic principles, such as those listed above, that are internally valid. We usually indicate our intention to employ this method of proof by saying "We carry out the proof in the internal logic" or words to that effect. Once this is said, we feel free to use all the usual modes of deduction (e.g. to prove $\varphi \to \psi$, assume φ and prove ψ; to prove $\forall x \varphi$, fix an arbitrary x and prove φ) insofar as they are intuitionistically correct. We also feel free to use standard notation, such as $f(x)$ when we know that $f \varepsilon \mathscr{P}(A \times B)$ is a function and $x \in \mathrm{domain}(f)$.

We close this section with some comments on the relation between the notion of function defined by (13) and the notion of function as something of sort B^A. There is a canonical morphism Graph: $B^A \to \mathscr{P}(A \times B) = \Omega^{A \times B}$, obtained by exponential adjunction from

$$B^A \times A \times B \xrightarrow[\mathrm{ev} \times \mathrm{id}]{} B \times B \xrightarrow[=_B]{} \Omega .$$

We have (for $a \varepsilon A$, $b \varepsilon B$, $f \varepsilon B^A$) $\langle a, b \rangle \in \mathrm{Graph}(f) \equiv (f \mathrm{ev} a) = b$, from which it easily follows that $\models \mathrm{Graph}(f)$ is a function $\wedge$ $\mathrm{domain}(\mathrm{Graph}(f)) = \mathrm{all}_A$. Thus, every function in the B^A sense gives rise to a function in the sense of (13) with domain all_A. Conversely, every such total function arises from an element of B^A; more precisely,

$$\models \forall z \varepsilon \mathscr{P}(A \times B)[(z \text{ is a function } \wedge \text{ domain } (z) = \mathrm{all}_A) \to \exists! f \varepsilon B^A \ \mathrm{Graph}(f) = z] .$$

The verification of this is essentially an internal version of Kock's construction [20] of exponentials from power objects; we leave the details to the reader.

3. Externalization. This section is devoted to pointing out the connections between properties of objects and morphisms expressed as validity of certain internal formulas and properties expressed in conventional category-theoretic terms. Proofs of most of the results are left to the reader, with the suggestion that he consult [21] in case of difficulty.

Suppose a and b are terms on V of sort A and $\models_V a =_A b$. This means that

$$\mathrm{sort}(V) \xrightarrow[\langle |a|_V, |b|_V \rangle]{} A \times A \xrightarrow[=_A]{} \Omega$$

factors through $|\mathrm{true}|$: $1 \to \Omega$, which is equivalent, in view of the definition of $=_A$, to saying that $\langle |a|_V, |b|_V \rangle$ factors through the diagonal $A \to A \times A$, i.e. that $|a|_V = |b|_V$. So $\models_V a = b$ is equivalent to $a \equiv_V b$. In particular, to say that two

(global) elements ξ, η: $1 \to A$ are equal as morphisms is the same as to say that (it is valid that) they are internally equal as constant terms, $\models \xi =_A \eta$, since clearly $|\xi|_\varnothing = \xi$. By virtue of extensionality, it follows that two morphisms ξ, η: $A \to B$ are equal if and only if $\models \forall a \varepsilon A [\xi(a) = \eta(a)]$.

Composition of morphisms is also expressible internally. If f: $\prod_{j<n} B_j \to C$ and, for each $j<n$, g_j: $\prod_{i<m} A_i \to B_j$, then h is the composite of f with $\langle g_j \rangle_{j<n}$ if and only if

$$\models \forall a_0 \varepsilon A_0 \ldots \forall a_{m-1} \varepsilon A_{m-1} [h(a_0, \ldots, a_{m-1}) = f(g_0(a_0, \ldots, a_{m-1}), \ldots, g_{n-1}(a_0, \ldots, a_{m-1}))] .$$

Identity morphisms are characterized by $\models \forall a \varepsilon A \ \mathrm{id}_A(a) =_A a$.

A morphism f: $A \to B$ is monic if and only if

$$\models \forall a, a' \varepsilon A [f(a) =_B f(a') \to a =_A a'] .$$

It is epic if and only if

$$\models \forall b \varepsilon B \exists a \varepsilon A f(a) =_B b .$$

It is an isomorphism if and only if

$$\models \forall b \varepsilon B \exists! a \varepsilon A f(a) =_B b .$$

We leave it to the reader to give similar characterizations of products, coproducts, pullbacks, etc.; the natural descriptions in the category of sets, when expressed in the internal logic, work in any topos.

There are two important processes for constructing objects and morphisms of $\mathscr{E}$ when we are given descriptions of them in the internal logic. The first of these amounts to invoking the fact that Ω classifies subobjects. Suppose we have a closed term $t \varepsilon \mathscr{P}(A)$. Then its denotation $|t|_\varnothing$: $1 \to \mathscr{P}(A)$ yields by exponential adjointness a morphism $\hat{t}$: $A \to \Omega$, which classifies a subobject $[t] \underset{i}{\rightarrowtail} A$ of A. Thus, any closed term t of sort $\mathscr{P}(A)$ gives rise to an object $[t]$ with a canonical monomorphism i into A. These are characterized by $\models \forall x \varepsilon A [x \in t \leftrightarrow \exists y \varepsilon [t] x = i(y)]$, where the quantifier $\exists y$ can be replaced with $\exists ! y$ because i is monic. A morphism f: $X \to A$ factors through i: $[t] \rightarrowtail A$ if and only if $\models \forall x \varepsilon X f(x) \in t$. We abbreviate $[\{x \varepsilon A | \varphi\}]$ to $[x \varepsilon A | \varphi]$.

As an example, $[z \varepsilon \mathscr{P}(A \times B) | z$ is a function $\wedge$ $\mathrm{domain}(z) = \mathrm{all}_A]$ is, in virtue of the discussion at the end of Section 2, simply B^A, the inclusion into $\mathscr{P}(A \times B)$ being Graph. As another example, a pullback

$$\begin{tikzcd} P \arrow[r, "\beta"] \arrow[d, "\alpha"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

can be obtained as $P = [\langle a, b \rangle \varepsilon A \times B | f(a) = g(b)]$ with $\alpha = 1^{\mathrm{st}}\, i$ and $\beta = 2^{\mathrm{nd}}\, i$, where i: $P \rightarrowtail A \times B$ is the inclusion. In classical set theory, one always takes inclusion maps to be restrictions of the identity map. While this is not possible (or even meaningful) in the present context, it is useful to adopt a notation making these inclusion maps as unobtrusive as possible. A convenient choice is a dot over a symbol. If we abbreviate $i(x)$ as $\dot{x}$ (or as $(x)^{\cdot}$ when x is a complex expression) it is easy to ignore the dots when reading, and then the formulas look like classical set theory. (We try to avoid having several inclusion maps in the same context, lest we need several shapes of dots.)

The second externalization process that we wish to discuss produces morphisms; in its simplest form it produces morphisms with domain 1. Suppose $t\,\varepsilon\,\mathscr{P}(A)$ is a closed term such that $\models \exists ! x\, x \in t$. Then $|t|\colon 1 \to \mathscr{P}(A)$ factors through the singleton map $|\{x\}|_{\{x\}}\colon A \to \mathscr{P}(A)$ yielding a morphism $\xi\colon 1 \to A$. This fact is essentially Kock's principle that unique local (i.e. internal) existence implies global existence; see [19]. The defining property of ξ is expressed by $\models t = \{\xi\}$ or, in view of $\models \exists ! x\, x \in t$, simply $\models \xi \in t$. In applications, we frequently have a formula $\varphi\,\varepsilon_{\{x\}}\,\Omega$ such that $\models \exists ! x \varphi$; then the preceding applies to $t = \{x | \varphi\}$ and provides $\xi\colon 1 \to A$ such that $\models \varphi[\xi/x]$. We refer to this sort of externalization (and its generalizations described below) as Kock's principle.

To similarly define a morphism from Y to X, suppose we have a closed term $t\,\varepsilon\,\mathscr{P}(Y\times X)$ such that $\models \forall y\,\varepsilon\, Y \exists ! x\,\varepsilon\, X \langle y, x\rangle \in t$. Thus, $\models t$ is a function $\wedge$ domain $(t) = \mathrm{all}_Y$, so, by the discussion at the end of Section 2, $\models \exists ! f\,\varepsilon\, X^Y\ \mathrm{Graph}(f) = t$. By Kock's principle, there is a (unique) $\hat{\xi}\colon 1 \to X^Y$ such that $\models \mathrm{Graph}(\hat{\xi}) = t$. For its exponential adjoint $\xi\colon Y \to X$, we have $\models \forall y \forall x[\langle y, x\rangle \in t \leftrightarrow x = \xi(y)]$, or, in view of $\models \forall y \exists ! x \langle y, x\rangle \in t$, simply $\models \forall y \langle y, \xi(y)\rangle \in t$.

The two externalization principles can be combined to yield the following process for defining a partial morphism from Y to X. Suppose we have a closed term $t\,\varepsilon\,\mathscr{P}(Y\times X)$ such that $\models t$ is a function. Then there exist a subobject $W \underset{i}{\rightarrowtail} Y$, namely $[\mathrm{domain}(t)]$, and a morphism $\xi\colon W \to X$ such that

$$\models \forall y\,\varepsilon\, Y \forall x\,\varepsilon\, X[\langle y, x\rangle \in t \leftrightarrow \exists w\,\varepsilon\, W(y = i(w) \wedge x = f(w))].$$

To obtain ξ, just apply Kock's principle to $\{\langle w, x\rangle\,\varepsilon\, W\times X | \langle i(w), x\rangle \in t\}$.

To prepare for one of the calculations in Section 6, and also to give a non-trivial example of a construction where one first describes what one wants in the internal language and then externalizes, we present a characterization of the object $\Sigma_J \Pi_\tau \Delta_I A$, which, as we have seen, occurs in the definition of τ-algebra.

THEOREM. *Let $\tau\colon I \to J$ be a morphism and A an object of the topos $\mathscr{E}$. Define $S\colon \equiv \{\langle j, f\rangle \in J\times\mathscr{P}(I\times A) | f$ is a function $\wedge$ domain$(f) = \{y\,\varepsilon\, I|\ \tau(y) = j\}\}$. Then $[S] = \Sigma_J \Pi_\tau \Delta_I A$.*

Proof. $[S]$ comes equipped with an inclusion i into $J\times\mathscr{P}(I\times A)$; we write pr for its composite with the projection $J\times\mathscr{P}(I\times A) \to J$. We shall show that $\mathrm{pr}\colon [S] \to J$ is $\Pi_\tau \Delta_I A$, which suffices to complete the proof since Σ_J sends every morphism with codomain J to its domain. In view of the definition of Π_τ as right adjoint to Δ_τ we must establish a natural bijection between morphisms in $\mathscr{E}/J$ from any $X \underset{\xi}{\to} J$ to $[S] \underset{\mathrm{pr}}{\to} J$ and morphisms in $\mathscr{E}/I$ from $\Delta_\tau(X \underset{\xi}{\to} J)$ to $\Delta_I A$. We shall construct the bijec-

tion for a fixed $X \underset{\xi}{\to} J$; naturality is easy and will be left to the reader. Having fixed $X \underset{\xi}{\to} J$, we construct the pullback

$$\begin{tikzcd} P \arrow[r, "q"] \arrow[d, "p"'] & X \arrow[d, "\xi"] \\ I \arrow[r, "\tau"'] & J \end{tikzcd}$$

where $P \underset{p}{\to} I$ is $\Delta_\tau(X \underset{\xi}{\to} J)$. In view of the definitions of $\mathscr{E}/I$, Δ_I, etc., we are seeking a bijection between commutative triangles of the form

(1)
$$\begin{tikzcd} X \arrow[rr, "\alpha"] \arrow[dr, "\xi"'] & & {[S]} \arrow[dl, "\mathrm{pr}"] \\ & J & \end{tikzcd}$$

and commutative triangles of the form

(2)
$$\begin{tikzcd} P \arrow[rr, "\beta"] \arrow[dr, "p"'] & & I\times A \arrow[dl, "1\mathrm{st}"] \\ & I & \end{tikzcd}$$

Any β as in (2) gives rise to a map $\gamma\colon P \underset{\beta}{\to} I\times A \underset{2\mathrm{nd}}{\to} A$, which completely determines β since the commutativity of (2) is equivalent to $\beta = \langle p, \gamma\rangle$. Thus, commutative triangles of the form (2) correspond bijectively to morphisms $\gamma\colon P \to A$.

Now consider any α as in (1). Its composite with $i\colon [S] \rightarrowtail J\times\mathscr{P}(I\times A)$ can be written as $\langle \xi, \eta\rangle\colon X \to J\times\mathscr{P}(I\times A)$, where the first component is ξ because (1) commutes. The fact that $\langle \xi, \eta\rangle$ factors through $[S]$ can be expressed as

(3) $\models \forall x\,\varepsilon\, X[\eta(x)$ is a function $\wedge$ domain$(\eta(x)) = \{y\,\varepsilon\, I|\ \tau(y) = \xi(x)\}]$.

Let us abbreviate the term $\{\langle y, x\rangle\,\varepsilon\, I\times X|\ \tau(y) = \xi(x)\}$ as t. (By previous remarks, $[t] = P$.) The domain clause in (3) can be rewritten as

$$\forall y\,\varepsilon\, I[y \in \mathrm{domain}(\eta(x)) \leftrightarrow \langle y, x\rangle \in t].$$

Set

$$\bar{\eta}\colon \equiv \{\langle\langle y, x\rangle, u\rangle\varepsilon(I\times X)\times A|\ \tau(y) = \xi(x) \wedge \langle y, u\rangle \in \eta(x)\}.$$

By the preceding, $\models \bar{\eta}$ is a function $\wedge$ domain$(\bar{\eta}) = t$. By Kock's principle, $\bar{\eta}$ defines a morphism $\gamma\colon P = [t] \to A$. Conversely, any such γ corresponds to $\bar{\eta}$ for a unique η satisfying (3), namely

$$|\{\langle y, u\rangle\,\varepsilon\, I\times A|\ \exists z\,\varepsilon\, P[p(z) = y \wedge q(z) = x \wedge \gamma(z) = u]\}|_{\{x\}}.$$

Thus, morphisms $\gamma\colon P \to A$ correspond bijectively to morphisms $\eta\colon X \to \mathscr{P}(I\times A)$ satisfying (3) and therefore also to commutative triangles (1). ■

This theorem allows us to define a τ-algebra structure on A, i.e. a morphism $\Sigma_J \Pi_\tau \Delta_I A \to A$, by giving a closed term $\mu \varepsilon \mathscr{P}\big((J\times\mathscr{P}(I\times A))\times A\big)$ and proving

$$\models \forall j \varepsilon J \forall m \varepsilon \mathscr{P}(I\times A)\big[m \text{ is a function} \wedge \text{domain}(m) = \{i \varepsilon I \mid \tau(i) = j\} \\ \to \exists! a \varepsilon A \big\langle\langle j, m\rangle, a\big\rangle \in \mu\big],$$

for then Kock's principle can be applied to $t \upharpoonright S$ to yield the desired morphism. We shall in fact define the algebra structure of the object of words in just this way in Section 6. (Indeed, any such morphism arises in this way from some term μ.)

Notice also that a homomorphism from one algebra (A, μ) to another (B, ν) can be obtained by applying Kock's principle to any closed term $\alpha \varepsilon \mathscr{P}(A\times B)$ such that

$$\models \alpha \text{ is a function} \wedge \text{domain}(\alpha) = \text{all}_A \wedge \\ \wedge \forall j \varepsilon J \forall m \varepsilon \mathscr{P}(I\times A) \forall a \varepsilon A\big[\big\langle\langle j, m\rangle, a\big\rangle \in \mu \to \big\langle\langle j, \alpha\circ m\rangle, \alpha(a)\big\rangle \in \nu\big].$$

Furthermore, all homomorphisms arise in this way, and $\models \alpha = \alpha'$ if and only if the corresponding homomorphisms are the same.

4. Well-foundedness. This section is devoted to developing, within the internal logic of a topos $\mathscr{E}$, enough of the classical theory of well-founded relations to permit the definition of functions by induction over such relations. We shall work with a fixed object X and, to avoid repetitious description of sorts, we agree to use x, y, z as variables of sort X, a, b, c as variables of sort $\mathscr{P}(X)$, and $\prec$ as a variable of sort $\mathscr{P}(X)^X$. Furthermore, we shall write $\prec x$ for $\prec \text{ev} x$, and $y \prec x$ for $y \in (\prec x)$. As a final convention, we abbreviate $\forall x[x\in b \to \ldots]$ and $\exists x[x \in b \wedge \ldots]$ as $\forall x \in b \ldots$ and $\exists x \in b \ldots$ respectively.

DEFINITIONS. a is $\prec$ inductive on b: $\equiv \forall x \in b[\forall y \in b(y \prec x \to y \in a) \to x \in a]$. $\prec$ is well-founded on b: $\equiv \forall a[a \text{ is } \prec \text{ inductive on } b \to b \subseteq a]$. In accordance with the stylistic conventions of Section 2, the first of these is to be taken as the definition of a term inductive $\varepsilon_{\mathfrak{g}}\big((\Omega^{\mathscr{P}(X)})^{\mathscr{P}(X)^X}\big)^{\mathscr{P}(X)}$ together with a convention for writing "((inductive ev$a$)ev$\prec$)ev$b$" as "a is $\prec$ inductive on b"; similarly for the definition of well-founded and the further definitions below.

A few easy (intuitionistic) consequences of these definitions are worth pointing out before proceeding to the problem of defining functions by recursion. First, applying the definition of well-foundedness with $a = \{x \mid \varphi\}$, we obtain the principle of proof by induction

$$\models \prec \text{ is well-founded on } b \wedge \forall x \in b[\forall y \in b(y \prec x \to \varphi[y/x]) \to \varphi] \to \forall x \in b \varphi$$

where y is not free in φ and the substitution is legitimate. Second, if a is $\prec$ inductive on b, then so is $a \cap b$; therefore

$$\prec \text{ is well-founded on } b \equiv \forall a[a \subseteq b \wedge a \text{ is } \prec \text{ inductive on } b \to a = b].$$

Finally, as far as the definition of well-foundedness on b is concerned, the predecessors outside b of any x are irrelevant. More precisely,

$$\prec \text{ is well-founded on } b \equiv (x \mapsto b \cap \prec x) \text{ is well-founded on } b.$$

This is obvious since the same sets are inductive with respect to $x \mapsto (b \cap \prec x)$ as with respect to $\prec$.

It should be mentioned that our definition is classically but not intuitionistically equivalent to the more common definition requiring the existence of minimal elements in all nonempty subsets of b. It is fairly well known [35, 3] that the latter definition is inappropriate in intuitionistic contexts; we comment further on this in the appendix to this section.

We now turn to the formulation and proof of a theorem on definition by recursion. This theorem is intuitively very similar to one proved by Mikkelsen [25, Appendix], but, unlike that theorem, ours is formulated in the internal language, and its proof is essentially the same as one of the standard proofs of the corresponding theorem in ordinary set theory (see [9, 16]). We fix an object Y to serve as the codomain of the maps to be defined by recursion, and we let p, q be variables of sort Y, while f, g, h are of sort $\mathscr{P}(X\times Y)$.

DEFINITIONS. h is data under x (with respect to $\prec$ on b): $\equiv h$ is a function $\wedge$ $\wedge$ domain$(h) = b \cap \prec x$. G is a recursion condition (with respect to $\prec$ on b): $\equiv G$ is a function $\wedge$ domain$(G) = \{\langle x, h\rangle \mid x \in b \wedge h \text{ is data under } x\}$. The clauses in parentheses will be omitted when $\prec$ and b are clear from the context. It should also be clear from the context that G is of sort $\mathscr{P}\big(X\times\mathscr{P}(X\times Y)\times Y\big)$. The idea is that, if we are defining f: $[b] \to Y$ by recursion over $\prec$, then, at the stage where $f(x)$ is to be defined, the available data consists of x and the previous values of f, i.e. $f \upharpoonright b \cap \prec x$. These previous values constitute data under x, and the manner in which $f(x)$ is to be computed from x and $f \upharpoonright x$ is codified by G. Thus, we define

$$f \text{ satisfies } G \text{ (w.r.t. } \prec \text{ on } b\text{)}: \equiv f \text{ is a function} \wedge \text{domain}(f) = b \wedge \\ \wedge \forall x \forall p\big(\langle x, p\rangle \in f \to \big\langle\langle x, f\upharpoonright b \cap \prec x\rangle, p\big\rangle \in G\big).$$

THEOREM. $\models \prec$ *is well-founded on* $b \wedge G$ *is a recursion condition* $\to \exists! f$ *satisfies* G.

Proof. Replacing $\prec$ with $x \mapsto b \cap \prec x$, we may assume for simplicity that $\models \forall x(\prec x \subseteq b)$. Define

$$g \text{ is good (for } G, \prec, b\text{)}: \equiv g \text{ is a function} \wedge \text{domain}(g) \subseteq b \wedge \\ \wedge \forall x \forall p\big[\langle x, p\rangle \in g \to \big\langle\langle x, g \upharpoonright \prec x\rangle, p\big\rangle \in G\big].$$

Working in the internal logic, we are given that $\prec$ is well-founded on b and that G is a recursion condition, and we must prove $\exists! f[f \text{ is good} \wedge \text{domain}(f) = b]$.

LEMMA. g *is good* $\wedge\, g'$ *is good* $\wedge \langle x, p\rangle \in g \wedge \langle x, p'\rangle \in g' \to p = p'$.

Proof. We use induction on x, with respect to $\prec$. Assume all the hypotheses of the lemma and assume that the lemma holds with any $y \prec x$ in place of x. From the hypothesis $\langle x, p\rangle \in g$ we obtain, as g is good, $\big\langle\langle x, g\restriction\prec x\rangle, p\big\rangle \in G$. Looking at the clause about domain(G) in the definition of recursion condition, we see that $g\restriction\prec x$ is data under x, which implies $\prec x \subseteq \mathrm{domain}(g)$. Likewise, $\prec x \subseteq \mathrm{domain}(g')$. Now the induction hypothesis yields (via extensionality) $g\restriction\prec x = g'\restriction\prec x$. Finally, the assumption that G is a function gives $p = p'$, as desired. ■

In view of extensionality, the lemma immediately implies the uniqueness of the desired f; it remains therefore only to prove existence. The lemma also implies that the union g of any family Γ of good functions is a function (i.e. $\forall h \in \Gamma$ (h is good) $\to (\bigcup \Gamma)$ is a function, where $\Gamma \,\varepsilon\, \mathscr{P}\mathscr{P}(X\times Y)$). Also, by the argument in the proof of the lemma, we have that, if h is good and $x \in \mathrm{domain}(h)$, then $\prec x \subseteq \mathrm{domain}(h)$, so, if Γ is a family of good functions, $h \in \Gamma$, and $x \in \mathrm{domain}(h)$, then $(\bigcup \Gamma)\restriction\prec x = h\restriction\prec x$. It then follows easily that $\bigcup \Gamma$ is good.

Let $\Gamma :\equiv \{h | h$ is good$\}$ and $f :\equiv \bigcup \Gamma$. By the above, f is good, and it remains to prove only that $b \subseteq \mathrm{domain}(f)$, i.e., $\forall x \in b\, \exists p \langle x, p\rangle \in f$. We prove this by induction on x. So suppose $x \in b$ and, for all $y \in \prec x$, $\exists q \langle y, q\rangle \in f$. Of course q is unique as f is a function. So $f\restriction\prec x$ is data under x. Since G is a recursion condition, there is p such that $\big\langle\langle x, f\restriction\prec x\rangle, p\big\rangle \in G$. Then clearly $f \cup \{\langle x, p\rangle\}$ is good, hence $\in \Gamma$, hence $\subseteq f$. So $\langle x, p\rangle \in f$. ■

It should be mentioned that, instead of building f up out of smaller good function, one can also obtain f from above, as the intersection

$$\bigcap \{r \,\varepsilon\, \mathscr{P}(X\times Y) | \forall x \in b \forall p \forall h [h \subseteq r \wedge \big\langle\langle x, h\rangle, p\big\rangle \in G \to \langle x, p\rangle \in r]\}.$$

The proof that this gives an f as required is similar to the classical one [14], though some care is needed to avoid using the law of the excluded middle. This approach was used by Mikkelsen [25] in proving a topos version of the recursion theorem; his proof used a "translation" into the standard language of category theory (commutative diagrams, etc.) rather than the simpler translation into the internal language.

Appendix to Section 4. More on well-foundedness. This appendix contains results on well-foundedness which seem to be worth mentioning although they are not needed for the applications to free algebras and coequalizers.

We begin by explaining why the classical "minimal element" definition of well-foundedness is inappropriate when the logic is intuitionistic rather than classical. (It is, of course, equivalent to the definition we have used, in the presence of classical logic.) We shall show that, to the extent that this minimal element definition is satisfied in a non-trivial way, the logic *must* be classical. The argument is essentially the same as one given by Myhill [28]. Retaining our conventions about the sorts of $\prec$, a, b, etc., we define

$$\prec \text{ satisfies the minimality condition on } b :\equiv \forall a [a \subseteq b \wedge \exists x\, x \in a \to \exists x (x \in a \wedge \neg \exists y (y \in a \wedge y \prec x))].$$

Proposition. $\models [\prec$ *satisfies the minimality condition on* $b \wedge x \in b \wedge y \in b \wedge x \prec y] \to (\forall z \,\varepsilon\, \Omega)[z \vee \neg z]$.

Proof. Working in the internal logic, let $\prec$, b, x, y be as in the hypothesis, and let $z \,\varepsilon\, \Omega$. Apply the minimality condition with

$$a = \{v \,\varepsilon\, X |\ v = y \vee [(v = x) \wedge z]\}.$$

Then $a \subseteq b$ because $x \in b \wedge y \in b$. Further, $\exists x\, x \in a$ because $y \in a$. So, by the minimality condition, let $u \,\varepsilon\, X$ be such that $u \in a \wedge \neg \exists v (v \in a \wedge v \prec u)$. By definition of a, $u \in a$ leads to two cases.

Case 1. $u = y$. Then $\neg \exists v (v \in a \wedge v \prec u)$ and $x \prec y$ yield $\neg (x \in a)$. But, by definition of a, $z \to (x \in a)$, so we have $\neg z$.

Case 2. $(u = x) \wedge z$. Therefore z.

Thus, we have $z \vee \neg z$. ■

Our goal in the remainder of this appendix is to show that well-foundedness is exactly the right hypothesis for the justification of recursive definitions. We shall show that the uniqueness of functions satisfying recursion conditions implies the well-foundedness of $\prec$. This is the internal form of a result of Mikkelsen [25]. (In the appendix to the next section, we shall obtain the same conclusion from the hypothesis of existence of such functions. Thus, well-foundedness is necessary for both the existence part and the uniqueness part of the theorem of Section 4.) For simplicity, we shall consider well-foundedness and recursion with respect to $\prec$ on all$_X$; it would be straightforward to do the same with an arbitrary $b \,\varepsilon\, \mathscr{P}(X)$. We shall also work within the internal logic in most of the discussion, although we state the main results in the meta-theory (i.e. with $\models$ in front).

We begin by showing that the classical notion of the well-founded part of a relation still makes sense in the present intuitionistic context.

Definition. $x \in$ the well-founded part of $\prec :\equiv \exists a [x \in a \wedge \forall y \in a (\prec y \subseteq a) \wedge \wedge \prec$ is well-founded on $a]$.

Proposition. (i) $\models \prec$ *is well-founded on the well-founded part of* $\prec$.

(ii) $\models x \in$ *the well-founded part of* $\prec \leftrightarrow (\prec x) \subseteq$ *the well-founded part of* $\prec$.

Proof. (i) Write w for the well-founded part of $\prec$. Assume c is $\prec$ inductive on w, and let $x \in w$; we must prove $x \in c$. As $x \in w$, fix a such that $x \in a$ and $(\forall y \in a) \prec y \subseteq a$ and $\prec$ is well-founded on a. It will suffice to show that c is $\prec$ inductive on a, for then well-foundedness on a gives $x \in c$ as desired. So fix $y \in a$ and assume $\forall z \in a (z \prec y \to z \in c)$; we must show $y \in c$. As c is $\prec$ inductive on w, it suffices to show $\forall z \in w (z \prec y \to z \in c)$. But if $z \prec y$, then $z \in a$ by choice of a, so the hypothesis $\forall z \in a (z \prec y \to z \in c)$ applies and yields $z \in c$ as desired.

(ii) The implication from left to right is an immediate consequence of the definition of the well-founded part w of $\prec$. For the converse, we assume $(\prec x) \subseteq w$, and show that $w \cup \{x\}$ is an a of the sort required in the definition of $x \in w$. The only non-trivial verification is the well-foundedness of $\prec$ on $w \cup \{x\}$. So suppose c is $\prec$ inductive on $w \cup \{x\}$. Then, in virtue of the left-to-right part of (ii), c is also

$\prec$ inductive on w. By (i), $w \subseteq c$. In particular, $\prec x \subseteq c$ since $\prec x \subseteq w$. By inductiveness of c, we infer $x \in c$. So $w \cup \{x\} \subseteq c$. ■

COROLLARY. $\vDash$ *The well-founded part of* $\prec$ *is the largest* $w \,\varepsilon\, \mathscr{P}(X)$ *such that* $(\forall x \in w) \prec x \subseteq w$ *and* $\prec$ *is well-founded on* w. ■

COROLLARY. $\prec$ *is well-founded on* $\mathrm{all}_X \equiv$ *the well-founded part of* $\prec \; = \mathrm{all}_X$. ■

THEOREM. $\vDash \prec$ *is well-founded on* $\mathrm{all}_X \leftrightarrow \forall G \,\varepsilon\, \mathscr{P}(X \times \mathscr{P}(X \times \Omega) \times \Omega) \forall f, f' \,\varepsilon\, \mathscr{P}(X \times \Omega)$ [G *is a recursion condition* $\wedge f$ *satisfies* $G \wedge f'$ *satisfies* $G \to f = f'$].

Proof. The left-to-right implication is just the uniqueness part of the theorem of Section 4. For the converse, apply the assumption with G, f, f' defined as follows.

$$\langle\langle x, h\rangle, p\rangle \in G \equiv h \text{ is data under } x \wedge p =_\Omega (\forall y \in \prec x)\langle y, \text{true}\rangle \in h,$$

$$\langle x, p\rangle \in f \equiv p,$$

$$\langle x, p\rangle \in f' \equiv p =_\Omega (x \in \text{the well-founded part of } \prec).$$

It is easy to verify that G is a recursion condition satisfied by both f and f' (use (ii) of the preceding proposition and the fact that $p =_\Omega \text{true} \equiv p$). So $f = f'$, which means, since $\forall x \langle x, \text{true}\rangle \in f$, that $\forall x [\text{true} =_\Omega (x \in \text{the well-founded part of } \prec)]$. So the well-founded part of $\prec \; = \mathrm{all}_X$ which, by the last corollary, is the desired conclusion. ■

5. Trees and words. As we indicated in Section 1, we shall construct the initial algebra for the variety defined by $\tau\colon I \to J$ as an algebra of words, where a word is a function into J from a tree of finite sequences of elements of I. The purpose of this section is to formulate this notion of word in the internal logic and to prove a few properties of words. In the next section we shall define a τ-algebra structure on the object of words and show that we thereby obtain the initial τ-algebra.

Let $\mathscr{E}$ be a topos with a natural numbers object N. We recall that the definition of a natural numbers object requires the existence of morphisms $1 \xrightarrow[0]{} N \xrightarrow[S]{} N$ such that for any $1 \xrightarrow[a]{} X \xrightarrow[f]{} X$ there is a unique $g\colon N \to X$ making the diagram

$$\begin{tikzcd} & N \arrow[r, "S"] \arrow[d, "g"] & N \arrow[d, "g"] \\ 1 \arrow[ur, "0"] \arrow[r, "a"'] & X \arrow[r, "f"'] & X \end{tikzcd}$$

commute. This definition, in the presence of the cartesian closed structure of $\mathscr{E}$ lets us define morphisms in $\mathscr{E}$ corresponding to those functions and predicates on the natural numbers which, in ordinary (set-based) arithmetic, would be defined by primitive recursion. In particular, we have $+, \cdot\colon N \times N \to N$ and the order relation $<\colon N \times N \to \Omega$ (defined by the recursion $\neg(x<0)$, $x < Sy \leftrightarrow (x < y \vee x = y)$). Furthermore, the uniqueness clause in the definition of N permits proofs by ordinary induction in the internal logic;

$$\vDash \forall a \,\varepsilon\, \mathscr{P}(N)[0 \in a \wedge \forall x \in a Sx \in a \to a = \mathrm{all}_N],$$

and

$$\vDash x \mapsto \{y \,\varepsilon\, N | \; Sy = x\} \text{ is well-founded on } \mathrm{all}_N.$$

Induction allows us to prove the basic laws of arithmetic, including the commutative, associative, and distributive laws for $+$ and $\cdot$ and (despite the possibly non-Boolean logic of $\mathscr{E}$)

$$\vDash \forall x, y \,\varepsilon\, N[x<y \vee x = y \vee y<x] \text{ (where } x<y \text{ means } <(x, y))$$

and

$$\vDash \forall x \,\varepsilon\, N[x = 0 \vee \exists y\, x = Sy]$$

which will be important in justifying some proofs by cases later on. For a more detailed treatment of natural number objects, see Chapter 6 of [17].

Our chief use for N is in the following definitions of the notions of a finite sequence from an object I and the length of such a sequence.

$$\mathrm{length}_I\colon \equiv \{\langle x, n\rangle \,\varepsilon\, \mathscr{P}(N \times I) \times N | x \text{ is a function} \wedge \mathrm{domain}(x) = \{p \,\varepsilon\, N | p<n\}\}.$$

Note that $\vDash \mathrm{length}_I$ is a function.

$$\mathrm{Seq}_I\colon \equiv \mathrm{domain}(\mathrm{length}_I).$$

The following discussion takes place within the internal logic. It is phrased in an informal style that approximates ordinary mathematical usage, but it could easily be formalized in the internal logic with the definitions already introduced.

We write ϕ for $\{x \,\varepsilon\, N \times I | \text{false}\}$ and observe $\phi \in \mathrm{Seq}$ and $\mathrm{length}(\phi) = 0$ (i.e. $\langle \phi, 0\rangle \in \mathrm{length}$). For $i \,\varepsilon\, I$, we write $\langle i\rangle$ for $\{\langle 0, i\rangle\}$, so $\langle i\rangle \in \mathrm{Seq}$ and length $(\langle i\rangle) = 1(\colon \equiv S(0))$. For $x, y \in \mathrm{Seq}$, we define the concatenation

$$x^\frown y\colon \equiv x \cup \{\langle \mathrm{length}(x)+p, i\rangle | \langle p, i\rangle \in y\},$$

and, with the help of $\forall q \,\varepsilon\, N[q<a+b \leftrightarrow q<a \vee \exists p \,\varepsilon\, N(p<b \wedge q = a+p)]$, we find that $x^\frown y \in \mathrm{Seq}$ and length $(x^\frown y) = \mathrm{length}(x) + \mathrm{length}(y)$. From associativity of $+$, we obtain associativity of $^\frown$, and clearly ϕ is a two-sided unit for $^\frown$. (By externalization, we obtain a monoid structure $1 \xrightarrow[\phi]{} [\mathrm{Seq}]$, $[\mathrm{Seq}] \times [\mathrm{Seq}] \xrightarrow[\frown]{} [\mathrm{Seq}]$ and a morphism $I \xrightarrow[\langle\,\rangle]{} [\mathrm{Seq}]$. In fact, $[\mathrm{Seq}]$ is the free monoid on I constructed in [23]. An alternate description of $[\mathrm{Seq}]$ is $\Sigma_N \Pi_\nu \Delta_R I$, where $R = [\langle x, y\rangle \,\varepsilon\, N \times N | x<y]$ and where $\nu\colon R \rightarrowtail N \times N \xrightarrow[2\text{nd}]{} N$ is the generic natural number.) Notice also that, for $x \in \mathrm{Seq}$,

$$\text{length } (x) = 0 \to x = \phi,$$

and

$$\text{length } (x) = S_n \to \exists y \in \mathrm{Seq}\, \exists i \,\varepsilon\, I[\mathrm{length}(y) = n \wedge x = y^\frown\langle i\rangle] \wedge$$
$$\wedge\; \exists y \in \mathrm{Seq}\, \exists i \,\varepsilon\, I[\mathrm{length}(y) = n \wedge x = \langle i\rangle^\frown y].$$

The first of these is obvious. For the second, to get $x = y^\frown\langle i\rangle$ take $i = x(n)$ and set $y = x \restriction \{p | p<n\}$; to get $x = \langle i\rangle^\frown y$, take $i = x(0)$ and

$$y = \{\langle p, z\rangle \,\varepsilon\, N \times I | \langle S(p), z\rangle \in x\}.$$

We call $s(\varepsilon\mathscr{P}\mathscr{P}(N\times I))$ a *tree* iff $s\subseteq$Seq and $\forall x\in \mathrm{Seq}\,\forall i\varepsilon I\,[x^\frown\langle i\rangle\in s\to x\in s]$. (This defines a term Tres$\,\varepsilon\mathscr{P}\mathscr{P}\mathscr{P}(N\times I)$.) If s is a tree, it follows by induction on length(y) that $\forall x, y\in \mathrm{Seq}\,[x^\frown y\in s\to x\in s]$. A tree s is *well-founded* iff the relation Extensions defined by

$$x\in \mathrm{Extensions}(y):\equiv \exists i\varepsilon I\, x = y^\frown\langle i\rangle$$

is well-founded on s. (Note that longer sequences are "lower" with respect to Extensions. Also note that no harm results from the lack of transitivity of Extensions.)

Until now, we have used only the domain I of the given morphism $\tau\colon I\to J$ relative to which we wish to construct an initial algebra. At this point, τ and J enter the picture. We call $w\varepsilon\mathscr{P}(\mathscr{P}(N\times I)\times J)$ a *word* (with respect to τ) iff w is a function and domain (w) is a well-founded tree s such that $\phi\in s$ and

$$\forall x\in \mathrm{Seq}\,\forall i\varepsilon I[x^\frown\langle i\rangle\in s\leftrightarrow \langle x, \tau(i)\rangle\in w]\,.$$

The reader should convince himself that this notion of word agrees, when $\mathscr{E}$ is the topos of sets, with the notion presented in Section 1. The term

$$\mathrm{Words}\ \varepsilon\mathscr{P}\mathscr{P}(\mathscr{P}(N\times I)\times J)$$

defined here will yield, by externalization, the underlying object [Words] of the initial τ-algebra.

For any word w, we define the *principal operator* of w to be $w(\phi)$. If $i\varepsilon I$ and $\tau(i)$ is the principal operator of w, the *i-constituent* of w is defined to be

$$\{\langle x, q\rangle|\big\langle\langle i\rangle^\frown x, q\big\rangle\in w\}\,.$$

The *constituent list* of w is the function whose domain is $\{i\varepsilon I|\ \tau(i) = w(\phi)\}$ and whose value at such an i is the i-constituent of w. The range of this constituent list (the set of all constituents of w) will be written Constituent (w).

PROPOSITION 1. *If w is a word and $\tau(i) = w(\phi)$, then the i-constituent of w is also a word.*

Proof. Let $u = \{\langle x, q\rangle|\big\langle\langle i\rangle^\frown x, q\big\rangle\in w\}$ be the i-constituent of w, let $s=$ domain(w), and let $t=$ domain(u). So $t = \{x|\langle i\rangle^\frown x\in s\}$. Clearly u is a function. Since $\langle\phi, \tau(i)\rangle\in w$ by assumption, the fact that w is a word implies $\phi^\frown\langle i\rangle\in s$, so $\langle i\rangle^\frown\phi\in s$ (as ϕ is the unit for $^\frown$) and therefore $\phi\in t$. Also, for any $a\varepsilon I$, $x\in$ Seq,

$$\begin{aligned} x^\frown\langle a\rangle\in t &\leftrightarrow \langle i\rangle^\frown x^\frown\langle a\rangle\in s\\ &\leftrightarrow \big\langle\langle i\rangle^\frown x, \tau(a)\big\rangle\in w\\ &\leftrightarrow \langle x, \tau(a)\rangle\in u\,.\end{aligned}$$

All that remains to be shown is that t is a well-founded tree. That it is a tree is immediate, since s is one (and $^\frown$ is associative). To prove well-foundedness, suppose $e\subseteq t$ is Extensions inductive on t. Let

$$e' = \{\phi\}\cup\{\langle a\rangle^\frown x|\ x\in \mathrm{Seq}\wedge(a = i\to x\in e)\}\,.$$

(Intuitively, it seems natural to take $\{\phi\}\cup\{\langle i\rangle^\frown x|\ x\in e\}\cup\{\langle a\rangle^\frown x|\ a\neq i\wedge x\in \mathrm{Seq}\}$, but, in the absence of the law of the excluded middle, this alternate definition seems not to work in the argument that follows. Of course the two formulations are classically equivalent.) We claim that e' is Extensions inductive on s.

To prove the claim, suppose $y\in s$ and, for every $z\in s\cap \mathrm{Extensions}(y)$, $z\in e'$. By definition of Extensions, this means that, for all $b\varepsilon I$, if $y^\frown\langle b\rangle\in s$ then $y^\frown\langle b\rangle\in e'$. We must show $y\in e'$, which is clear if length$(y) = 0$ (as then $y = \phi\in e'$), so we may assume (by one of the properties of natural numbers mentioned earlier) that length$(y) = S(n)$ for some n, so $y = \langle a\rangle^\frown x$ for some $a\varepsilon I$ and some $x\in$ Seq. By definition of e', what we must prove is $a = i\to x\in e$, so let us assume $a = i$. As $y = \langle i\rangle^\frown x\in s$, we have $x\in t$. Also, for any $b\varepsilon I$ we have the chain of implications

$$\begin{aligned} x^\frown\langle b\rangle\in t &\to y^\frown\langle b\rangle = \langle i\rangle^\frown x^\frown\langle b\rangle\in s\\ &\to y^\frown\langle b\rangle = \langle i\rangle^\frown x^\frown\langle b\rangle\in e'\\ &\to x^\frown\langle b\rangle\in e\,.\end{aligned}$$

But e is Extensions inductive on t, so $x\in e$ as desired. This completes the proof of the claim.

Since Extensions is well-founded on s (because w is a word), the claim implies $e'\supseteq s$. Thus, for any $x\in$ Seq,

$$\begin{aligned} x\in t &\leftrightarrow \langle i\rangle^\frown x\in s\\ &\to \langle i\rangle^\frown x\in e'\\ &\leftrightarrow x\in e\,,\end{aligned}$$

so $t\subseteq e$. This completes the proof that t is well-founded. ■

COROLLARY. *If w is a word and $z\in$ domain(w) then*

$$w*z:\equiv\{\langle x, q\rangle|\ \langle z^\frown x, q\rangle\in w\}$$

is also a word.

Proof. We use induction on length (z). If length $(z) = 0$ then $z^\frown x = \phi^\frown x = x$, so $w*z = w$, a word. Now suppose length $(z) = Sn$. Then $z = y^\frown\langle i\rangle$ for some $y\in$ domain (w) (because domain (w) is a tree) and some $i\varepsilon I$. By induction hypothesis, $w*y$ is a word. Also, because w is a word, it follows from $y^\frown\langle i\rangle = z\in$ domain (w) that $\langle y, \tau(i)\rangle\in w$, so $\langle\phi, \tau(i)\rangle\in w*y$. Now the proposition tells us that the i-constituent of $w*y$ is a word, and it is easy to check that this constituent is precisely $w*z$. ■

Recall that, for any word w,

Constituent $(w)\equiv\{u|(\exists i\varepsilon I)[\langle\phi, \tau(i)\rangle\in w\wedge u$ is the i-constituent of $w]\}$.

PROPOSITION 2. *Constituent is well-founded on Words.*

Proof. Let $a\subseteq$Words be Constituent inductive on Words, and let w be any word; we must prove $w\in a$. Let $e = \{z\in \mathrm{domain}\,(w)|\ w*z\in a\}$; we must prove $\phi\in e$. Since $\phi\in$ domain (w) and Extensions is well-founded on domain (w), it suffices to prove that e is Extensions inductive on domain (w). So suppose

$z \in \operatorname{domain}(w)$ and $\forall i \varepsilon I[z^\frown\langle i\rangle \in \operatorname{domain}(w) \to z^\frown\langle i\rangle \in e]$; we must prove $z \in e$, i.e., $w * z \in a$. For this it suffices, since a is Constituent inductive on Words and $w * z$ is a word, to show that every constituent u of $w * z$ is in a. So suppose $\langle \phi, \tau(i)\rangle \in w * z$ and u is the i-constituent of $w * z$. The condition on $\tau(i)$ means $\langle z, \tau(i)\rangle \in w$. As w is a word, it follows that $z^\frown\langle i\rangle \in \operatorname{domain}(w)$. By our supposition about z, we infer $z^\frown\langle i\rangle \in e$, so $w * (z^\frown\langle i\rangle) \in a$. But, as remarked in the proof of the last corollary, $w * (z^\frown\langle i\rangle) = u$. ∎

PROPOSITION 3. *For every $j \varepsilon J$ and every function λ with domain $\{i \varepsilon I \mid \tau(i) = j\}$ and range $\subseteq$ Words, there is a unique word w whose principal operator is j and whose constituent list is λ.*

Proof. It is fairly clear that the only w that could satisfy the conditions is

$$w = \{\langle \phi, j\rangle\} \cup \{\langle \langle i\rangle^\frown x, q\rangle \mid \tau(i) = j \wedge \langle x, q\rangle \in \lambda(i)\}.$$

We need only check that this w is a word. Its domain is

$$s = \{\phi\} \cup \{\langle i\rangle^\frown x \mid \tau(i) = j \wedge x \in \operatorname{domain}(\lambda(i))\}.$$

To show that s is a tree, suppose $y^\frown\langle a\rangle \in s$. As $y^\frown\langle a\rangle$ cannot equal ϕ, it must equal $\langle i\rangle^\frown x$ for some i with $\tau(i) = j$ and some $x \in \operatorname{domain}(\lambda(i))$. If length $(y) = 0$, then $y = \phi \in s$, so we may assume length $(y) = Sn$, $y = \langle i\rangle^\frown z$, and $x = z^\frown\langle a\rangle$. As $\lambda(i)$ is a word, we deduce from $x \in \operatorname{domain}(\lambda(i))$ that $z \in \operatorname{domain}(\lambda(i))$, so $y \in s$. This completes the proof that s is a tree.

To show that s is well-founded, let $e \subseteq d$ be Extensions inductive on s. Temporarily fix an arbitrary i with $\tau(i) = j$, and set $e'_i = \{x \mid \langle i\rangle^\frown x \in e\}$. We shall show that e'_i is Extensions inductive on domain $(\lambda(i))$. Suppose that $x \in \operatorname{domain}(\lambda(i))$ and, for each a such that $x^\frown\langle a\rangle \in \operatorname{domain}(\lambda(i))$, we have $x^\frown\langle a\rangle \in e'_i$. In other words, $\langle i\rangle^\frown x \in s$ and, for each a such that $\langle i\rangle^\frown x^\frown\langle a\rangle \in s$, we have $\langle i\rangle^\frown x^\frown\langle a\rangle \in e$. By the assumed inductiveness of e, we have $\langle i\rangle^\frown x \in e$, so $x \in e'_i$, as desired. Since $\lambda(i)$ is a word, its domain is well-founded, so domain $(\lambda(i)) \subseteq e'_i$. Since this holds for every such i, we have the implication

$$\tau(i) = j \wedge x \in \operatorname{domain}(\lambda(i)) \to \langle i\rangle^\frown x \in e.$$

This shows that every element of the second set in the union defining s belongs to e. In particular, $\forall i \varepsilon I[\langle i\rangle \in s \to \langle i\rangle \in e]$, which implies $\phi \in e$ because e is inductive. This completes the proof that $s \subseteq e$, so s is well-founded.

Finally, we must verify that

$$\forall x \in \operatorname{Seq} \forall i \varepsilon I[x^\frown\langle i\rangle \in \operatorname{domain}(w) \leftrightarrow \langle x, \tau(i)\rangle \in w].$$

Suppose first that length $(x) = 0$, so $x = \phi$ and $x^\frown\langle i\rangle = \langle i\rangle$. Then,

$$\begin{aligned} x^\frown\langle i\rangle \in \operatorname{domain}(w) &\leftrightarrow \langle i\rangle^\frown\phi \in \operatorname{domain}(w) \\ &\leftrightarrow \tau(i) = j \wedge \phi \in \operatorname{domain}(\lambda(i)) \quad \text{(by definition of } w) \\ &\leftrightarrow \tau(i) = j \quad \text{(as } \lambda(i) \text{ is a word)} \\ &\leftrightarrow \langle \phi, \tau(i)\rangle \in w \quad \text{(as } j \text{ is the unique value of the function } w \text{ at } \phi) \\ &\leftrightarrow \langle x, \tau(i)\rangle \in w, \end{aligned}$$

so we have desired conclusion in this case. It remains to consider the case length $(x) = S(n)$. Then $x = \langle a\rangle^\frown y$ for some $a \varepsilon I$ and some $y \in \operatorname{Seq}$. We have

$$\begin{aligned} x^\frown\langle i\rangle \in \operatorname{domain}(w) &\leftrightarrow \tau(a) = j \wedge y^\frown\langle i\rangle \in \operatorname{domain}(\lambda(a)) \quad \text{(by definition of } w) \\ &\leftrightarrow \tau(a) = j \wedge \langle y, \tau(i)\rangle \in \lambda(a) \quad \text{(as } \lambda(a) \text{ is a word)} \\ &\leftrightarrow \big\langle \langle a\rangle^\frown y, \tau(i)\big\rangle \in w \quad \text{(by definition of } w) \\ &\leftrightarrow \langle x, \tau(i)\rangle \in w, \end{aligned}$$

which completes the proof that w is a word. ∎

Appendix to Section 5. Still more on well-foundedness.

There are results about well-founded trees analogous to the results of Section 5 about words. The simplest way to obtain these is to delete everything in Section 5 that refers to the values of the (functions that are) words, keeping only their domains. In the resulting arguments, words are replaced with rooted well-founded trees, where a tree s is *rooted* if $\phi \in s$. The i-constituent of such a tree s is $\{x \mid \langle i\rangle^\frown x \in s\}$ if $\langle i\rangle \in s$ and undefined otherwise. (The requirement $\langle i\rangle \in s$ corresponds to $\tau(i) = w(\phi)$ in the earlier discussion.) The analogs of the propositions of Section 5 assert that all the constituents of a rooted well-founded tree are also rooted well-founded trees, that Constituent is a well-founded relation on rooted well-founded trees, and that, every function λ with domain $\varepsilon\mathscr{P}(I)$ and with rooted well-founded trees as values is the constituent list of a unique rooted well-founded tree. These results can also be obtained as consequences rather than analogs of the results in Section 5 by observing that rooted well-founded trees are in canonical (constituent-preserving) one to one correspondence with words relative to

$$\tau\colon [\langle x, y\rangle \varepsilon I \times \mathscr{P}(I) \mid x \in y] \rightarrowtail I \times \mathscr{P}(I) \underset{2^{\text{nd}}}{\to} \mathscr{P}(I).$$

We leave the details of this approach to the reader.

Rooted well-founded trees can be used as a substitute for the ordinal numbers of classical set theory. Although direct analogs of ordinal numbers can be defined in topoi, they do not seem to fulfill their classical roles in the absence of classical logic and the axiom of replacement. To illustrate how trees can be used to fulfill these roles, we prove the following theorem, which complements the results of the appendix to Section 4.

THEOREM. *Let $\prec \in \mathscr{P}(X)^X$ and let* R∧T *be the object of rooted well-founded trees over X. If every recursion condition G from X to* RWT *is satisfied by some f, then $\prec$ is well-founded.*

Proof. Define

$$G\colon \equiv \{\langle \langle x, h\rangle, s\rangle \mid h \text{ is data under } x \wedge s \text{ is the (unique) rooted well-founded tree with constituent list } h\}.$$

Then G is a recursion condition; suppose f is a function satisfying it. Thus, if $y \prec x$ then $f(y)$ is a constituent (specifically the y-constituent) of $f(x)$. Since the constituent relation is well-founded, we are reduced to proving the following lemma.

LEMMA. *Suppose that f is a function from X to X', that $\prec$ and $\prec'$ are relations on X, X' respectively, that $x\prec y\to f(x)\prec' f(y)$, and that $\prec'$ is well-founded. Then $\prec$ is well-founded.*

Proof. Let e be $\prec$ inductive and let

$$e' :\equiv \{z\,\varepsilon\, X' |\ \forall x\,\varepsilon\, X[f(x)=z\to x\in e]\}.$$

It suffices to show e' is $\prec'$ inductive, for then $\forall z\,\varepsilon\, X'\ z\in e'$ and therefore $\forall x\,\varepsilon\, X\ x\in e'$ as desired. So let $z\,\varepsilon\, X'$ and assume $\forall u\,\varepsilon\, X'[u\prec' z\to u\in e']$. We wish to show $z\in e'$, so we consider an arbitrary $x\,\varepsilon\, X$ with $f(x)=z$, and we wish to show $x\in e$. Since e is $\prec$ inductive, it suffices to show that $y\in e$ for every $y\,\varepsilon\, X$ with $y\prec x$. But $y\prec x$ implies $f(y)\prec f(x)=z$, so $f(y)\in e'$ by assumption on z, so $y\in e$. This proves the lemma and thus also the theorem. ∎

We close this appendix with the observation that well-foundedness of rooted trees, though apparently only a special case of well-foundedness, is in fact "equivalent" to the general case. Specifically, we can define for any $\prec\,\varepsilon\,\mathscr{P}(X)^X$, a rooted tree of decreasing sequences.

$$\mathrm{Decr}(\prec) :\equiv \{v|\ v\in \mathrm{Seq}\wedge(\forall n\,\varepsilon\, N)[Sn\in \mathrm{domain}(v)\to v(Sn)\prec v(n)]\}.$$

PROPOSITION. *$\prec$ is well-founded if and only if* $\mathrm{Decr}(\prec)$ *is well-founded.*

Proof sketch. We only indicate how to transform a $\prec$ inductive set e into an inductive subset e' of $\mathrm{Decr}(\prec)$ and vice versa; the detailed verifications are left to the reader. Given e, set

$$e' = \{\phi\}\cup\{v|\ \exists n\,\varepsilon\, N[\mathrm{length}(v)=S(n)\wedge v(n)\in e]\}.$$

Conversely, given e', set

$$e=\{x|\ \forall v\in \mathrm{Seq}\,\forall n\,\varepsilon\, N[(\mathrm{length}(v)=S(n)\wedge v(n)=x)\to v\in e']\}. \ ∎$$

6. The initial algebra. We continue to work in the internal logic. By virtue of Proposition 3 of Section 5,

$$\mu :\equiv \{\langle\langle j,\lambda\rangle, w\rangle|\ w \text{ is a word with principal connective } j \text{ and constituent list } \lambda\}$$

is a function with domain

$$\{\langle j,\lambda\rangle|\ \lambda \text{ is a function with domain } \{i\,\varepsilon\, I|\ \tau(i)=j\} \text{ and range} \subseteq \text{Words}\}.$$

Let $F=[\text{Words}]$. The externalization of the domain of μ is $[\mathrm{domain}(\mu)]=\Sigma_J\Pi_\tau\Delta_I F$, by the theorem in Section 3. Externalizing μ we get a morphism, which we still call μ, from $\Sigma_J\Pi_\tau\Delta_I F$ to F, in other words a τ-algebra structure on F. The purpose of this section is to prove that (F,μ) is the initial τ-algebra, thereby completing the construction of free algebras begun in Section 1.

Let (A,ν) be any τ-algebra. Recall from Section 3 that homomorphisms from (F,μ) to (A,ν) are the externalizations of closed terms $\alpha\,\varepsilon\,\mathscr{P}(F\times A)$ such that it is valid that

(1) α is a function $\wedge$ domain $(\alpha)=\mathrm{all}_F\wedge\forall j\,\varepsilon\, J\,\forall m\,\varepsilon\,\mathscr{P}(I\times F)\,\forall w\,\varepsilon\, F$
$$[\langle\langle j,m\rangle, w\rangle\in\mu\to\langle\langle j,\alpha\circ m\rangle,\alpha(w)\rangle\in \mathrm{Graph}(\nu)].$$

And two homomorphisms are equal if and only if the corresponding terms satisfy $\vDash\alpha=\alpha'$. Using the definitions of F and μ, we see (in the internal logic) that α's satisfying (1) correspond canonically to β's$\big(\varepsilon\mathscr{P}(\mathscr{P}(\mathscr{P}(N\times I)\times J)\times A)\big)$ such that

(2) β is a function $\wedge$ domain $(\beta)=$ Words $\wedge\,\forall\langle j,m\rangle\in$ domain $(\mu)\ \forall w\in$ Words
$$[\langle\langle j,m\rangle, w\rangle\in\mu\to\langle\langle j,\beta\circ m\rangle,\beta(w)\rangle\in\mathrm{Graph}(\nu)].$$

Since every word has a (unique) principal operator and constituent list, (2) is equivalent to

(3) $\forall w\in$ Words $[\beta(w)=\nu(\langle$principal connective of w,
$$(\beta\restriction \mathrm{Constituent}(w))\circ(\text{constituent list of } w)\rangle)].$$

But (3) says that β satisfies a certain recursion condition with respect to Constituent, which is well-founded on Words by Proposition 2 of Section 5. Therefore, there is a unique such β. This completes the proof that (F,μ) is the initial τ-algebra.

It is perhaps worth pointing out that the close connection between initial algebras and well-foundedness is not merely an artifact of our construction but is inherent in the situation. Indeed, by combining our construction of initial algebras with the remarks in the appendix to Section 5 about words with respect to

$$\tau:\ [\{\langle x,y\rangle\,\varepsilon\, I\times\mathscr{P}(I)|\ x\in y\}]\rightarrowtail I\times\mathscr{P}(I)\xrightarrow[2^{\mathrm{nd}}]{}\mathscr{P}(I),$$

we see that the object of well-founded rooted trees on I is (canonically isomorphic to) the initial algebra for this τ. We leave to the reader the task of describing the structure of trees in terms of the algebra structure. Note, incidentally, that τ could be described as having exactly one q-ary operation symbol for every $q\subseteq I$.

7. Coequalizers. In this section, we show that, in a Boolean topos with natural numbers object, coequalizers of τ-algebras exist, and can be obtained by a word construction similar to the construction of free algebras. Rosebrugh [34] has shown the existence of coequalizers of τ-algebras in any topos with natural numbers object that is Grothendieck over a topos satisfying the axiom of choice. In an appendix to this section, we shall show that Rosebrugh's result does not subsume ours; it is obvious that ours does not subsume his, since there are many non-Boolean topoi Grothendieck over the topos of sets. As far as I know, it may be the case that coequalizers of τ-algebras exist in all topoi with natural numbers objects.

As in ordinary set-based universal algebra, it is convenient to reduce the construction of coequalizers to a special case, the construction of quotients of congruence relations. Let $\tau: I\to J$, and let (A,μ) be a τ-algebra. Then $A\times A$ has a τ-algebra structure μ^2 which can be described in either of the following ways. If $\mu:\Sigma_J\Pi_\tau\Delta_I A\to A$ corresponds, under the adjunction $\Sigma_J\dashv\Delta_J$, to $\tilde\mu:\Pi_\tau\Delta_I A\to\Delta_J A$, then μ^2 corresponds, under the same adjunction, to

$$\Pi_\tau\Delta_I(A\times A)\cong(\Pi_\tau\Delta_I A)\times(\Pi_\tau\Delta_I A)\xrightarrow[\tilde\mu\times\tilde\mu]{}(\Delta_J A)\times(\Delta_J A)\cong\Delta_J(A\times A),$$

where the isomorphisms are due to the existence of left adjoints for Π_τ, Δ_I, and Δ_J. Alternately, $\models \mu^2(\langle j, \lambda\rangle) = \langle \mu(\langle j, 1^{\text{st}} \circ \lambda\rangle), \mu(\langle j, 2^{\text{nd}} \circ \lambda\rangle)\rangle$.

We define, in the internal logic,

$$\text{Equivalence}\,(r \,\varepsilon\, \mathscr{P}(A\times A)) :\equiv \forall x \,\varepsilon\, A\,[r(x,x)] \wedge \forall x, y \,\varepsilon\, A\,[r(x,y) \to r(y,x)] \wedge \\ \wedge \forall x, y, z \,\varepsilon\, A\,[r(x,y)\wedge r(y,z) \to r(x,z)]\,,$$

and

$$\text{Congruence}\,(r) :\equiv \text{Equivalence}\,(r) \wedge \\ \wedge \forall j \,\varepsilon\, J\, \forall\lambda\forall\lambda'\big[\forall i \,\varepsilon\, I\big(\tau(i) = j \to r(\lambda(i), \lambda'(i))\big) \to r\big(\mu(\langle j, \lambda\rangle), \mu(\langle j, \lambda'\rangle)\big)\big]\,.$$

We call a subobject $R \rightarrowtail A\times A$ an equivalence (resp. congruence) relation on A iff the adjoint $r\colon 1 \to \mathscr{P}(A\times A)$ of its classifying map $A\times A \to \Omega$ satisfies $\models$ Equivalence (r) (resp. $\models$ Congruence (r)). Note that a congruence is simply an equivalence that is closed under μ^2, i.e. defines a subalgebra of $(A\times A, \mu^2)$; we often write μ^2 for the restriction of μ^2 to $\Sigma_J \Pi_\tau \Delta_I R$. A congruence R is equipped with two canonical homomorphisms $R \rightarrowtail A\times A \overset{1^{\text{st}}}{\underset{2^{\text{nd}}}{\rightrightarrows}} A$, whose τ-algebra coequalizer, if it exists, is called the quotient of A by R; we use the notation $A/\!/R$ (or, more precisely, $(A, \mu)/\!/R$) for this quotient, reserving the more common notation A/R for the "set theoretic" quotient, i.e. the coequalizer of $R \rightrightarrows A$ in $\mathscr{E}$ rather than in the category of τ-algebras.

It is easy to verify, in the internal logic, that the intersection of any family of congruences is again a congruence and that the kernel of any homomorphism $\alpha\colon (A, \mu) \to (B, \nu)$, defined by

$$\text{kernel}\,(\alpha) \equiv \{\langle x, y\rangle \,\varepsilon\, A\times A \mid \alpha(x) = \alpha(y)\}\,,$$

is a congruence. Now let $(X, \varrho) \overset{f}{\underset{g}{\rightrightarrows}} (A, \mu)$ be a pair of τ-algebra homomorphisms. A homomorphism $\alpha\colon (A, \mu) \to (B, \nu)$ coequalizes these if and only if its kernel includes the image of $\langle f, g\rangle$ in $A\times A$, i.e. $\{z \,\varepsilon\, A\times A \mid (\exists x \,\varepsilon\, X)\, z = \langle f(x), g(x)\rangle\}$. But this is, by the preceding remarks, the same as saying that the kernel of α includes the intersection $r \,\varepsilon\, \mathscr{P}(A\times A)$ of all congruences that include this image, i.e. that α coequalizes the canonical pair of morphisms $[r] \rightrightarrows A$. Thus, the coequalizer of f and g is the same as the quotient $A/\!/[r]$, in the sense that if either exists then so does the other and they are the same. Therefore, in constructing coequalizers of τ-algebras, we shall confine our attention to constructing quotients by congruence relations.

It was pointed out to me by Paré (and it is hinted at in [18, 30, 34]) that the usual construction of a quotient algebra $A/\!/R$, by transferring the algebra structure from A to A/R, fails in general. To define the value of the quotient algebra structure $\bar{\mu}$ at $\langle j, \lambda\rangle$, where $\lambda\colon \tau^{-1}\{j\} \to A/R$, one takes an arbitrary $\lambda'\colon \tau^{-1}\{j\} \to A$ whose composite with the projection $\pi\colon A \to A/R$ is λ, and one sets $\bar{\mu}(\langle j, \lambda\rangle) = \pi\mu(\langle j, \lambda'\rangle)$. It does not matter which λ' one takes, since R is a congruence, but one must have at least one such λ' available, which involves the axiom of choice. The following proposition shows that there is no way to circumvent this use of the axiom of choice.

THEOREM. *Suppose $\mathscr{E}$ is a topos such that, for every $\tau\colon I \to J$ in $\mathscr{E}$, every τ-algebra (A, μ), and every congruence relation $R \rightarrowtail A\times A$, the quotient τ-algebra $A/\!/R$ exists and the canonical homomorphism $\theta\colon A \to A/\!/R$ is an epimorphism of objects in $\mathscr{E}$. Then $\mathscr{E}$ satisfies the internal axiom of choice.*

Remarks. It will be convenient to take the internal axiom of choice in the form that says, for every epimorphism $p\colon Y \to X$,

$$\models \exists f \,\varepsilon\, \mathscr{P}(X\times Y)\,\big[f \text{ is a function} \wedge \text{domain}\,(f) = \text{all}_X \wedge (\forall x \,\varepsilon\, X)\, p(f(x)) = x\big]\,.$$

Other internal formulations of the axiom of choice are equivalent to this one (as are some external formulations, such as the axiom (IC) of [17], p. 143).

Our proof of the theorem will in fact use a weaker hypothesis than is stated. Instead of assuming that $A/\!/R$ always exists and θ is always epi, we need only know that θ is epi whenever $A/\!/R$ happens to exist.

Proof of theorem. Let $p\colon Y \to X$ be an epimorphism. Let q_1 and q_2 be the two projections of $Y \underset{X}{\times} Y$ to Y; then p is their coequalizer. For any $\tau\colon I \to J$, the diagram of free τ-algebras

$$F(Y \underset{X}{\times} Y) \overset{F(q_1)}{\underset{F(q_2)}{\rightrightarrows}} F(Y) \xrightarrow[F(p)]{} F(X)$$

is also a coequalizer since F is a left adjoint. By assumption, $F(p)$ is epi. It follows easily from our construction of free algebras that $F(p)$ is the map that composes any word in $F(Y)$ (a function into $J+Y$) with $\mathrm{id}_J + p\colon J+Y \to J+X$ to obtain a word in $F(X)$. (Just check that this composition defines a homomorphism of τ-algebras mapping the free generators Y of $F(Y)$ correctly.)

Let us specialize to the case where τ is the unique map $X \to 1$. (We consider algebras with a single X-ary operation.) Set

$$w :\equiv \{\langle \phi, *\rangle\} \cup \{\langle\langle x\rangle, x\rangle \mid x \,\varepsilon\, X\}\,;$$

then $\models w$ is a word with respect to $X \to 1 \rightarrowtail 1+X$. Since $F(p)$ is epi, we have

$$\models \exists u\,[u \text{ is a word with respect to } X \to 1 \rightarrowtail 1+Y \wedge (\mathrm{id}_1 + p)\circ u = w]\,.$$

But then, working in the internal logic, we have, for any such u, domain (u) $=$ domain $(w) = \{\phi\} \cup \{\langle x\rangle \mid x \,\varepsilon\, X\}$, and $(\forall x \,\varepsilon\, X)\, p(u(\langle x\rangle)) = w(\langle x\rangle) = x$, so $\{\langle x, u(\langle x\rangle)\rangle \mid x \,\varepsilon\, X\}$ is a choice function for p. ∎

It should be pointed out that the converse of the theorem also holds. If the internal axiom of choice holds in $\mathscr{E}$ then the usual construction of the algebra structure on A/R, as outlined just before the theorem, can be carried out in the internal logic. Kock's principle then gives an actual (external) algebra structure, and it is straightforward to verify that the resulting algebra is the desired quotient.

To construct coequalizers in the absence of the axiom of choice, we begin with

A/R and define as much as we can of an algebra structure in the usual way. But when we have a $\lambda\colon \tau^{-1}\{j\} \to A/R$ that cannot be lifted to a function into A, we freely adjoin a new element to serve as $\mu(\langle j, \lambda\rangle)$. This is the idea behind the following construction.

THEOREM. *Let $\tau\colon I \to J$ be a morphism in a Boolean topos $\mathscr{E}$ with natural numbers object. The category of τ-algebras has coequalizers.*

Proof. Let (A, μ) be a τ-algebra and $R \rightarrowtail A\times A$ a congruence on it. Let $\pi\colon A \to A/R$ be the projection to the set-theoretic quotient, and construct the free τ-algebra F on A/R. By Section 6 (and Section 1), we may view F as [Words], where the term Words is defined with respect to $\tau'\colon I \xrightarrow[\tau]{} J \rightarrowtail J+(A/R)$.

Working in the internal logic, call a word (with respect to τ') a *redex* if it has the form $\{\langle\phi, j\rangle\} \cup \{\langle i, \pi(f(i))\rangle |\ \tau(i) = j\}$ for some $j\,\varepsilon J$ and some function f with domain $\{i\,\varepsilon I|\ \tau(i) = j\}$ and values in A. For such a redex, define its *reduct* to be $\{\langle\phi, \pi(\mu(\langle j, f\rangle))\rangle\}$; since R is a congruence, every redex has just one reduct. (Intuitively, a redex is a word for which a value in A/R can be defined by the usual procedure described before, and its reduct is that value, considered as a word.) By induction with respect to Constituent, define a *reduced* word to be one which is not a redex and all of whose constituents are reduced. (Intuitively, w is reduced iff it includes no redex.) For $j\,\varepsilon J$ and λ a function from $\tau^{-1}\{j\}$ to reduced words, define

$$\nu(\langle j, \lambda\rangle) = \begin{cases}\text{the unique word } w \text{ with principal connective } j \text{ and constituent}\\ \text{list } \lambda, \text{ if this word is reduced (i.e. if it is not a redex, since}\\ \text{its constituents are reduced), and its reduct otherwise.}\end{cases}$$

Because $\mathscr{E}$ is a Boolean topos, this defines ν at every such pair $\langle j, \lambda\rangle$, so, by externalizing, we get a τ-algebra (B, ν) where $B =$ [reduced words] and where $\nu\colon \Sigma_J \Pi_\tau \Delta_I B \to B$ is the externalization of the term ν defined above.

We define a morphism $\theta\colon A \to B$ by externalizing $\theta :\equiv \{\langle a, \{\langle\phi, \pi(a)\rangle\}\rangle|\ a\,\varepsilon A\}$. The reduct clause in the definition of ν makes θ a homomorphism of τ-algebras. Clearly $\models aRa' \to \pi(a) = \pi(a') \to \theta(a) = \theta(a')$. Finally, suppose $\langle B', \nu'\rangle$ is another τ-algebra and $\theta'\colon A \to B'$ is a homomorphism such that $\models aRa' \to \theta'(a) = \theta'(a')$. We must prove that there is a unique homomorphism $h\colon B \to B'$ such that $\theta' = h\theta$. By Kock's principle, it suffices to prove this in the internal logic, which we now proceed to do. For each reduced word w, we must define $h(w)\,\varepsilon B'$. If the principal connective of w is some $\pi(a)\,\varepsilon A/R$, then $w = \{\langle\phi, \pi(a)\rangle\} = \theta(a)$, so we must set $h(w) = h\theta(a) = \theta'(a)$; this is well-defined by our assumption on θ'. On the other hand, if w has principal connective $j\,\varepsilon J$ and constituent list λ, then $w = \nu(\langle j, \lambda\rangle)$ and, in order that h be a homomorphism, we must set $h(w) = \nu'(\langle j, h\circ\lambda\rangle)$. By recursion with respect to Constituent, there is a unique such h. This proves the uniqueness of h and almost proves existence; we must still check the homomorphism property $h(\nu(\langle j, \lambda\rangle)) = \nu'(\langle j, h\circ\lambda\rangle)$ in the case where $\nu\langle j, \lambda\rangle$ is not the word w with principal connective j and constituent list λ but its reduct. In this case, we have $\lambda = \theta\circ f$ for some function f from $\tau^{-1}\{j\}$ into A, and $\nu(\langle j, \lambda\rangle) = \langle\phi, \pi(\mu(\langle j, f\rangle))\rangle$. Therefore

$$\begin{aligned} h(\nu(\langle j, \lambda\rangle)) &= \theta'(\mu(\langle j, f\rangle)) \text{ by definition of } h\\ &= \nu'(\langle j, \theta'\circ f\rangle) \text{ as } \theta' \text{ is a homomorphism}\\ &= \nu'(\langle j, h\circ\lambda\rangle) \text{ as } \theta'\circ f = h\circ\theta\circ f = h\circ\lambda\,. \ \blacksquare\end{aligned}$$

COROLLARY. *For any $\tau\colon I \to J$ in a Boolean topos $\mathscr{E}$ with natural numbers object, the category of τ-algebras is internally cocomplete, in the sense of Paré and Schumacher* [30].

Proof. Since we have already constructed coequalizers, it suffices, by III. 3.6 of [30], to construct internally-indexed coproducts. Suppose we have an internal family of τ-algebras (A_p, μ_p) indexed by an object P of $\mathscr{E}$. In other words, we have a $\Delta_P(\tau)$-algebra (A, μ) in $\mathscr{E}/P$. To form the coproduct τ-algebra in $\mathscr{E}$, divide the free algebra F on $\Sigma_P A$ by the smallest congruence making the canonical injection of each A_P a τ-homomorphism (more precisely, making the canonical injection $A \to \Delta_P F$ a $\Delta_P(\tau)$-algebra homomorphism). $\blacksquare$

A few words are in order about our use of Booleanness. This assumption was needed because we cut down the algebra of words to the algebra of reduced words, essentially throwing away all redexes, and then needed to be able to decide whether the result of an operation is a redex. It is, of course, more natural not to throw away redexes but to identify them with their reducts, thus obtaining a quotient, rather than a subset, of the set of words. Unfortunately, putting a τ-algebra structure on such a quotient is a particular case of the theorem being proved. So we arrive at a vicious circle. I do not know whether one can break the circle and show that coequalizers of τ-algebras exist under no special hypotheses on $\mathscr{E}$ (except the existence of natural number objects). Even if this turns out to be impossible, one might still obtain a common generalization of the present result and Rosebrugh's [34] by requiring only that $\mathscr{E}$ be Grothendieck over a Boolean topos.

Appendix to Section 7. In this appendix we show that not every Boolean topos is Grothendieck over a topos satisfying the axiom of choice; thus, the main theorem of Section 7 is not covered by the theorem of Rosebrugh [34]. In fact, we show that the "best" Boolean topoi without choice, namely the well-pointed ones (essentially models of set theory [11, 26]) can be used as the required example.

PROPOSITION. *Let $f\colon \mathscr{E} \to \mathscr{F}$ be a geometric morphism, where $\mathscr{E}$ is well-pointed and $\mathscr{F}$ satisfies the axiom of choice. Then $\mathscr{E}$ also satisfies the axiom of choice.*

Proof. It is well-known [11] that well-pointedness is equivalent to the conjunction of three conditions: Booleanness, two-valuedness (0 and 1 are the only subobjects of 1), and supports (epimorphisms to subobjects of 1) split. It therefore implies that subobjects of 1 generate [17, Lemma 5.33].

Let $\mathscr{E} \xrightarrow[q]{} \mathscr{S} \xrightarrow[i]{} \mathscr{F}$ be the factorization of f as a surjection q followed by an inclusion [17, Section 4.1]. We claim first that $\mathscr{S}$ inherits the axiom of choice from $\mathscr{F}$. Indeed, let $\alpha\colon A \twoheadrightarrow B$ be an epimorphism in $\mathscr{S}$. Its image under i_* need not be epi, but we consider its epi-mono factorization

$$i_*(\alpha)\colon i_*(A) \xrightarrow[\beta]{} J \xrightarrow[\gamma]{} i_*(B)\,.$$

Since i^* preserves both epimorphisms and monomorphisms, and since i is an inclusion, we have the following factorization of α:

$$A \cong i^* i_*(A) \underset{i^*(\beta)}{\twoheadrightarrow} i^*(J) \underset{i^*(\gamma)}{\rightarrowtail} i^* i_*(B) \cong B ,$$

where the isomorphisms are given by the counit of the adjunction $i^* \dashv i_*$. Since α is epi, $i^*(\gamma)$ must be iso. And the axiom of choice in $\mathscr{F}$ guarantees that the epimorphism β is split. It follows that $i^*(\beta)$ and therefore α are split. This completes the proof of the axiom of choice in $\mathscr{S}$. Two thirds of well-pointedness of $\mathscr{S}$ follows, for the axiom of choice implies that supports split (trivially) and that $\mathscr{S}$ is Boolean (by Diaconescu's theorem [7]). We next show that $\mathscr{S}$ inherits the remaining third of well-pointedness, namely two-valuedness, from $\mathscr{E}$. Indeed, suppose U is a subobject of 1 in $\mathscr{S}$. Since $\mathscr{E}$ is two-valued, one of $0 \rightarrowtail U$ and $U \rightarrowtail 1$ must be sent to an isomorphism by q^*, and must therefore already be an isomorphism since q is a surjection.

We have shown that $\mathscr{S}$ is a well-pointed topos satisfying the axiom of choice. It remains only to point out that any such topos can play the role of the topos of sets in Theorem 5.39 of [17], so $\mathscr{E}$, being a Boolean $\mathscr{S}$-topos where subobjects of 1 generate, must satisfy the axiom of choice. ■

8. Free algebras subject to identities. We turn to the study of varieties whose definition involves identities. Formally, an identity for τ-algebras is a triple (V, w, w'), usually written $w = w'(V)$, where V is a set (of "variables") and w and w' are words in the free τ-algebra $F(V)$. A τ-algebra (A, μ) *satisfies* this identity if, for every function from V into A, the induced homomorphism f: $F(V) \to A$ has $f(w) = f(w')$. This definition makes sense in any topos, and so does its extension to an (internally) indexed family of identities and an algebra satisfying such a family. In the topos of sets, it is customary to omit reference to V and write an identity simply as $w = w'$; this is legitimate because if $V \subseteq W$ then any $V \to A$ can be extended to $W \to A$, but in other topoi the reference to V is essential.

We are interested in constructing free algebras for any variety, i.e., for any τ and any family of identities. As in Section 1, we reduce the problem to constructing initial algebras for all varieties.

This problem contains, as a special case, the construction of coequalizers of τ-algebras. Indeed, the coequalizer of $(X, \varrho) \rightrightarrows (A, \mu)$ may be viewed as the initial algebra for τ': $I \underset{\tau}{\to} J \rightarrowtail J + A$ subject to identities guaranteeing that the canonical map from A into the algebra is a τ-homomorphism and coequalizes the two given maps from X to A. The results of Section 7 therefore show that, in the absence of the axiom of choice, we cannot construct the desired algebras in the usual fashion, by starting with the initial τ-algebra A_0 and identifying words when forced to do so by the identities. At best, we must expect to add new words, to serve as values of the operations j on argument lists λ that cannot be lifted to A_0.

But the situation here is worse than in Section 7 in that, after we add the necessary new words, obtaining a τ-algebra A_1 with a canonical homomorphism α: $A_0 \to A_1$, the identities will force more identifications — identifications among the new words, and possibly even among the "old" ones if there are maps $V \to \alpha(A_0)$ that cannot be factored through α. These identifications will, like the earlier ones, force the addition of more words, requiring more identifications, requiring more words, ... *ad infinitum*. The question is, can we somehow bound this *infinitum*. We shall show in the next section that we cannot (provided certain large cardinal axioms are consistent with set theory). Free algebras simply need not exist. Furthermore, the topos in which this happens has logical properties as nice as possible except for the failure of the axiom of choice; it is a model of Zermelo-Fraenkel set theory (ZF). In particular, in contrast to the situation in Section 7, Booleanness (or even well-pointedness) cannot overcome the difficulties presented by the absence of the axiom of choice.

In this section, we present some positive results on existence of free algebras. The hardest of these, Proposition 2, gives a sufficient condition in the context of ZF; its necessity is the main result of the next section. We begin, however, with a much simpler result for more general topoi.

PROPOSITION 1. *In a topos with natural numbers object and the internal axiom of choice, free algebras exist for any variety.*

The proof is left to the reader since it is merely the translation into the interna logic of the familiar construction [13, 31] of such free algebras as algebras of equiv alence classes of elements of the free τ-algebras without identities.

Rosebrugh [33] has shown that the same result holds in any topos with natural numbers object that is Grothendieck over a topos satisfying the external axiom of choice.

From here on, we work in ZF rather than in topoi.

PROPOSITION 2. (ZF). *If the regular cardinals are cofinal in the class of ordinals, then free algebras exist for every variety.*

Proof. As remarked earlier, it suffices to construct an initial algebra for every variety. So let a variety be given, determined by τ: $I \to J$ and a set of identities $w = w'(V)$, where we may assume that V is the same for all the identities because we can always increase V without affecting satisfaction of the identity.

We construct a direct system of τ-algebras A_α (not necessarily satisfying the given identities) and τ-homomorphisms $f_{\beta\alpha}$: $A_\alpha \to A_\beta$, for $\alpha < \beta$, indexed by the ordinals. (Readers who worry that such a system is a proper class and therefore outside the scope of ZF set theory are referred to [16] for an account of how one can fit such notions into ZF. They may also find it reassuring that, from an appropriate ordinal on, all the $f_{\beta\alpha}$ are isomorphisms.)

We begin by letting A_0 be the initial τ-algebra (which exists by Section 6). At limit ordinals λ, we let A_λ be the direct limit of the system of τ-algebras A_α $(\alpha < \lambda)$ with respect to the homomorphisms $f_{\beta\alpha} (\alpha < \beta < \lambda)$. The corollary in Section 7 guarantees the existence of this direct limit, but the reader should note that A_λ is *not* in general the direct limit of the *sets* A_α. Finally, for a successor ordinal $\alpha + 1$, we let $A_{\alpha+1}$ be the quotient of A_α by the smallest congruence R such that for all the given identities

$w = w'$ and all homomorphisms $h\colon F(V) \to A_\alpha$, $\langle h(w), h(w')\rangle \in R$. Let $f_{\alpha+1,\alpha}$ be the canonical homomorphism from A_α to $A_{\alpha+1}$, and, for $\gamma < \alpha$, let $f_{\alpha+1,\gamma} = f_{\alpha+1,\alpha} f_{\alpha,\gamma}$ as required for a direct system. The results of Section 7 guarantee the existence of $A_{\alpha+1}$, but they also show that $f_{\alpha+1,\alpha}$, though an epimorphism of τ-algebras, need not be surjective as a map of sets.

Let B be any algebra of our variety, i.e., any τ-algebra satisfying the given identities. By definition of A_0, there is a unique τ-homomorphism $g_0\colon A_0 \to B$. Its kernel K has the property stipulated in the definition of A_1, that $\langle h(w), h(w')\rangle \in K$ for all identities $w = w'$ and all homomorphisms $h\colon F(V) \to A_0$. Indeed, since $w = w'$ is satisfied by B, $g_0 h(w) = g_0 h(w')$. So g_0 factors through $f_{1,0}$ yielding a τ-homomorphism $g_1\colon A_1 \to B$. The uniqueness of the factorization and the uniqueness of g_0 combine to imply that g_1 is in fact the only τ-homomorphism from A_1 to B. Continuing in this manner, and using the universal property of direct limits at limit stages, we find by induction on α that every A_α has a unique τ-homomorphism g_α to B, and $g_\alpha \cdot f_{\alpha\beta} = g_\beta$ for all $\beta < \alpha$. Thus, if any one of the τ-algebras A_α satisfies the given identities and is therefore in our variety, then it will be the desired initial algebra of the variety. (Furthermore, $f_{\beta\alpha}$ will be an isomorphism for every $\beta > \alpha$, so the construction essentially terminates once it achieves its goal.) The proof will thus be complete once we show that some A_α satisfies the identities.

For any α, the kernels R_β of the homomorphisms $f_{\beta\alpha}\colon A_\alpha \to A_\beta$ for $\beta > \alpha$ form a non-decreasing sequence of congruence relations on A_α. Since there are only a set of congruence relations on A_α and since no set is cofinal in the ordinals, there must be an ordinal $\gamma > \alpha$ such that $R_\beta = R_\gamma$ for all $\beta \geqslant \gamma$. Write α^* for the least such γ, and define a non-decreasing sequence of ordinals α_ξ as follows.

$$\alpha_0 = 0;$$
$$\alpha_{\xi+1} = (\alpha_\xi)^*;$$
$$\alpha_\lambda = \text{supremum } \{\alpha_\xi \mid \xi < \lambda\} \text{ for limit ordinals } \lambda.$$

To simplify notation, let us agree that a superscript ξ has the same meaning as a subscript α_ξ; thus, for example, $f^{\eta\xi}\colon A^\xi \to A^\eta$ means $f_{\alpha_\eta\alpha_\xi}\colon A_{\alpha_\xi} \to A_{\alpha_\eta}$. The defining property of α^* guarantees that, if two elements of A^ξ have the same image (under $f^{\eta\xi}$) in some later A^η then they have the same image already in $A^{\xi+1}$. It follows that, for a limit ordinal λ, every element a of the *set-theoretic* (not τ-algebra) direct limit of the $A^\xi (\xi < \lambda)$ has a canonically chosen precursor b in some A^η, $\eta < \lambda$ (i.e. b maps to a under the canonical map of A^η into the direct limit). To find b, first find the smallest ξ such that a has a precursor in A^ξ; all its precursors in A^ξ have the same image in $A^{\xi+1}$, by the above, and this image serves as b. (The point of all this is to be able to choose a specific precursor b for a without invoking the axiom of choice.)

If there is no cofinal map of I (the domain of τ) into the limit ordinal λ, then the set-theoretic direct limit A' of A^ξ $(\xi < \lambda)$ can be equipped with a τ-algebra structure μ' as follows. For any operation symbol $j \in J$ and any list of arguments $g\colon \tau^{-1}\{j\} \to A'$, let $b(i) \in A^{\eta(i)}$ be the canonically chosen precursor of $g(i)$ for each $i \in \tau^{-1}\{j\}$. By the assumption on I and λ, the supremum ξ of all the $\eta(i)$ is smaller than λ. Let

$c(i) = f^{\xi\eta(i)}(b(i))$, and set $\mu'(\langle j, g\rangle)$ equal to the image in A' of $\mu^\xi(\langle j, c\rangle)$ (where μ^ξ is the algebra structure of A^ξ). It is straightforward to verify that A', with this τ-algebra structure, has the universal property of a direct limit of τ-algebras, so $A' = A^\lambda$.

Now assume, in addition to the hypothesis of the preceding paragraph, that V cannot be mapped cofinally into λ. Consider any one of the given identities, say $w = w'$, and any map $h\colon V \to A' = A^\lambda$. For each $v \in V$, let $b(v) \in A^{\eta(v)}$ be the canonically chosen precursor of $h(V)$. By the new assumption on V and λ, the supremum ξ of all the $\eta(v)$ is smaller than λ. Let $c(v) = f^{\xi\eta(v)}(b(v))$, so $c\colon V \to A^\xi$ and $f^{\lambda\xi} \cdot c = h$. For the homomorphism $\bar{c}\colon F(V) \to A^\xi$ extending c, we have that $\bar{c}(w)$ and $\bar{c}(w')$, though not necessarily equal in A^ξ, have the same image in the next algebra $A_{\alpha_\xi+1}$ of our original sequence and, *a fortiori*, also in $A^{\xi+1}$. So, if $\bar{h}\colon F(V) \to A^\lambda$ is the homomorphism extending h,

$$\bar{h}(w) = f^{\lambda,\xi+1} f^{\xi+1,\xi} \bar{c}(w) = f^{\lambda,\xi+1} f^{\xi+1,\xi} \bar{c}(w') = \bar{h}(w').$$

Thus, to complete the proof of Proposition 2, all we need is a limit ordinal λ into which neither I nor V can be cofinally mapped.

By Hartogs's theorem, there is an ordinal α that admits no one-to-one map into $\mathscr{P}(I \cup V)$; the hypothesis of the proposition provides a regular cardinal $\lambda \geqslant \alpha$. Then $I \cup V$ cannot be mapped onto α or any larger ordinal, for if f were such a map then the function $\xi \mapsto f^{-1}\{\xi\}$ would contradict the choice of α. In particular, neither I nor V (nor even $I \cup V$) can be mapped onto a cofinal subset of λ, so A^λ is the initial algebra for our variety. ∎

9. Consistency of non-existence of free algebras. This section is devoted to the proof of the converse of Proposition 2 of the preceding section.

THEOREM. (ZF) *If the regular cardinals are not cofinal in the class of all ordinals, then there is a variety with no initial algebra.*

Gitik [12] has constructed a model of ZF in which $\omega (= \aleph_0 = N)$ is the only infinite regular cardinal, assuming the existence of a model of ZFC (= ZF plus the axiom of choice) in which the strongly compact cardinals [9] are cofinal in the ordinals. This construction allows us to convert our theorem into a relative consistency result.

COROLLARY. *If it is consistent with* ZFC *that the strongly compact cardinals are cofinal in the ordinals, then it is consistent with* ZF *that there is a variety with no initial algebra.* ∎

To avoid notational complications, we shall prove the theorem under the stronger hypothesis that ω is the only infinite regular cardinal. This special case suffices to yield the corollary. For the general case, one would replace the ω-ary operation in the proof below with a set of $\varkappa$-ary operations, one for each regular cardinal $\varkappa$.

Suppose, then, that ω is the only infinite regular cardinal. Since the cofinality of any limit ordinal is always such a cardinal, our assumption implies that every limit ordinal is the supremum of a sequence of order type ω.

The variety which, we shall show, has no initial algebra has three operation symbols, a 0-ary (constant) symbol 0, a unary symbol S, and an ω-ary symbol sup. Before completing the definition of the variety by specifying a set of identities, we indicate the intended interpretation of the operation symbols to motivate the choice of identities. We intend 0 to denote the ordinal number zero, S to denote the successor operation on ordinals, and sup to denote the least upper bound operation on ω-sequences of ordinals. If they formed a set rather than a proper class, the ordinals would constitute a τ-algebra for the τ described here. The simplest way to describe the identities we want is to say that they are all the identities, in ω or fewer variables, that are satisfied in this interpretation. (Although satisfaction of first-order formulas in proper classes is not definable in ZF, there is no such problem with satisfaction of identities.) For readers who feel uncomfortable with this definition, we give a more explicit list of identities; other readers can skip the next paragraph except for the notational conventions following (1).

The following identities (1) through (5) suffice to define the desired variety. The x's and y's in these identities are intended to be variables taken from a fixed countable list. First, we have $2^{\aleph_0}$ identities saying that sup does not depend on the order of its arguments nor on repetitions among them.

(1) $\quad \sup(x_0, x_1, \ldots) = \sup(y_0, y_1, \ldots) \quad$ if $\quad \{x_0, x_1, \ldots\} = \{y_0, y_1, \ldots\}$ (as sets of variables).

This identity allows us to unambiguously write $\sup Q$, for any nonempty countable set Q, to mean $\sup(q_0, q_1, \ldots)$ where the q's are an enumeration of Q in any order (possibly with repetitions). In the following identities, X and Y stand for $\{x_0, x_1, \ldots\}$ and $\{y_0, y_1, \ldots\}$ or subsets thereof

(2) $\quad \sup\big(X \cup \{\sup(X \cup Y)\}\big) = \sup(X \cup Y)$,

(3) $\quad \sup\big(X \cup \{S(\sup(X \cup Y))\}\big) = S(\sup(X \cup Y))$,

(4) $\quad \sup P = \sup Q$ where every element of $P \cup Q$ is of the form $\sup R$, and every element of P (resp. Q) occurs as an element of R for some $\sup R \in Q$ (resp. P),

(5) $\quad \sup\{0\} = 0$.

Suppose this variety had an initial algebra F. We note first that F has no proper subalgebra A, for, by initiality, F would admit a homomorphism into A, whose composite with the inclusion $A \to F$ would, by initiality again, have to be id_F, so the inclusion must be surjective. It follows in particular that every element of F is 0 or Sa for some $a \in F$, or $\sup X$ for some $X \subseteq F$, since the elements of these forms clearly constitute a subalgebra. To analyze the structure of F in more detail, we introduce the operator Γ on subsets $X \subseteq F$ that applies the algebra operations once to elements of X, and we formalize the notion that iterated application of this operator should produce the whole algebra F from the empty set. The definition of Γ is

$$\Gamma(X) = X \cup \{0\} \cup \{S(a) \mid a \in X\} \cup \{\sup(A) \mid A \text{ countable}, \emptyset \neq A \subseteq X\}.$$

By induction on ordinals, define $\Gamma^0 = \emptyset$, $\Gamma^{\alpha+1} = \Gamma(\Gamma^\alpha)$, and $\Gamma^\lambda = \bigcup_{\alpha<\lambda} \Gamma^\alpha$ for limit ordinals λ. The sequence of subsets Γ^α of F is clearly non-decreasing. For some α, the difference $\Gamma^{\alpha+1} - \Gamma^\alpha$ must be empty, as otherwise we would have a proper class of distinct such differences (one for each ordinal α) which is absurd as F is a set. Let θ be the smallest ordinal for which $\Gamma^{\theta+1} = \Gamma^\theta$. This equation implies that Γ^θ is a subalgebra of F, so by our earlier remarks $\Gamma^\theta = F$. For each element $a \in F$, there is a first $\beta \leqslant \theta$ such that $a \in \Gamma^\beta$; inspection of the definition of Γ^0 and Γ^λ for λ a limit shows that this β must be the successor of some α which we call the *rank* of a. Thus, $a \in \Gamma^{\mathrm{rank}(a)+1} - \Gamma^{\mathrm{rank}(a)}$. By definition of $\Gamma^{\alpha+1}$, we see that a is either 0, or $S(b)$ with $\mathrm{rank}(b) < \mathrm{rank}(a)$, or $\sup(B)$ with $\mathrm{rank}(b) < \mathrm{rank}(a)$ for all $b \in B$. Note also that all ranks are $< \theta$.

Fix an arbitrary limit ordinal $\varkappa$. We define an algebra structure on the set $\varkappa+1$ of all ordinals $\leqslant \varkappa$ exactly like the intended interpretation in the class of all ordinal numbers except that $S(\varkappa)$ is defined to be $\varkappa$ (rather than $\varkappa+1$ which is not in our set). That this algebra, $\varkappa+1$, is in our variety can be seen either by inspection of the explicit list of identities (1) through (5) or by noting that $\varkappa+1$ is the quotient of the "algebra" of all ordinal numbers obtained by identifying everything beyond $\varkappa$ with $\varkappa$. (An identity satisfied in an algebra A remains satisfied in a quotient B *provided* the projection $A \to B$ admit a set-theoretic section. The proviso is satisfied here, and there are no difficulties due to the fact that the ordinals form a proper class.) Since F is assumed to be initial, let $\pi\colon F \to \varkappa+1$ be the unique homomorphism.

CLAIM. *Every ordinal $\alpha < \varkappa$ is the image under π of a unique element of F.*

We prove the claim by induction on α.

Case 1. $\alpha = 0$. The existence is clear, as π, being a homomorphism, must send 0 to 0. To prove uniqueness, suppose $\pi(a) = 0$ but $a \neq 0$. We may assume that a has been chosen to have the smallest possible rank. Clearly, $a \neq S(b)$ for any $b \in F$, as $\pi(S(b)) = S(\pi(b)) \neq 0$. So a must be $\sup(B)$ with $\mathrm{rank}(b) < \mathrm{rank}(a)$ for every $b \in B$. As π is a homomorphism, we have $0 = \pi(a) = \pi(\sup B) = \sup(\pi(B))$. In view of the definition of sup in the algebra $\varkappa+1$, we have $\pi(b) = 0$ for every $b \in B$. By the minimality of $\mathrm{rank}(a)$, we may infer $b = 0$ in F for every $b \in B$. But then $a = \sup\{0\} = 0$ by one of the defining identities (5) of our variety.

Case 2. $\alpha = \beta+1$. By induction hypothesis, there is a unique $b \in F$ with $\pi(b) = \beta$. Then $\pi(S(b)) = S(\pi(b)) = S(\beta) = \alpha$ (as $\beta < \varkappa$). This proves existence. For uniqueness, suppose that a is a counterexample ($\pi(a) = \alpha$ but $a \neq Sb$) of minimum rank. Clearly $a \neq 0$, and almost as clearly $a \neq S(c)$ for any c; indeed, $\beta+1 = \alpha = \pi(S(c)) = S(\pi(c)) = \pi(c)+1$ would imply $\beta = \pi(c)$, so by the uniqueness part of the induction hypothesis $c = b$ and $a = Sb$. So $a = \sup C$ for some $C \subseteq F$ with $\mathrm{rank}(c) < \mathrm{rank}(a)$ for all $c \in C$. As π is a homomorphism, $\beta+1 = \sup(\pi(C))$, which is possible only if $\beta+1 \in \pi(C)$, since the successor ordinal $\beta+1$ cannot be the supremum of strictly smaller ordinals. By the minimality of $\mathrm{rank}(a)$, we must have $S(b) \in C$. Let $C' = C - \{S(b)\}$. Thus, $\pi(C')$ consists of ordinals $\leqslant \beta$, and

$$\pi(\sup(C' \cup \{b\})) = \sup(\pi(C') \cup \{\beta\}) = \beta = \pi(b)$$

so, by induction hypothesis, $\sup(C' \cup \{b\}) = b$. Now

$$\begin{aligned} a = \sup(C' \cup \{S(b)\}) &= \sup(C' \cup \{S(\sup(C' \cup \{b\}))\}) \\ &= S(\sup(C' \cup \{b\})) \text{ by identity (3)} \\ &= S(b)\,. \end{aligned}$$

Case 3. α is a limit ordinal. Since ω is the only infinite regular cardinal, there is an increasing ω-sequence β_n with limit α. By induction hypothesis, each β_n is $\pi(b_n)$ for a unique $b_n \in F$. Let $B = \{b_n|\ n \in \omega\}$. Then

$$\pi(\sup B) = \sup\{\beta_n|\ n \in \omega\} = \alpha\,,$$

which proves existence. For uniqueness, suppose again that a is a counterexample ($\pi(a) = \alpha$ but $a \neq \sup B$ for our fixed B) of minimum rank. Since α is a limit ordinal, a cannot have the form 0 or $S(c)$, so $a = \sup(C)$ for some set C of elements of F of strictly lower rank than a.

Subcase 1. For each $c \in C$, $\pi(c) < \alpha$. We shall rewrite every element of $B \cup C$ as a supremum in such a way that identity (4) will become applicable. Let $\{z_\sigma|\ \sigma < \gamma\}$ be an enumeration of $B \cup C$ in order of increasing values of π, i.e. $\sigma < \sigma'$ implies $\pi(z_\sigma) \leqslant \pi(z_{\sigma'})$. Note that, since $\sup(\pi(B)) = \alpha$ and $\sup(\pi(C)) = \pi(\sup(C)) = \pi(a) = \alpha$ and since $\alpha \notin \pi(C)$, both B and C must be cofinal in the z-enumeration. Also note that each z_σ equals $\sup\{z_{\sigma'}|\ \sigma' \leqslant \sigma\}$ in F, because of the trivial equation $\pi(z_\sigma) = \sup\{\pi(z_{\sigma'})|\ \sigma' \leqslant \sigma\}$ and the uniqueness part of the induction hypothesis. Let all the z's be expressed as sups in this way. Then the mutual cofinality of B and C allows us to apply the identity (4) to conclude $\sup B = \sup C = a$.

Note that the argument in Subcase 1 did not use the fact that all elements of C have lower rank than a.

Subcase 2. For some $c \in C$, $\pi(c) = \alpha$. The minimality of rank(a) shows that the only such c is $\sup B$. Let $C' = C - \{\sup B\}$. By subcase 1, $\sup(C' \cup B) = \sup B$, and therefore

$$\begin{aligned} a = \sup C = \sup(C' \cup \{\sup B\}) &= \sup(C' \cup \{\sup(C' \cup B)\}) \\ &= \sup(C' \cup B) \text{ by identity (2)} \\ &= \sup B\,. \end{aligned}$$

This completes the proof of the claim.

If we now let $\varkappa$ vary, we find that arbitrarily large ordinals $\varkappa$ admit one-to-one mappings into F. This contradicts Hartogs's theorem, so there cannot exist an initial algebra for our variety. ■

We close this section with a brief summary of what is known concerning the set-theoretic strength of the proposition that the regular cardinals are not cofinal in the ordinals (or, equivalently, that some variety lacks free algebras). As mentioned earlier, Gitik [12] has proved the consistency of this proposition relative to a proper class of strongly compact cardinals. Some large cardinal axiom is certainly needed for any such consistency proof, since it follows from Jensen's covering lemma [6]

that the proposition implies the existence of $0^{\#}$. In fact, by work of Jensen and Dodd [8], the proposition implies the existence of inner models with measurable cardinals, and I believe that Mitchell's results [27] will yield inner models with many measurable cardinals.

10. Consistency of non-existence of coequalizers.

THEOREM. *If it is consistent with* ZFC *that the strongly compact cardinals are cofinal in the ordinals, then it is consistent with* ZF *that there is a variety of algebras in which coequalizers do not always exist.*

Proof. We shall work in a universe satisfying ZF plus

(a) ω is the only infinite regular cardinal, and

(b) there is a Dedekind-finite set D with a linear ordering $\preccurlyeq$ and a map p of D onto ω such that $d_1 \preccurlyeq d_2$ implies $p(d_1) \leqslant p(d_2)$.

(A Dedekind-finite set is one into which ω cannot be injectively mapped.) Our first task is to obtain a model in which these hypotheses hold.

It is easy to satisfy (b) in a Fraenkel–Mostowski model (see [15] for definitions). Take a countable set A of atoms and label them a_{nq} with $n \in \omega$ and $q \in \boldsymbol{Q}$ the set of rationals. Let $\preccurlyeq$ be the lexicographic order of the atoms, with respect to the standard orderings of ω and $\boldsymbol{Q}$, and let $p\colon A \to \omega$ map each a_{nq} to its first subscript n. Let G be the group of permutations of A that preserve both $\preccurlyeq$ and p, and let $\mathscr{F}$ be the finite-support filter, generated by the subgroups $H_E = \{g \in G|\ g \text{ fixes } E \text{ pointwise}\}$ for finite $E \subseteq A$. Then the permutation model determined by G and $\mathscr{F}$ is easily seen to satisfy (b), with D being the set A of atoms and $\preccurlyeq$ and p being as defined above.

By the Jech–Sochor theorem [15], every model M of ZF has a symmetric Cohen extension M' satisfying (b). The hypothesis of our theorem permits us to apply Gitik's construction, already used in Section 9, to obtain an M satisfying (a). The Cohen extension M' satisfying (b) also satisfies (a), since it has no new ordinals and contains all the cofinal maps in M from ω into each limit ordinal. Thus, M' is the desired model.

From now on, we work in a universe satisfying ZF and (a) and (b). The variety which lacks coequalizers will be a slight variation of the one shown to lack free algebras in Section 9. It has a unary operation symbol S and an ω-ary operation symbol sup, but instead of a single constant symbol it has a D-indexed family of constant symbols $\bar{d}$ for $d \in D$. Its identities are all identities in a fixed countable set of variables, not involving constants, that are satisfied in the class of ordinal numbers with the standard interpretations of S and sup (or, more explicitly, identities (1) through (4) from Section 9), plus the identities

$$S(\bar{d}) = \bar{d} \quad \text{for each } d \in D\,, \tag{6}$$

and

$$\sup(\bar{d}_0, \bar{d}_1, \ldots) = \max\{d_0, d_1, \ldots\} \quad \text{for each } d_0, d_1, \ldots \in D\,. \tag{7}$$

In (7), the maximum is taken with respect to the linear ordering $\preccurlyeq$ on D given in (b), and it exists because $\{d_0, d_1, \ldots\}$, being the range of an ω-sequence in a Dedekind-finite set, is finite.

For each ordinal α, we define an algebra B_α as follows. Its underlying set is the ordinal $\omega+\alpha+1$. Each constant $\bar{d}$ is interpreted as $p(d) \in \omega$. Sup is interpreted as least upper bound with respect to the usual ordering of $\omega+\alpha+1$, and S is interpreted as the identity map on the first ω elements and the last element but as the successor map on the intervening α elements. The strange definition of S on the first ω elements guarantees that (6) is satisfied, while (7) is satisfied because p preserves order. (1) through (4) are easy to see directly, or one can use the fact that B_α is isomorphic to the quotient of the class of all ordinals (with standard S and sup) by the congruence that identifies two distinct ordinals iff either they are less than ω^2 and their difference is finite, or they are both $\geqslant \omega^2+\alpha$.

There is also an algebra structure on D, interpreting $\bar{d}$ as d, sup as maximum with respect to $\preccurlyeq$ (since any ω-sequence in D has finite range), and S as the identity. All the defining identities of our variety clearly hold in D. (This is obvious in the formulation using (1) through (4). In the formulation using the identities true in the class of ordinals, one observes first that they are satisfied in each B_α and second that, when such an identity, involving countably many variables, is interpreted in the Dedekind finite set D, there can be only finitely many distinct values of variables, so the relevant part of D, a finite linearly ordered set, is isomorphic to a finite part of (the initial ω-segment of) B_α.) It is trivial to check that, because of identities (6) and (7), the algebra D is the initial algebra in our variety.

Suppose D had a quotient algebra A by the congruence

$$R = \{\langle d_1, d_2\rangle \mid p(d_1) = p(d_2)\},$$

and let $q\colon D \to A$ be the canonical homomorphism. For any ordinal α, the unique homomorphism $h\colon D \to B_\alpha (d \mapsto p(d))$ has kernel R, so it factors through A, say as $D \xrightarrow[q]{} A \xrightarrow[\pi]{} B_\alpha$. We write $\hat{n}$ for the element $q(d) \in A$ (the denotation of $\bar{d}$) for any $d \in D$ with $p(d) = n$; this is well-defined because q has kernel R. As in Section 9, we define an operation Γ on subsets of A and a rank function on A. We set

$$\Gamma(X) = X \cup \{\hat{n} \mid n \in \omega\} \cup \{S(x) \mid x \in X\} \cup \{\sup(C) \mid C \text{ countable}, C \subseteq X\},$$

$$\Gamma^0 = \varnothing, \quad \Gamma^{\alpha+1} = \Gamma(\Gamma^\alpha), \quad \Gamma^\lambda = \bigcup_{\alpha<\lambda} \Gamma^\alpha \text{ for limit } \lambda.$$

Let θ be the least ordinal such that $\Gamma^{\theta+1} = \Gamma^\theta$; it exists because A is a set. Then Γ^θ is a subalgebra of A containing the range of q. By the universal property of A, there is a homomorphism $A \to \Gamma^\theta$ whose composite with the inclusion $\Gamma^\theta \rightarrowtail A$ is the identity map of A; therefore $\Gamma^\theta = A$. The rank of $x \in A$ is the unique α such that $x \in \Gamma^{\alpha+1} - \Gamma^\alpha$.

Still proceeding as in Section 9, we claim that every element ξ of $\omega+\alpha \subseteq B_\alpha$ is $\pi(a)$ for a unique $a \in A$. The proof is by induction on ξ.

Case 0. $\xi = n < \omega$. By definition of π, we have $\pi(\hat{n}) = n$. To prove uniqueness, suppose $a \in A$ were a counterexample, $\pi(a) = n$ but $a \neq \hat{n}$, with rank(a) as small as possible. If $a = S(b)$ with rank$(b)<$rank(a), then $n = \pi(a) = S(\pi(b))$, which implies $\pi(b) = \hat{n}$ by definition of B_α. Minimality of rank(a) gives $b = \hat{n}$, and then (6) gives $a = S(\hat{n}) = \hat{n}$. On the other hand, if $a = \sup(C)$ with all elements of C of lower rank than a, then $n = \sup(\pi(C))$, so, by definition of B_α, $\pi(C)$ is a subset of $\{m \mid m \leqslant n\}$ containing n. By induction hypothesis and minimality of rank(a), C is a subset of $\{\hat{m} \mid m \leqslant n\}$ containing $\hat{n}$. By (7), $a = \sup(C) = \hat{n}$.

Case 1. $\xi = \omega$. The existence is clear as $\pi(\sup\{\hat{n} \mid n \in \omega\}) = \sup\{n \mid n \in \omega\} = \omega$. The uniqueness is proved exactly as in Case 3 of the proof in Section 9.

Case 2 (resp. Case 3). $\xi = \omega+\eta$ with η a successor (resp. limit) ordinal. These cases are treated exactly like the corresponding cases in Section 9.

Having established the claim, we let α vary and, as in Section 9, obtain a contradiction with Hartogs's theorem.

References

[1] K. J. Barwise, *Infinitary logic and admissible sets*, J. Symb. Logic 34 (1969), pp. 226–252.
[2] G. Birkhoff, *On the structure of abstract algebras*, Proc. Cambridge Phil. Soc. 31 (1935), pp. 433–454.
[3] T. Brook, *Order and Recursion in Topoi*, Notes on Pure Mathematics 9, Australian National Univ., Canberra 1977.
[4] J. Cole, *Categories of sets and models of set theory*, Proc. Bertrand Russell Memorial Logic Conf., Uldum 1971 (ed. J. Bell and A. Slomson) Leeds (1973), pp. 351–399.
[5] M. Coste, *Langage interne d'un topos*, Seminaire Bénabou, Univ. Paris-Nord 1972.
[6] K. Devlin and R. B. Jensen, *Marginalia to a theorem of Silver*, ISILC Logic Conf., Kiel 1974 (ed. G. H. Müller, A. Oberschelp, and K. Potthoff) Springer Lecture Notes in Math. 499 (1975), pp. 115–142.
[7] R. Diaconescu, *Axiom of choice and complementation*, Proc. Amer. Math. Soc. 51 (1975), pp. 176–178.
[8] A. Dodd and R. B. Jensen, *The Core Model*, Annals of Math. Logic 20 (1981), pp. 43–75.
[9] F. Drake, *Set Theory, an Introduction to Large Cardinals*, North-Holland, 1974.
[10] M. Fourman, *The logic of topoi*, Handbook of Mathematical Logic (ed. J. Barwise), North-Holland (1978), pp. 1053–1090.
[11] P. Freyd, *Aspects of topoi*, Bull. Austral. Math. Soc. 7 (1972), pp. 1–76 and 467–480.
[12] M. Gitik, *All uncountable cardinals can be singular*, Israel J. Math. 35 (1980), pp. 61–88.
[13] G. Grätzer, *Universal Algebra*, van Nostrand, 1968.
[14] P. Halmos, *Naive Set Theory*, van Nostrand, 1960.
[15] T. Jech, *The Axiom of Choice*, North-Holland, 1973.
[16] R. B. Jensen, *Modelle der Megenlehre*, Springer Lecture Notes in Math. 37 (1967).
[17] P. T. Johnstone, *Topos Theory*, Academic Press, 1977.
[18] — and R. Paré, eds., *Indexed Categories and their Applications*, Springer Lecture Notes in Math. 661 (1978).
[19] A. Kock, *Linear Algebra and Projective Geometry in the Zariski Topos*, Aarhus Univ. Preprint Series (1974/75), no. 4.
[20] — and C. J. Mikkelsen, *Non-standard Extensions in the Theory of Toposes*, Aarhus Univ. Preprint Series (1971/72), no. 25.
[21] J. Lambek, *From types to sets*, Advances in Math. 36 (1980), pp. 113–164.

[22] F. W. Lawvere, *Quantifiers and Sheaves,* Actes du Congrès International des Mathématiciens, Nice 1970, pp. 329–334.
[23] B. Lesaffre, *Structures algébriques dans les topos élémentaires,* C. R. Acad. Sci. Paris 277 (1973), ser. A, pp. 663–666.
[24] S. MacLane, *Categories for the Working Mathematician,* Springer Graduate Texts in Mathematics 5 (1971).
[25] C. J. Mikkelsen, *Lattice-theoretic and Logical Aspects of Elementary Topoi,* Aarhus Univ. Various Publ. Series (1976), no. 25.
[26] W. Mitchell, *Boolean topoi and the theory of sets,* J. Pure and Applied Algebra 2 (1972), pp. 261–274.
[27] — *Constructibility and Large Cardinals,* Lectures at Cambridge Summer Institute in Set Theory (1978).
[28] J. Myhill, *Some properties of intuitionistic Zermelo-Fraenkel set theory,* Cambridge Summer School in Mathematical Logic 1971 (ed. A. R. D. Mathias and H. Rogers) Springer Lecture Notes in Mathematics 337 (1973), pp. 206–231.
[29] G. Osius, *Logical and set theoretic tools in elementary topoi,* Model Theory and Topoi, Springer Lecture Notes in Math. 445 (1975), pp. 297-346.
[30] R. Paré and D. Schumacher, *Abstract families and the adjoint functor theorem,* [18], pp. 1–125.
[31] R. S. Pierce, *Introduction to the Theory of Abstract Algebras,* Holt Rinehart Winston, 1968.
[32] D. Pincus, *Cardinal representatives,* Israel J. Math. 18 (1974), pp. 321–344.
[33] R. Rosebrugh, *Abstract families of algebras,* thesis, Dalhousie Univ. (1977).
[34] — *Coequalizers in algebras for an internal type,* [18], pp. 243–260.
[35] I. Sols, *Bon ordre dans l'objet des nombres naturels d'un topos booléen,* C. R. Acad. Sci. Paris 281 (1975) sér. A, pp. 601–603.
[36] H. Volger, *Completeness theorem for logical categories,* Model Theory and Topoi, Springer Lecture Notes in Math. 445 (1975), pp. 51–86.
[37] — *Logical categories, semantical categories, and topoi,* Ibid., pp. 87–100.

UNIVERSITY OF MICHIGAN
Ann Arbor, Michigan

Accepté par la Rédaction le 15. 12. 1980