A constructive version of Birkhoff's theorem

Jesper Carlström*

Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden

Received 16 April 2007, accepted 28 June 2007
Published online 21 January 2008

Key words Variety, universal algebra, predicativity.
MSC (2000) 03C05, 03F65

A version of Birkhoff's theorem is proved by constructive, predicative, methods. The version we prove has two conditions more than the classical one. First, the class considered is assumed to contain a generic family, which is defined to be a set-indexed family of algebras such that if an identity is valid in every algebra of this family, it is valid in every algebra of the class. Secondly, the class is assumed to be closed under inductive limits.

1 Background and remarks

Birkhoff's theorem states that if a class of algebras (of a fixed signature) is closed under homomorphic images, subalgebras, and products, it can be axiomatized by a set of equations. One might suspect that the constructive content of it is nil: how could it be possible, given only some closure properties of a class of algebras, to extract a set of axioms for it? To investigate this, it is natural first to turn to the original proof by Birkhoff [1, p. 441]. It turns out that he uses products indexed by arbitrary classes, which is not acceptable even to the classical set theorist.[1)] The version reproduced in textbooks [2, p. 68], [3, pp. 166 – 167] uses products that can be proved to be set-indexed by the power set axiom and the full separation axiom (with quantification over all sets). The crucial product is taken over all quotients of a term algebra that are members of the class considered. From the predicative point of view this is not a set. Thus, the standard proof is far from being constructive.

Is there any hope in finding a constructive proof of Birkhoff's theorem? First, we must ask how it should be stated precisely in a constructive context. As is known since Bishop [4], an algebra should be seen constructively as a set together with an equivalence relation and operations which preserve this equivalence relation. In the empty signature, this is simply a *setoid*, as it is often called among constructivists (Bishop called them *sets*). Notice that the only class of setoids that contains the setoid of natural numbers, and is axiomatizable by equations, is the class of *all* setoids. This is so, because the only equational axioms that are satisfied by the setoid of natural numbers are the ones of the form $x = x$. Therefore, Birkhoff's theorem implies that if a class of setoids contains the setoid of natural numbers and is closed under the taking of surjective images, subsetoids, and arbitrary set-indexed products of setoids, it contains all setoids. In particular, Birkhoff's theorem implies that if N^I is subcountable for every set I, then every set is subcountable. I conjecture that this principle is not constructive, although my attempts to demonstrate this have failed. Several people have pointed out to me that the category of assemblies over the natural numbers forms a model of constructive mathematics [5, 6], and that the *modest* sets (those for which each number realizes at most one element) are a class that is closed under the operations considered. However, the proof that M^I is modest whenever M is modest uses the extensionality principle that if functions are pointwise equal, then they are themselves equal. Thus, if we by "constructive" mean also intensional and predicative, that is, formalizable in Martin-Löf's type theory, it is not clear that we can prove that the modest sets do form a class

* e-mail: jesper@math.su.se

1) In fact, the picture is a bit more complicated. The proof itself is acceptable because it says things about "sets" of algebras. However, the theorem stated is about proper classes of algebras. It is impredicative to accept the proof as a proof of the theorem. Clearly Birkhoff does not distinguish between "sets" and "classes". He writes, for instance, "the set $\mathcal{F}(\mathcal{B})$ of all algebras which can be constructed from algebras of $\mathcal{B}$ by the taking of subalgebras, homomorphic images, and direct products (...)" [1, p. 440]. This class is not a set even in ZFC.

that is closed under the operations considered. However, the observations made are enough for suspecting that Birkhoff's theorem is hardly constructive. We therefore look for alternative, but similar, formulations.

The usual way of finding constructive versions of theorems is to try to get rid of the "magical" part. For instance, Bishop often replaced the assumption that a function be *continuous* by the requirement that it be *uniformly continuous*. The key observation defending this habit is that given a function we can often check that it is in fact uniformly continuous, and then we can apply the theorem, so it is not necessary to use the weaker assumption of continuity. To do the analogous thing in our case, we look at a typical application of Birkhoff's theorem. Let us assume that we start with the ring of integers, in the signature $(0, 1, +, -, *)$. We realize that we need to be able to form other structures by taking substructures, quotients, limits, colimits, et cetera. By Birkhoff's theorem, already substructures, quotients, isomorphisms, and products are enough for obtaining all commutative rings (with unity). Thus, if we want to be able to make such natural constructions, we must consider the whole class of commutative rings. The situation here is similar to Bishop's: given a uniformly continuous function, the classical mathematician concludes that it is continuous and applies his theorem. The constructive mathematician applies his theorem using the information that the function is *uniformly* continuous. In our example, the classical mathematician concludes that if the class should be closed under many natural operations, it has, in particular, to be closed under homomorphic images, subalgebras, and products, hence by Birkhoff's theorem it is axiomatizable by equations. The constructive version we are about to present requires more. It is not sufficient to know only these three closure properties, but if we know also that the class we consider is closed under inductive limits (colimits of systems indexed by directed preorders) and contains a *generic* algebra (in the above-mentioned example the ring of integers), we have enough to conclude that the set is axiomatized by equations. Thus, in particular, the example given above is one where the constructive theorem can be applied.

The proof we will give is constructive and predicative[2)]. It is based on a different idea than the usual proofs: that of generalizing the well-known method of representing rings of polynomials by rings of polynomial functions. In other words, we generalize the observation that $\mathbb{Z}[x_1, \ldots, x_n]$ is isomorphic to a subring of $\mathbb{Z}^{\mathbb{Z}^n}$. The proof we give yields also a simple classical proof of the original theorem by using the classical principle that all equalities are either true or false.

2 Preliminaries

We will reason, in this article, in an informal style, but every step is straightforward to formalize in Martin-Löf's type theory. For details about a possible formalization of subsets, images, and partial functions, see [8]. For details about inductive limits and products, see [9]. Following the tradition of informal mathematics, we will not emphasize intensional aspects like the distinction between the notation $x \in X$ and $x \varepsilon X$ as in [8]. Moreover, we will freely assume that a *subset* is a *set*, although it is more natural in many cases to view it as a propositional function on a set. It is possible to keep the distinction between subsets and sets, but because we work with classes that are closed under isomorphisms, this distinction is not useful.

An *algebra* A is a set $|A|$ equipped with an equivalence relation $=_A$ and some (possibly infinitely many) operations (endo-functions) of finitely many variables, respecting the equality $=_A$. We will keep the signature fixed, which allows us not to mention signatures at all. Thus we assume that the operations are of the same arity for each algebra. In type theory we would define the type of algebras as a record:

$$\begin{aligned}
&X : \mathrm{set}\\
&=_X : (X, X)\mathrm{prop}\\
&f_1 : (X, \ldots, X)X\\
&\vdots\\
&f_m : (X, \ldots, X)X
\end{aligned}$$

together with proofs that $=_X$ is an equivalence relation and that the operations respect this relation. A *class of algebras* will in the following mean a class of algebras of the signature considered, that is, a collection of such algebras singled out by a propositional function, as in [8].

[2)] The word *predicative* is used here in the generalized sense often used in the type theory community, rather than in the sense of Feferman-Schütte. It means roughly that the power set axiom is avoided so that the results can be formalized in Martin-Löf's type theory [7].

A *homomorphism* (or *morphism*) $A \longrightarrow B$ is a function $|A| \longrightarrow |B|$ that respects the equalities and the operations. A *subalgebra of* A is a subset of A that respects $=_A$ and is closed under all operations. A *homomorphic image of* A is the image of a homomorphism from A: the subalgebra $\{y \mid (\exists x \in |A|)(y = f(x))\}$, where the equality is the one that is defined on the codomain. Hence a homomorphic image always respects the equality. If $\{A_i\}_{i \in I}$ is a family of algebras, then ΠA_i is the algebra whose underlying set is $\Pi|A_i|$ and with the equality and the operations defined point-wise. Notice that the operations respect the equality. If the family is constant, we write A^I. In the special case when $I = N_n = \{1, \dots, n\}$, we write A^n. Thus in particular A^0 is a *trivial algebra* consisting of one element only. If $a \in A^n$, we use the notation $a = (a_1, \dots, a_n)$, with $a_i = a(i)$.

If both A and B are algebras, then A^B is the subalgebra of $A^{|B|}$ consisting of the extensional functions, i. e., those that preserve the equalities. They need not be homomorphisms.

An important algebra is *the term algebra*. For every natural number n, let $|T_n|$ be the set of terms generated inductively as follows:

$$\begin{array}{ll} x_i \in |T_n| & (i \in N_n), \\ f_1(t_1, \dots, t_{n_1}) \in |T_n| & (t_1 \in |T_n|, \dots, t_{n_1} \in |T_n|), \\ \vdots & \\ f_m(t_1, \dots, t_{n_m}) \in |T_n| & (t_1 \in |T_n|, \dots, t_{n_m} \in |T_n|). \end{array}$$

Let the equality be the finest equivalence relation and the operations be composition of terms. We denote by x the function from N_n to T_n, taking every number to its corresponding variable symbol. In correspondence with the notation mentioned above, we may write $x = (x_1, \dots, x_n)$ and we have $x \in T_n^n$.

Any function $a \in A^n$ extends, by a recursive definition, to a function $\widehat{a} \in A^{T_n}$ such that $\widehat{a} \circ x =_{A^n} a$. The extension is unique in the sense that if $f \in A^{T_n}$ and $f \circ x =_{A^n} a$, then $f =_{A^n} \widehat{a}$.

A function $a \in A^n$ is said to be a *family of generators* of the image of $\widehat{a}$, which is called *the subalgebra generated by* $a_1, \dots, a_n$. An n-generated algebra is therefore the same as a homomorphic image of T_n. An algebra is *finitely generated* if it is n-generated for some natural number n. Classically it is common to generalize this notion of generated algebras also to arbitrary, not necessarily finite, sets of generators by considering functions $a \in A^I$ for arbitrary sets I. That is not so fruitful constructively, because partial functions cannot in general be extended to total functions, which means that a valuation of the variables of a term need not extend to a function $a \in A^I$. Therefore, we define the subalgebra of A generated by the subset $U \subseteq |A|$ to be the union of all algebras generated by families a with $a_i \in U$. It is straightforward to check that it respects the equality and is closed under the operations. Notice the fact that the subalgebra generated by U is the inductive limit of the subalgebras generated by finitely enumerable subsets of U. This indicates how functions are naturally defined on such subalgebras, a fact that will be used later on.

For any algebra A we shall introduce the notation $\operatorname{Ids} A$ for the *identities of* A, that is, for the equations that are *valid in* A: true in A for *all* substitutions for the variables. For instance, in group theory,

$$x_1 \circ (x_2 \circ x_3) = (x_1 \circ x_2) \circ x_3$$

is an identity for every group, while $x_1 \circ x_2 = x_2 \circ x_1$ is an identity in Abelian groups only. To get an algebraic representation of identities, we will consider them as pairs of terms, rather than as equations. We proceed as follows. For any $a \in A^n$, *the kernel of* $\widehat{a}$ is the subset of $T_n \times T_n$ defined by

$$(s, t) \in \ker \widehat{a} \quad \Leftrightarrow \quad \widehat{a}(s) =_A \widehat{a}(t)$$

and the intersections of all such kernels is called $\operatorname{Ids}_n A$:

$$\operatorname{Ids}_n A \stackrel{\text{def}}{=} \textstyle\bigcap_{a \in A^n} \ker \widehat{a}.$$

$\operatorname{Ids} A$ is now the inductive limit of the sequence $\operatorname{Ids}_n A$. In fact, we will not need this construction in the following, it suffices to consider $\operatorname{Ids}_n A$ for different n. However, it is convenient to write things like "$\operatorname{Ids} A \subseteq \operatorname{Ids} B$" rather than "$\operatorname{Ids}_n A \subseteq \operatorname{Ids}_n B$ for all n". Therefore we use $\operatorname{Ids} A$ as an informal construction, although formally it is in fact not needed.

3 The proof of Birkhoff's theorem

We will base our proof of Birkhoff's theorem on the following lemma.

Lemma 3.1 (Free algebras) $T_n/\mathrm{Ids}_n\, A$ *is isomorphic to a subalgebra of* A^{A^n}.

P r o o f. Let F be the subalgebra of A^{A^n} generated by the projections π_i, which are given by $\pi_i(a) \stackrel{\text{def}}{=} a_i$ for each $i \in N_n$. F is "freely generated over the class $\{A\}$" in the sense that if $a \in A^n$, then there is a unique morphism $a^* : F \longrightarrow A$ such that $a^* \circ \pi =_{A^n} a$:

$$\begin{tikzcd}
N_n \arrow[d, "x"] \arrow[dr, bend left, "a"] \arrow[dd, bend right=60, "\pi"'] & \\
T_n \arrow[r, "\widehat{a}"] \arrow[d, "\widehat{\pi}"] & A. \\
F \arrow[ur, bend right, "a^*"'] &
\end{tikzcd}$$

The uniqueness is obvious because F is generated by the projections, it suffices to prove the existence. But simply let $a^*(f) \stackrel{\text{def}}{=} f(a)$ and we have

$$a^*(\pi_i) \stackrel{\text{def}}{=} \pi_i(a) \stackrel{\text{def}}{=} a_i. \tag{1}$$

The diagram above commutes, because the only triangle left unchecked so far is $a^* \circ \widehat{\pi} =_{A^{T_n}} \widehat{a}$, but this one follows from the uniqueness of $\widehat{a}$ as an extension of a to T_n. Indeed, this property tells us that it is enough to verify that $a^* \circ \widehat{\pi} \circ x =_{A^n} a$ is true, but we checked that already in (1). One also has to check that a^* preserves the operations, a straightforward task.

We shall prove that $\ker \widehat{\pi} = \mathrm{Ids}_n\, A$. We use in one direction

$$\mathrm{Ids}_n\, A \subseteq \mathrm{Ids}_n\, A^{A^n} \subseteq \mathrm{Ids}_n\, F \subseteq \ker \widehat{\pi}$$

and, for the other direction, that

$$\ker \widehat{\pi} \subseteq \ker(a^* \circ \widehat{\pi}) = \ker \widehat{a}$$

for any $a \in A^n$, so that $\ker \widehat{\pi} \subseteq \mathrm{Ids}_n\, A$. □

This lemma has the following immediate corollary.

Corollary 3.2 *If* A, B *are algebras,* B *is* n*-generated, and* $\mathrm{Ids}_n\, A \subseteq \mathrm{Ids}_n\, B$*, then* B *is a homomorphic image of an* n*-generated subalgebra of* A^{A^n}.

P r o o f. In view of the lemma, it is sufficient to prove that B is a homomorphic image of $T_n/\mathrm{Ids}_n\, A$. But B was assumed to be n-generated, which means that B is the image of some $\widehat{b} : T_n \longrightarrow B$. To see that $\widehat{b}$ factors over $T_n/\mathrm{Ids}_n\, A$, it suffices to prove that $\mathrm{Ids}_n\, A \subseteq \ker \widehat{b}$. But $\mathrm{Ids}_n\, A \subseteq \mathrm{Ids}_n\, B \subseteq \ker \widehat{b}$. □

This corollary is sufficient to prove a version of Birkhoff's theorem. This theorem states a fact about axiomatization, a concept that has not been defined yet. We will avoid the introduction of many logical concepts by using the following definition.

Definition 3.3 Let G be an algebra. A class $\mathcal{C}$ is *axiomatized by* $\mathrm{Ids}\, G$ if, for every algebra A,

$$A \in \mathcal{C} \Leftrightarrow \mathrm{Ids}\, G \subseteq \mathrm{Ids}\, A. \tag{2}$$

This relates to the logical notion of axioms in the following way. If a set of equational axioms is given, we can construct a generic algebra G for it as the term algebra generated by countably many variables, but with the equality defined to mean that terms can be proved to be equal from the axiom set with equational logic. Then G is a model of the axioms. The class $\mathcal{C}$ described by the axioms will be "axiomatized by $\mathrm{Ids}\, G$" in the sense of (2).

On the other hand, if (2) holds for some G, then $\mathcal{C}$ is clearly described by the equational axioms we can extract from $\mathrm{Ids}\, G$. Hence "axiomatized by equations" in the logical sense is here faithfully represented as "axiomatized by $\mathrm{Ids}\, G$, for some algebra G". It should be remarked that from the predicative point of view "$\mathcal{C}$ is axiomatized by $\mathrm{Ids}\, G$" is not a proposition, because there is no *set* of proofs of it. Rather, such an utterance expresses that the class $\mathcal{C}$ comes equipped with a function that to every algebra $A \in \mathcal{C}$ gives a proof of $\mathrm{Ids}\, G \subseteq \mathrm{Ids}\, A$.

We generalize the notion of generic algebra in the usual way of Universal Algebra: an algebra $G \in \mathcal{C}$ is said to be *generic for* $\mathcal{C}$ if $\mathrm{Ids}\, G \subseteq \mathrm{Ids}\, A$ is true for every $A \in \mathcal{C}$. It is often easy to find generic algebras: the ring of integers is generic for commutative rings, the field of rationals is generic for fields, the two-element Boolean algebra is generic for Boolean algebras, et cetera.

We can now state and prove the following version of Birkhoff's theorem:

Theorem 3.4 *If $\mathcal{C}$ is a class of algebras that*

1. *is closed under homomorphic images,*
2. *contains precisely those algebras whose finitely-generated subalgebras are in $\mathcal{C}$,*
3. *contains every power of a generic algebra G,*

then $\mathcal{C}$ is axiomatized by $\mathrm{Ids}\, G$.

Proof. We shall prove that if $\mathrm{Ids}\, G \subseteq \mathrm{Ids}\, A$, then $A \in \mathcal{C}$. It suffices to prove that the finitely-generated subalgebras of A are in $\mathcal{C}$. Fix such an algebra B. Because $\mathrm{Ids}\, A \subseteq \mathrm{Ids}\, B$ we have that $\mathrm{Ids}\, G \subseteq \mathrm{Ids}\, B$. Hence, according to Corollary 3.2, B is a homomorphic image of a finitely-generated subalgebra of G^{G^n} for some n. Because $G^{G^n} \in \mathcal{C}$, we are done. $\square$

Although this version is often as powerful as Birkhoff's original theorem, it does not have the same flavour. Birkhoff's theorem is stated in a way that connects algebraic properties of a class on the one hand (closure under certain operations: homomorphic images, subalgebras, products) with logical properties on the other hand (axiomatizability by equations); the version above has more complicated conditions. We therefore perform an analysis of the differences.

The first condition is identical with one of the conditions of Birkhoff's theorem. The second condition is a compact way of stating two conditions: closure under finitely-generated subalgebras and the converse. The converse is hard (at least for the author) to state in a nice algebraic way, but it will turn out to be equivalent to *closure under inductive limits*. To see that, begin by recalling that every algebra is the inductive limit of its finitely-generated subalgebras. On the other hand, consider a directed system of algebras and its limit A. Assume that every member of the system is in $\mathcal{C}$ and that the conditions of Theorem 3.4 are fulfilled. We shall prove that $A \in \mathcal{C}$. It is sufficient to prove that every finitely-generated subalgebra of A is in $\mathcal{C}$. So take such a subalgebra, say it is generated by $a_1, \dots, a_n$. For each $i \in N_n$, take A_i from the system such that the image of A_i in A contains a_i. Because the system is directed, there exists an algebra B in it with morphisms $A_i \longrightarrow B$. Hence the image of B in A has $\{a_1, \dots, a_n\}$ as a subset. Because the image of B is in $\mathcal{C}$, so is the subalgebra generated by $a_1, \dots, a_n$.

Under the condition of closure under inductive limits it is clear that closure under finitely-generated subalgebras implies closure under arbitrary subalgebras, because to prove that a subalgebra is in $\mathcal{C}$ it is sufficient to prove that its finitely-generated subalgebras are in $\mathcal{C}$. Hence closure under subalgebras could be stated for finitely-generated ones only. One could also state closure under inductive limits for systems of finitely-generated algebras only – indeed we used only such systems in the argument above.

The third condition is also a compact way of writing two conditions: one says that $\mathcal{C}$ has a generic algebra and one says that it also has all of its powers. This is at the same time a weakening and a strengthening of Birkhoff's condition of closure under products. If we add closure under products, we may obviously reduce our condition to the requirement of the existence of a generic algebra. Let us see why this requirement can be avoided in an impredicative setting.

First, notice that, assuming we require closure under products, it is enough to require a *generic family* of algebras, that is, a set-indexed family $\{G_i\}$ of algebras of $\mathcal{C}$ such that an equation is an identity of every algebra of $\mathcal{C}$ if it is an identity of every G_i. This is so because if $\{G_i\}$ is a generic family, ΠG_i is a generic algebra. From the impredicative point of view, such families always exist.

Proposition 3.5 (Using powerset and full separation) *Every class that is closed under isomorphisms and subalgebras contains a generic family.*

Proof. We use a variant of the method used to save Birkhoff's original proof. Let K_n be the set of kernels of morphisms of the form $T_n \longrightarrow A$ for algebras $A \in \mathcal{C}$ (that K_n is indeed a set is the impredicative step). We claim that $\{T_n/k\}_{n\in N, k\in K_n}$ is a generic family, with $I \stackrel{\text{def}}{=} (\Sigma n : N)K_n$ as an indexing set. To prove this claim, we have to show two things. First that $T_n/k \in \mathcal{C}$ for every $(n,k) \in I$, and secondly, that the identities of the algebras of this family are identities of all algebras of the class. The first claim is proved via the observation that our assumption of closure under isomorphisms and subalgebras reduces the problem to showing that T_n/k can be injectively embedded into an algebra that is in $\mathcal{C}$. But since k is the kernel of a morphism $T_n \longrightarrow A$, where $A \in \mathcal{C}$, we can choose that morphism. To prove the second claim, we must show that, for every m, if $(s,t) \in \mathrm{Ids}_m(T_n/k)$ for all n, k, then $(s,t) \in \mathrm{Ids}_m A$, for every algebra $A \in \mathcal{C}$, that is, that $(s,t) \in k$ for every $k \in K_m$. So let us assume $(s,t) \in \mathrm{Ids}_m(T_n/k)$ for all n, k and take an arbitrary $k \in K_m$. Because $(s,t) \in \mathrm{Ids}_m(T_m/k)$, it follows that $(s,t) \in k$. □

It seems impossible to prove this in a constructive way, a fact that we do not consider to be a drawback. In practice it is often easy to find generic families. It is more useful to use a well-described generic family than one that is asserted to exist by the previous proposition, because our version of Birkhoff's theorem tells us what the axioms are in terms of the generic family. In lucky cases, as the one of commutative rings, and the one of Boolean algebras, there is a simple decision procedure for testing identities of the generic family.

The analysis above lets us restate the theorem in a way that reminds more about Birkhoff's original one, because it relates closure properties to logical properties. However, we remark that there are at least three reasons to favour Theorem 3.4 as it was stated:

1) the conditions are often easier to check in applications,
2) they state exactly that which is used in the proof,
3) it is easier to see that they follow from the conditions of the following theorem than conversely.

The following theorem is, on the other hand, cleaner and easier to remember. The condition $(+)$ here is dealt with in the way we always use in predicative mathematics: we cannot state it as a proposition quantifying over all inductive systems, but we have to interpret it as follows: the class comes equipped with a function that given any directed system of algebras of the class gives a proof that the inductive limit is in the class.

Theorem 3.6 *If a class $\mathcal{C}$ contains a generic family $\{G_i\}$ and is closed under*

$(\div)$ *homomorphic images,*

$(-)$ *subalgebras,*

$(+)$ *inductive limits,*

$(\times)$ *products,*

then it is axiomatized by $\mathrm{Ids}(\Pi G_i)$.

Notice that, conversely, if a class $\mathcal{C}$ is axiomatized by $\mathrm{Ids}\, G$, it satisfies the conditions of the theorem. So the conditions are necessary in the logical sense.

4 About the use of inductive limits

The assumption of closure under inductive limits was used above to reduce the proof that an algebra is in $\mathcal{C}$ to proving that every finitely-generated subalgebra of it is. That was easier because Corollary 3.2 could then be applied. So the first question is if this corollary, or something similar, can be proved for arbitrary, not necessarily finitely-generated, algebras. Because the power mentioned in the corollary depends on n, it would be necessary to modify also this power. The natural modification is to replace N_n by an arbitrary set I in the following way, which however is not provable constructively.

Proposition 4.1 (Classical) *If A, B are algebras, B generated by the image of a set I, and* $\mathrm{Ids}\, A \subseteq \mathrm{Ids}\, B$, *then B is a homomorphic image of a subalgebra of A^{A^I}.*

Classical proof. Let the function $\pi : I \longrightarrow A^{A^I}$ be given by $\pi_i(a) \stackrel{\text{def}}{=} a(i)$. Let F be the subalgebra generated by the image of π. By definition F is then the union of the finitely-generated subalgebras given by generators of the form π_i, for $i \in I$. Hence, to construct a morphism $F \longrightarrow B$ it suffices to construct a compatible family of morphisms from those subalgebras. This is done as in the proof of Corollary 3.2. □

The non-constructive step is hidden in the last sentence. One has to use that the finitely-generated algebras that build up F are isomorphic to those algebras that were called F in the proof of Corollary 3.2. But this is not always true in models of constructive mathematics where all real endo-functions are continuous. Indeed, the function $\pi : I \longrightarrow A^{A^I}$ is constant in such models if A is finite and I is the set of real numbers. Hence the algebra F will be 1-generated in this example. Moreover, A^{A^I} will be finite, so the proposition itself, not only its proof, fails in such models. Classically, this proposition immediately gives the classical version of Birkhoff's theorem.

We observe that the proof of Proposition 4.1 works classically because it can be *proved* classically that F is the inductive limit of the system of finitely-generated algebras that we considered before. This explains why it is possible classically to avoid the assumption that the class is closed under inductive limits. Constructively, this way is closed, since F is too small in some models to be an inductive limit of that system.

However, all that is needed to make the proof of Proposition 4.1 constructive is to replace the set I with a discrete setoid I (a set with an equivalence relation that, for all arguments, is either true or false). The algebra A^I is then the subalgebra of $A^{|I|}$ consisting of the extensional (equality-preserving) functions. Hence the following version is constructive. In particular, it can be used when B is discrete, or (sub)countable, or (sub)countably generated.

Proposition 4.2 *If A, B are algebras, B generated by the image of a discrete setoid I, and* $\operatorname{Ids} A \subseteq \operatorname{Ids} B$, *then B is a homomorphic image of a subalgebra of A^{A^I}.*

P r o o f. We shall repair the proof of Proposition 4.1 by proving that every subalgebra which is generated by $(\pi_{i_1}, \dots, \pi_{i_n})$ for some $i = (i_1, \dots, i_n) \in |I|^n$ is isomorphic to some algebra of the form $T_m/\operatorname{Ids}_m A$. We begin by studying the case $n = 0$. There are unique morphisms $T_0 \longrightarrow A$ and $T_0 \longrightarrow A^{A^I}$. The images of these morphisms are isomorphic, because the image of the latter consists of constant functions. Hence the kernels are both equal to $\operatorname{Ids}_0 A$.

Now consider the case $n \geq 1$. Because every finitely enumerable subsetoid of I is finite, we may assume also that $i_1, \dots, i_n$ are distinct so that there is a left inverse j to i. For every $a \in A^n$, let

$$(a \circ j)^*(f) \stackrel{\text{def}}{=} f(a \circ j).$$

Then $(a \circ j)^* \circ \widehat{\pi \circ i} = \widehat{a}$ by the uniqueness of $\widehat{a}$ as an extension of a to T_n:

$$((a \circ j)^* \circ \widehat{\pi \circ i} \circ x)(m) = \pi_{i_m}(a \circ j) = (a \circ j)(i_m) = a_m.$$

It is now clear that the following diagram commutes:

$$\begin{tikzcd}
& N_n \arrow[dl, bend right, "i"'] \arrow[d, "x"] \arrow[dr, bend left, "a"] & \\
\operatorname{im} i \arrow[ur, "j"'] \arrow[dr, bend right, "\pi"'] & T_n \arrow[r, "\widehat{a}"] \arrow[d, "\widehat{\pi\circ i}"] & A. \\
& A^{A^I} \arrow[ur, bend right, "(a\circ j)^*"'] &
\end{tikzcd}$$

Hence $\ker \widehat{\pi \circ i} \subseteq \ker \widehat{a}$ for every $a \in A^n$. It follows that $\ker \widehat{\pi \circ i} \subseteq \operatorname{Ids}_n A$. □

It is reasonable to view Theorems 3.4 and 3.6 as natural constructive counterparts to Birkhoff's classical theorem, in particular because the two conditions we had to add are typical: one takes care of the predicativity part by assuring that there is a *set* of equations that form an axiom system, the other one is explicitly stating closure under some colimits where this could be proved to follow from closure under some limits in classical mathematics (much like Σ-types can be defined in terms of Π-types in impredicative systems). It seems not to be trivial to prove classically that closure under inductive limits follows from the other closure conditions (the simplest proof seems to be via Birkhoff's theorem), but it is trivial to see that our proof would not need this condition in a classical setting, because Proposition 4.2 could then be applied in all cases.

Acknowledgements I would like to thank Thierry Coquand for suggestions leading to considerable simplifications of the proof, and Thomas Streicher, Michael Rathjen, and Benno van den Berg for interesting discussions about possible models where powers of N, but not all sets, are subcountable. I also thank the referees for helpful comments.

References

[1] G. Birkhoff, On the structure of abstract algebras. Proc. Cambridge Phil. Soc. **31**, 433 – 454 (1935).
[2] S. Burris and H. P. Sankappanavar, A Course in Universal Algebra (Springer, 1981).
[3] G. Grätzer, Universal Algebra, second edition (Springer, 1979).
[4] E. Bishop, Foundations of Constructive Analysis (McGraw-Hill Book Co., 1967).
[5] T. Streicher, Realizability. Lecture notes (2004/05). Available at `http://www.mathematik.tu-darmstadt.de/~streicher/`.
[6] A. Bauer, The realizability approach to computable analysis and topology. Ph. D. thesis, School of Computer Science, Carnegie Mellon University (2000). Available at `http://andrej.com/thesis/`.
[7] B. Nordström, K. Petersson, and J. Smith, Programming in Martin-Löf's Type Theory (Oxford University Press, 1990). Available at `http://www.cs.chalmers.se/Cs/Research/Logic/book/`.
[8] J. Carlström, Subsets, quotients and partial functions in Martin-Löf's type theory. In: Types for Proofs and Programs, Second International Workshop, TYPES 2002, Berg en Dal, The Netherlands, April 24 – 28, 2002, Selected Papers (H. Geuvers and F. Wiedijk, eds.). Lecture Notes in Computer Science 2646, pp. 78 – 94 (Springer, 2003).
[9] J. Carlström, Formalized limits and colimits of setoids. Technical Report 9, Department of Mathematics, Stockholm University (2003). Available at `http://www.math.su.se/~jesper/research/limitsinsetoid/`.

