Minimum bases for equational theories of groups and rings: The work of Alfred Tarski and Thomas Green

George F. McNulty

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

Abstract

Suppose that T is an equational theory of groups or of rings. If T is finitely axiomatizable, then there is a least number μ so that T can be axiomatized by μ equations. This μ can depend on the operation symbols that occur in T. In the 1960's, Alfred Tarski and Thomas C. Green completely determined the values of μ for arbitrary equational theories of groups and of rings. While Tarski and Green announced the results of their collaboration in 1970, the only fuller publication of their work occurred as part of a seminar led by Tarski at Berkeley during the 1968-69 academic year. The present paper gives a full account of their findings and their proofs.

Key words: Equational logic, equational bases, groups, rings
2000 MSC: 03C05, 08B05

1 Equational logic and equational theories of algebras

Equational logic can be viewed as that fragment of first-order logic in which the only sentences are universal sentences whose quantifier-free part is an equation between terms. The familiar distributive law $\forall x \forall y \forall z [x \cdot (y + z) \approx x \cdot y + x \cdot z]$ is an example of such a sentence. Equational logic has no logical connectives, no relation symbols apart from the logical equality symbol $\approx$, and its sole quantifier is the universal quantifier which plays such a restrained role that it is usually surpressed—the distributive law, for example, is expressed simply as $x \cdot (y + z) \approx x \cdot y + x \cdot z$. In comparison to first-order logic, equational logic is equipped with an apparently meager means of expression. Still, many classes of algebras that have found important places in mathematics can be specified by means of equations and certainly reasoning about equations is ubiquitous.

Equational logic can be developed as a formal system in its own right. Garrett Birkhoff (1935) proved a completeness theorem for equational logic using a system of simple rules of inference referring only to equations. He also proved that the classes of algebras axiomatized by equations are exactly those which are closed with respect to the formation of homomorphic images, subalgebras, and arbitrary direct products—the earliest preservation theorem.

Two formalisms for equational logic can differ only in their operation symbols. While for more general considerations arbitrary systems of operation symbols are appropriate, in this paper we restrict our attention to those equational formalisms provided with only systems of finitely many operation symbols. Suppose a formalism has been specified by selecting a system of operation symbols. We will say that a set T of equations is an **equational theory** if and only if it is closed under logical consequence. A set Σ of equations is a **base** for T provided T is the set of all logical consequences of Σ. Thus Σ is a set of equational axioms for T. We say that the equational theory T if **finitely based** if it has a finite base. A set Σ of equations is **irredundant** if and only if Σ is not logically equivalent to any of its proper subsets.

In practice, equational theories arise in two ways: as the set of consequences of some particular set Σ of equations, and as the set of all equations true in all the algebras belonging to some class $\mathcal{K}$ of algebras. Ring theory probably arose in the first way while the theory of groups arose in the second way with $\mathcal{K}$ being the class of all groups of permutations. These two ways in which equational theories ordinarily arise correspond to two purposes to which equational bases are put: to provide a basis upon which to construct proofs, and to provide a means to determine whether an algebra belongs to a particular class $\mathcal{K}$ of algebras. The work of Tarski and Green, which is at the heart of this paper, concerned a third purpose.

In first order logic, any finitely axiomatizable theory can be axiomatized by a single sentence. Because equational logic lacks connectives, many finitely based equational theories fail to be based a just a single equation. For an equational theory T we let μT be the least among all cardinals κ so that T has a base of cardinality κ. This parameter μT offers a means to differentiate among equational theories. Alfred Tarski and Thomas Green took on the task of determining μT in the cases when T is either a theory of rings or a theory of groups.

2 Term equivalence of equational theories

The theory of groups has been formalized as an equational theory using a number of systems of operation symbols. Certain of these formalizations differ

from others only cosmetically. For example, one could be formalized using a binary operation symbol $\cdot$ and a unary operation symbol $^{-1}$ while another formalization might use a binary operation symbols $*$ and a unary operation symbol $^{-}$. While the equational theories arising from these two formulations are certainly different, the difference carries no mathematically interesting information. Such theories are said to be **literally similar**. More interesting are other equational formalizations. For example, Marshall Hall (1976) provides two formalizations of group theory using different systems of operation symbols—the first formalization uses the customary symbols $\cdot, ^{-1}$ and 1 while the second uses only one symbol / to stand for division. In this paper we use symbol $\dashv$ in place of /. The connection between these two formulations of the theory of groups can be described as follows. Let T_0 be the equational theory of groups in the first formulation. Following Hall, this theory is based on

$$\{x \cdot (y \cdot z) \approx (x \cdot y) \cdot z, 1 \cdot x \approx x, x^{-1} \cdot x \approx 1\}.$$

Let T_1 be the equational theory of groups in the second formulation. Again following Hall (but using $\dashv$ in place of /), this theory is based on

$$\{x \dashv x \approx y \dashv y, x \dashv (y \dashv y) \approx x, (x \dashv x) \dashv (y \dashv z) \approx z \dashv y, (x \dashv y) \dashv (z \dashv y) \approx x \dashv z\}.$$

Let T_2 be the equational theory based on

$$T_0 \cup \{x \dashv y \approx x \cdot y^{-1}\}.$$

The theory T_2 is a **definitional extension** of T_0 because it is obtained from T_0 by adding equations which define the new operation symbol(s) by means of term(s) built up from variables and the operation symbols of T_0. It can be proved that T_2 is also based on

$$T_1 \cup \{1 \approx x \dashv x, x^{-1} \approx (x \dashv x) \dashv x, x \cdot y \approx x \dashv ((y \dashv y) \dashv y)\}.$$

So T_0 and T_1 have a common definitional extension. An important technical point illustrated by the second base for T_2 is that the term $x \dashv x$ involves more variables than 1, the term it defines. In order for this to be legitimate we require that $x \dashv x \approx y \dashv y \in T_1$. In other words, for T_2 to be a definitional extension of T_1 it is necessary that T_1 contain equations which assert that the terms used in the definitions do not depend on such surplus variables.

We say that equational theories T_0 and T_1 are **term equivalent** if and only if there are equational theories T_0', T_1' and T_2' so that T_0 is literally similar to T_0', that T_2' is a common definitional extension of T_0' and T_1', and that T_1' is literally similar to T_1. Tarski referred to this concept as *definitional equivalence* and Maltsev used the phrase *rational equivalence*. Term equivalence can also be developed from a connection between the models of T_0 and the models of T_1, see for example (McKenzie et al., 1987).

Suppose that T_2 is a definitional extension of T_1. The definitions of the new operation symbols offer a means to eliminate the new symbols from the terms of T_2 to obtain equivalent terms of T_1. Here are the details for the case given above. The elimination map λ from the set of terms appropriate for T_2 to terms appropriate for T_1 is defined by the following recursion, where x is any variable and s and t are any terms appropriate for T_2:

$$\begin{aligned}
\lambda(x) &= x \\
\lambda(1) &= x \dashv x \\
\lambda(t^{-1}) &= (\lambda(t) \dashv \lambda(t)) \dashv \lambda(t) \\
\lambda(s \cdot t) &= \lambda(s) \dashv ((\lambda(t) \dashv \lambda(t)) \dashv \lambda(t)) \\
\lambda(s \dashv t) &= \lambda(s) \dashv \lambda(t)
\end{aligned}$$

It is not hard to prove that if Σ is a base for T_2 then $\{\lambda(s) \approx \lambda(t) \mid s \approx t \in \Sigma\}$ is a based for T_1. The same applies to any pair of theories such that one is a definitional extension of the other.

3 Tarski and Irredundant Bases of Equational Theories

Alfred Tarski (1938) considered the equational theory of Abelian groups construed as algebras with a single operation $\dashv$ to stand for right subtraction. That is $a \dashv b = a + (-b)$ where $+$ and $-$ represent the customary Abelian group operations of addition and additive inversion (negation). He demonstrated that the equational theory of Abelian groups so construed can, in fact, based on some one equation. Higman and Neumann (1952) extended this result to all finitely based equational theories of groups where $\dashv$ stands for right division, i.e. $a \dashv b = ab^{-1}$ using the more customary operations. In that paper, Higman and Neumann raise the problem of discovering the cardinalities of all irredundant bases of a finitely based equational theory. As Higman and Neumann point out, all these cardinalities must be finite (a simple consequence of the Compactness Theorem). But they noted that these cardinalities may have no finite upper bound. They also observed that these problems were then open for the variety of groups and the variety of Abelian groups, among others. They also note that there seemed to be no example of an infinite irredundant set of equations using just a single operation which is binary. In a footnote, Higman and Neumann say that this last problem had been solved recently. This is probably a reference to the work of Jan Kalicki (1955) carried out under Tarski's influence at Berkeley.

Tarski and his collaborators made decisive inroads on the problems raised by Higman and Neumann. Their results are spelled out in Tarski (1968) and in the two abstracts of Green and Tarski (1970a,b). Proofs for two of these results were later published, but proofs for the remaining results have not been

published in the ensuing thirty years. During the late 1960's Thomas C. Green pursued a Ph.D. under Tarski's supervision, but left Berkeley (and apparently mathematics) before a final version of their joint work was completed and submitted for publication. It seems to me appropriate for this volume to finally place these proofs in the literature.

My knowledge of these proofs stems from notes from a seminar conducted by Tarski at Berkeley during the 1968-1969 academic year. I was a participant in that seminar, as were Fred Backer, Gary Cooper, Steven Givant, Thomas C. Green, Joel Karnofsky, Michael Kwatinetz, Charles Martin, Don Pigozzi, Kan Ching Ng, William Wadge, and Benjamin F. Wells. My role is expositor—neither the theorems nor any of the proofs given here should be credited to me. Any errors below are mine.

Theorem 0 (Tarski) *If T is a finitely based equational theory with irredundant bases of cardinalities m and n, and k is a natural number such that $m < k < n$, then T has an irredundant base of cardinality k.*

Tarski (1975) gives a proof of this theorem has a consequence of a more general result about certain kinds of closure operators. In the same issue of Discrete Mathematics one can find the related papers McNulty and Taylor (1975) and Givant (1975).

Tarski introduced ∇T to stand for the set of cardinalities of irredundant bases of the equational theory T. As a consequence of the theorem above ∇T must either be empty, a finite interval of natural numbers, an infinite interval of natural numbers, or the set $\{\omega\}$. Equational theories of the first kind and of the last kind cannot be finitely based, while those of the other kinds are finitely based. In this way, to any finitely based equational theory T one can associate a pair $m, n \in \mathbb{N} \cup \{\infty\}$ of parameters so that $\nabla T = [m, n)$. As described in (Tarski, 1968), all the possibilities for ∇T left open by the theorem above have actually been realized by examples, using just one binary operation symbol, constructed by Tarski, Judith Ng, and Ralph McKenzie.

Theorem 1 (Tarski) *Let T be a finitely based equational theory such that $t \approx x \in T$ for some term t in which the variable x occurs at least twice. Under these assumptions ∇T is infinite.*

Tarski never published a proof of this theorem. A proof of the following modest extension can be found in (McNulty, 1976).

Theorem 2 *Let T be a finitely based equational theory such that $t \approx x \in T$ for some term t in which some operation symbol of rank at least two occurs or else in which two different unary operation symbols occur. Under these assumptions ∇T is infinite.*

Consequently, the most commonly encountered finitely based equational theories will have irredundant bases of all large enough finite cardinalities. For such equational theories determining the least among the cardinalities of irredundant bases becomes a natural question. Tarski used μT to stand for this least cardinality.

Higman and Neumann had already determined that $\mu T = 1$ for any finitely based theory T of groups construed using only the operation symbol $-$ standing for right division. During the last half of the 1960's Tarski undertook the determination of μT for equational theories T of groups or of rings subject only to the restriction that the number of basic operation symbols should be finite. Green joined in this enterprise. The results of their efforts are definitive. Some of these findings are due to Tarski, some to Green, and some to both jointly. I have tried to attribute the results appropriately, but the notes I have leave some of attributions in doubt.

Here are the findings about groups. This is joint work of Tarski and Green.

Theorem 3 *Let T be a finitely based theory of groups in which exactly n different operation symbols appear and such that T has a model with more than one element. Then $\mu T = \max\{1, n-1\}$.*

The notion of ring used here is distinct from the notion of ring with unit in that the latter has a distinguished constant denoting a unit element, while rings of the former sort may have no element playing the role of a multiplicative unit. An equational theory T of rings is said to be of the **first kind** provided T has a model with more than one element in which the multiplication of any elements results in the additive unit 0; if T is not of the first kind it is said to be of the **second kind**. Here are the findings about rings.

Theorem 4 *Let T be a finitely based equational theory of rings with exactly n operation symbols which has a model with more than one element.*

(i) *If T is a theory of the first kind, then $\mu T = \max\{1, n-1\}$.*
(ii) *If T is a theory of the second kind, then $\mu T = 1$.*
(iii) *If T is a theory of rings with unit, then $\mu T = 1$.*

Theorem 4 (i) and (iii) were established by Tarski with part (iii) done independently and by different means by Ralph McKenzie (see the abstract (Grätzer and McKenzie, 1967) and the paper (Grätzer and Padmanabhan, 1978)). Theorem 4 (ii) is due to Green.

These results about groups and rings are particular cases of more general results proven by Tarski and Green, as described in the next section.

Ralph McKenzie (1970) proved that if T is a finitely based equational theory

of lattices, then $\mu T = 1$ if T is the theory of all lattices or if T is the theory of one element lattices and in all other cases at least two equations are needed—Padmanabhan (1969c) had proven that two equations suffice. Subsequently, Padmanabhan and his collaborators published a series of papers concerning μT. For example, (Padmanabhan and Quackenbush, 1973) proved that if T is finitely based and every model of T is both congruence distributive and congruence permutable, then $\mu T = 1$. Further reference to this body of work can be found in the bibliography.

4 Results from the 1968 Seminar

To establish the theorems about minimum bases for equational theories of groups and rings requires proving two sorts of things: the existence of bases of the requisite sizes and the nonexistence of any smaller bases. We take up the existence proofs first.

As in the earlier works (Tarski, 1938) and (Higman and Neumann, 1952), an analysis of the operation $\dashv$ of right division (alias right subtraction) is the point of departure. The key result is the following theorem.

Theorem 5 (Tarski) *Let T be an equational theory based on the finite set Σ of equations and suppose T contains the following equation:*

$$y \approx [(z \dashv z) \dashv (x \dashv y)] \dashv [(w \dashv w) \dashv x]. \qquad (\varepsilon)$$

Assume that there is a nonempty set $\Gamma \subseteq T$ such that each equation in Σ has a substitution instance which is a logical consequence of $\Gamma \cup \{\varepsilon\}$. Under these assumptions, T has a base with no more than $|\Gamma|$ equations.

The proof of this theorem relies on the four lemmas which follow.

Lemma 6 (The Cancellation Lemma) *Let p, s, and t be any terms. The following cancellation laws hold:*

(a) $\varepsilon,\ p \dashv s \approx p \dashv t \vdash s \approx t$
(b) $\varepsilon,\ s \dashv p \approx t \dashv p \vdash s \approx t.$

PROOF. Here is the reasoning to establish (a.):

$s \approx [(z \dashv z) \dashv (p \dashv s)] \dashv [(w \dashv w) \dashv p]$	a substitution instance of ε
$\approx [(z \dashv z) \dashv (p \dashv t)] \dashv [(w \dashv w) \dashv p]$	since $p \dashv s \approx p \dashv t$
$\approx t$	a substitution instance of ε

To establish (b.) observe that the next two equations are substitution instances of ε.

$$[(z \dashv z) \dashv (s \dashv p)] \dashv [(w \dashv w) \dashv s] \approx p$$
$$p \approx [(z \dashv z) \dashv (t \dashv p)] \dashv [(w \dashv w) \dashv t]$$

Consequently

$$[(z \dashv z) \dashv (s \dashv p)] \dashv [(w \dashv w) \dashv s] \approx [(z \dashv z) \dashv (t \dashv p)] \dashv [(w \dashv w) \dashv t].$$

But in view of $s \dashv p \approx t \dashv p$ we obtain

$$[(z \dashv z) \dashv (t \dashv p)] \dashv [(w \dashv w) \dashv s] \approx [(z \dashv z) \dashv (t \dashv p)] \dashv [(w \dashv w) \dashv t].$$

Now two applications of the cancellation law (a.) give first

$$(w \dashv w) \dashv s \approx (w \dashv w) \dashv t$$

and then the desired result

$$s \approx t$$

□

Lemma 7 *The equations* $(z \dashv z) \dashv ((z \dashv z) \dashv x) \approx x, y \dashv y \approx z \dashv z$, *and* $x \dashv (z \dashv z) \approx x$ *are logical consequences of the equation* ε.

PROOF. This short proof relies on three different substituion instances of ε. The first is

$$x \approx [[(x \dashv x) \dashv (x \dashv x)] \dashv [(w \dashv w) \dashv x]] \dashv [(w \dashv w) \dashv (w \dashv w)]$$

which arises from ε by the substitution

$x \mapsto w \dashv w$
$y \mapsto x$
$z \mapsto x \dashv x.$

The second substitution instance is

$$x \approx [(x \dashv x) \dashv (x \dashv x)] \dashv [(w \dashv w) \dashv x] \qquad (*)$$

which arises from ε by the substitution

$y \mapsto x$
$z \mapsto x.$

The right side of the second substitution instance occurs in the first substitu-

tion instance. Hence,

$$x \approx x \dashv [(w \dashv w) \dashv (w \dashv w)]. \qquad (**)$$

The third substitution instance of ε we need is

$$y \dashv y \approx [(z \dashv z) \dashv ((y \dashv y) \dashv (y \dashv y))] \dashv [(y \dashv y) \dashv (y \dashv y)]$$

which arises from ε by the substitution

$$x \mapsto y \dashv y$$
$$y \mapsto y \dashv y$$
$$w \mapsto y$$

Now apply $(**)$ twice to this last substitution instance to obtain first

$$y \dashv y \approx (z \dashv z) \dashv [(y \dashv y) \dashv (y \dashv y)]$$

and then

$$y \dashv y \approx z \dashv z$$

The other two equations we need to prove follow from the last equation applied to $(*)$ and $(**)$. $\square$

Lemma 8 *Let s and t be any terms. The set $\{\varepsilon, s \approx t\}$ is logically equivalent with $\{\varepsilon, s \dashv t \approx z \dashv z\}$.*

PROOF. According to Lemma 7, we have

$$\varepsilon \vdash s \dashv s \approx z \dashv z.$$

Consequently,

$$\varepsilon,\ s \approx t \vdash s \dashv t \approx z \dashv z.$$

For the reverse direction, observe that

$$\varepsilon,\ s \dashv t \approx z \dashv z \vdash s \dashv t \approx t \dashv t.$$

Therefore, by the Cancellation Lemma

$$\varepsilon,\ s \dashv t \approx z \dashv z \vdash s \approx t$$

as desired. $\square$

Now for any terms s and t we will let $\delta_{s,t}$ stand for the following equation

$$y \approx [(z \dashv z) \dashv (x \dashv y)] \dashv [(s \dashv t) \dashv x].$$

Lemma 9 *Let s and t be any terms. The equation $\delta_{s,t}$ is logically equivalent to the set $\{\varepsilon, s \approx t\}$ of equations.*

PROOF. We assume, without loss of generality, that the variables x, y, z, and w do not occur in $s \approx t$. Evidently, $\varepsilon,\ s \approx t \vdash \delta_{s,t}$, so it remains to establish $\delta_{s,t} \vdash \varepsilon$ and $\delta_{s,t} \vdash s \approx t$. These derivations are accomplished at once by an argument like the proof of Lemma 7. We need three substitution instances of $\delta_{s,t}$ (rather than of ε). The first two are

$$z \approx [[(z \dashv z) \dashv (z \dashv z)] \dashv [(s \dashv t) \dashv z]] \dashv [(s \dashv t) \dashv (s \dashv t)]$$
$$z \approx [(z \dashv z) \dashv (z \dashv z)] \dashv [(s \dashv t) \dashv z]$$

The right side of the second equation occurs in the first equation, giving

$$z \approx z \dashv [(s \dashv t) \dashv (s \dashv t)]$$

Substitute $z \dashv z$ for z to obtain

$$z \dashv z \approx (z \dashv z) \dashv [(s \dashv t) \dashv (s \dashv t)]. \qquad (***)$$

The third substitution instance of $\delta_{s,t}$ is

$$s \dashv t \approx [(z \dashv z) \dashv [(s \dashv t) \dashv (s \dashv t)]] \dashv [(s \dashv t) \dashv (s \dashv t)]$$

Applying $(***)$ twice to this equation we obtain first

$$s \dashv t \approx (z \dashv z) \dashv [(s \dashv t) \dashv (s \dashv t)]$$

and then

$$s \dashv t \approx z \dashv z.$$

But z does not occur in $s \dashv t$, so substituting w for z gives

$$s \dashv t \approx w \dashv w.$$

Now replace $s \dashv t$ in $\delta_{s,t}$ by $w \dashv w$ to obtain ε. With ε in hand, the last equation displayed above yields $s \approx t$ by Lemma 8. $\square$

This lays the groundwork for the proof of Theorem 5.

PROOF. [Proof of Theorem 5] Let $\Sigma = \{u_0 \approx r_0, u_1 \approx r_1, \ldots, u_{m-1} \approx r_{m-1}\}$. For each $i < m$, let $u_i^* \approx r_i^*$ denote a substitution instance of $u_i \approx r_i$ such that $\Gamma \cup \{\varepsilon\} \vdash u_i^* \approx r_i^*$. We suppose, without loss of generality, that no variable that occurs in any one equation in Σ or in Γ or in $\{u_i^* \approx r_i^* \mid i < m\}$ or in ε

occurs also in any of the other equations. Pick $p \approx q \in \Gamma$ and let s be

$$p \dashv \Big([u_0 \dashv (\cdots \dashv (u_{m-2} \dashv u_{m-1}) \cdots)] \dashv [u_0^* \dashv (\cdots \dashv (u_{m-2}^* \dashv u_{m-1}^*) \cdots)]\Big)$$

and let t be

$$q \dashv \Big([r_0 \dashv (\cdots \dashv (r_{m-2} \dashv r_{m-1}) \cdots)] \dashv [r_0^* \dashv (\cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)]\Big)$$

Now let $\Delta = (\Gamma - \{p \approx q\}) \cup \{\delta_{s,t}\}$. The set Δ turns out to be a base, of the desired cardinality, for the equational theory T. It is evident that $|\Delta| \leq |\Gamma|$ and that $\Delta \subseteq T$. To complete the proof, it remains only to establish that $\Delta \vdash \Sigma$.

Because $\delta_{s,t} \in \Delta$, it follows from Lemma 9 that $\Delta \vdash \varepsilon,\ s \approx t$. So after the appropriate substitutions, we see that the following equation is a consequence of Δ.

$$\begin{aligned} p \dashv \Big([u_0^* \dashv (\cdots \dashv (u_{m-2}^* \dashv u_{m-1}^*) \cdots)] \dashv [u_0^* \dashv (\cdots \dashv (u_{m-2}^* \dashv u_{m-1}^*) \cdots)]\Big) \approx \\ q \dashv \Big([r_0^* \dashv (\cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)] \dashv [r_0^* \dashv (\cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)]\Big) \end{aligned}$$

Now, in view of Lemma 7 and the presence of ε we obtain

$$p \dashv (z \dashv z) \approx q \dashv (z \dashv z).$$

Therefore, according to the Cancellation Lemma, we have $\Delta \vdash p \approx q$ and also

$$\begin{aligned} \Big([u_0 \dashv (\cdots \dashv (u_{m-2} \dashv u_{m-1}) \cdots)] \dashv [u_0^* \dashv (\cdots \dashv (u_{m-2}^* \dashv u_{m-1}^*) \cdots)]\Big) \approx \\ \Big([r_0 \dashv (\cdots \dashv (r_{m-2} \dashv r_{m-1}) \cdots)] \dashv [r_0^* \dashv (\cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)]\Big) \end{aligned}$$

by cancellation in $\delta_{s,t}$. Now observe that $\Delta \vdash \Gamma$. Since $\Gamma \vdash u_i^* \approx r_i^*$ for all $i < m$, we find that the next equation is also a consequence of Δ.

$$\begin{aligned} \Big([u_0 \dashv (\cdots \dashv (u_{m-2} \dashv u_{m-1}) \cdots)] \dashv [r_0^* \dashv (\cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)]\Big) \approx \\ \Big([r_0 \dashv (\cdots \dashv (r_{m-2} \dashv r_{m-1}) \cdots)] \dashv [r_0^* \dashv (\cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)]\Big) \end{aligned}$$

Applying the Cancellation Lemma yet again, we arrive at the next consequence of Δ.

$$u_0 \dashv (u_1 \dashv \cdots \dashv (u_{m-2} \dashv u_{m-1}) \cdots) \approx r_0 \dashv (r_1 \dashv \cdots \dashv (r_{m-2} \dashv r_{m-1}) \cdots). \quad (\star)$$

Now substitution gives

$$u_0 \dashv (u_1^* \dashv \cdots \dashv (u_{m-2}^* \dashv u_{m-1}^*) \cdots) \approx r_0 \dashv (r_1^* \dashv \cdots \dashv (r_{m-2}^* \dashv r_{m-1}^*) \cdots)$$

but $\Gamma \vdash u_i^* \approx r_i^*$ so we get

$$u_0 \leftharpoonup (u_1^* \leftharpoonup \cdots \leftharpoonup (u_{m-2}^* \leftharpoonup u_{m-1}^*)\cdots) \approx r_0 \leftharpoonup (u_1^* \leftharpoonup \cdots \leftharpoonup (u_{m-2}^* \leftharpoonup u_{m-1}^*)\cdots)$$

The Cancellation Lemma applied to the equation above gives

$$u_0 \approx r_0$$

and, from $(\star)$, we get

$$u_1 \leftharpoonup \cdots \leftharpoonup (u_{m-2} \leftharpoonup u_{m-1})\cdots) \approx r_1 \leftharpoonup \cdots \leftharpoonup (r_{m-2} \leftharpoonup r_{m-1})\cdots).$$

We can repeat this process to obtain $\Delta \vdash u_i \approx r_i$ for all $i < m$. But this means that $\Delta \vdash \Sigma$, which is what was to be proved. $\square$

Corollary 10 (Tarski) *Let T be term equivalent to a finitely based equational theory with $\leftharpoonup, \cdot$, and 1 among its operation symbols to which that following equations belong:*

$$[(z \leftharpoonup z) \leftharpoonup (x \leftharpoonup y)] \leftharpoonup [(w \leftharpoonup w) \leftharpoonup x] \approx y$$
$$x \cdot (z \leftharpoonup z) \approx z \leftharpoonup z$$
$$x \cdot 1 \approx x$$

Under these assumptions, T is one-based.

PROOF. At first we suppose that T includes $\leftharpoonup, \cdot$, and 1 among its operation symbols. So the equations listed in the theorem actually belong to T. Let $\{p_0 \approx q_0, \ldots, p_{m-1} \approx q_{m-1}\}$ be a base for T. Let Σ be

$$\{p_0 \cdot z_0 \approx q_0 \cdot z_0, \ldots, p_{m-1} \cdot z_{m-1} \approx q_{m-1} \cdot z_{m-1}\} \cup \{x \cdot 1 \approx x, \varepsilon\}$$

where none of the distinct variables $z_0, \ldots, z_{m-1}$ occur in any of the $p_i \approx q_i$. The set Σ is a base for T. Take Γ to be $\{(x \cdot 1) \leftharpoonup x \approx (y \cdot (w \leftharpoonup w)) \leftharpoonup (z \cdot (w \leftharpoonup w))\}$. Evidently, $\Gamma \subseteq T$.

The set $\Gamma \cup \{\varepsilon\}$ has the following logical consequences.

$x \cdot 1 \approx x$
By substituting x for y, z, and w in $(x \cdot 1) \leftharpoonup x \approx (y \cdot (w \leftharpoonup w)) \leftharpoonup (z \cdot (w \leftharpoonup w))$ we obtain $(x \cdot 1) \leftharpoonup x \approx (x \cdot (x \leftharpoonup x)) \leftharpoonup (x \cdot (x \leftharpoonup x))$. By Lemma 7 it follows that $(x \cdot 1) \leftharpoonup x \approx z \leftharpoonup z$. So by Lemma 8 we get $x \cdot 1 \approx x$.

$y \cdot (w \leftharpoonup w) \approx z \cdot (w \leftharpoonup w)$
As just observed, we know that $(x \cdot 1 \leftharpoonup x \approx z \leftharpoonup z)$. Consequently, $(y \cdot (w \leftharpoonup w)) \leftharpoonup (z \cdot (w \leftharpoonup w)) \approx z \leftharpoonup z$. But then by Lemma 8 we arrive at $y \cdot (w \leftharpoonup w) \approx z \cdot (w \leftharpoonup w)$.

In particular, $p_i \cdot (z_i \dashv z_i) \approx q_i \cdot (z_i \dashv z_i)$ is a logical consequence of $\Gamma \cup \{\varepsilon\}$ for every $i < m$. Thus, each equation in Σ has a substitution instance which is a logical consequence of $\Gamma \cup \{\varepsilon\}$. So T is one-based according Theorem 5.

Now suppose that T is term-equivalent to an equational theory T_0 which has $\dashv, \cdot$, and 1 among its operation symbols and the three equations listed in the theorem among its equations. Thus, there are definitions for these three operation symbols in terms of the operation symbols of T. Let T_1 be the resulting definitional extension of T. By the argument above, there is a single equation $s \approx t$ that is a base for T_1. Then $\{\lambda(s) \approx \lambda(t)\}$ is a base of T where λ is the map that eliminates the new operation symbols from terms in favor of their definitions. $\square$

It is worth noting that if T satisfies the conditions of Corollary 10 and T' is another finitely based equational theory such that $T \subseteq T'$, then T' also satisfies the conditions of the Theorem, and so must also be one-based. We refer to one-based theories all of whose finitely based extensions are also one-based as **essentially one-based**. Thus, each finitely based equational theory of rings with unit is essentially one-based and Theorem 4 (iii) has been established.

A variant of Corollary 10 was discovered by Green and Tarski. They used this next corollary to establish that every finitely based equational theory of rings of the second kind is one-based.

Corollary 11 (Green and Tarski) *Let T be term equivalent to a finitely based equational theory with $\dashv$ and $*$ among its operation symbols to which that following equations belong:*

$$\begin{aligned} [(z \dashv z) \dashv (x \dashv y)] \dashv [(w \dashv w) \dashv x] &\approx y \\ x * (z \dashv z) &\approx z \dashv z \\ (z \dashv z) * x &\approx z \dashv z \\ x * t &\approx x \end{aligned}$$

where t is some term. Under these assumptions, T is one-based.

PROOF. As in the proof of the previous corollary, at first we suppose that T includes $\dashv$ and $*$ among its operation symbols. So the equations listed in the theorem actually belong to T. Let $\{p_0 \approx q_0, \dots, p_{m-1} \approx q_{m-1}\}$ be a base for T. Let Σ be

$$\{(p_0 * \dashv q_0) * z_0 \approx z \dashv z, \dots, (p_{m-1} \dashv q_{m-1}) * z_{m-1} \approx z \dashv z\} \cup \{x * t \approx x, \varepsilon\}$$

where none of the distinct variables $z_0, \dots, z_{m-1}$ occur in any of the $p_i \approx q_i$. The set Σ is a base for T. Take Γ to be $\{(x * t) \dashv x \approx (y * (z \dashv z)) * w\}$, where

we assume that y, x and w do not occur in the term t. Evidently, $\Gamma \subseteq T$.

The set $\Gamma \cup \{\varepsilon\}$ has the following logical consequences.

$y * (z \dashv z) \approx z \dashv z$
By substituting $y * (z \dashv z)$ for x and t^* for w in $(x * t) \dashv x \approx (y * (w \dashv w)) * z$ we obtain $(y * (z \dashv z) * t^*) \dashv y * (z \dashv z) \approx (y * (z \dashv z)) * t^*$, where t^* results from substituting $y * (z \dashv z)$ for x in t. By Lemma 7 it follows that $(y * (z \dashv z) * t^*) \dashv y * (z \dashv z) \approx (y * (z \dashv z)) * t^* \dashv (z \dashv z)$. By the Cancellation Lemma we get $y * (z \dashv z) \approx z \dashv z$.
$x * t \approx x$
Observe

$$\begin{aligned} x * t \dashv x &\approx (y * (z \dashv z)) * w \\ &\approx (z \dashv z) * w \\ &\approx (z \dashv z) * (z \dashv z) \\ &\approx z \dashv z \end{aligned}$$

So the equation we need follows from Lemma 8.

Notice, in particular, $(p_i \dashv q_i) * (z \dashv z) \approx z \dashv z$ is a logical consequence of $\Gamma \cup \{\varepsilon\}$ for every $i < m$. Thus, each equation in Σ has a substitution instance which is a logical consequence of $\Gamma \cup \{\varepsilon\}$. So T is one-based according Theorem 5.

Now suppose that T is term-equivalent to an equational theory T_0 which has $\dashv$ and $*$ among its operation symbols and the three equations listed in the theorem among its equations. Thus, there are definitions for these two operation symbols in terms of the operation symbols of T. Let T_1 be the resulting definitional extension of T. By the argument above, there is a single equation $s \approx t$ that is a base for T_1. Then $\{\lambda(s) \approx \lambda(t)\}$ is a base of T where λ is the map that eliminates the new operation symbols from terms in favor of their definitions. $\square$

To complete the proof of Theorem 4 (ii) we need to show that any equational theory of rings of the second kind satisfies the hypotheses of Corollary 11. Our line of reasoning is essentially that of Thomas Green.

PROOF. [Proof of Theorem 4 (ii)] Suppose that T is an equational theory of rings of the second kind. Because we need only concern ourselves with theories up to term equivalence, we assume without loss of generality that the standard ring operation symbols $+, -, 0$ and $\cdot$ as well as $\dashv$ are the operation symbols of T; moreover, we suppose that the equation $x \dashv y \approx x + (-y)$ belongs to T.

Actually, the only properties of rings that we will need are:

$$\begin{aligned} x + (y + z) &\approx (x + y) + z \\ x + y &\approx y + x \\ x + (-x) &\approx 0 \\ x + 0 &\approx x \\ x \cdot (y + z) &\approx (x \cdot y) + (x \cdot z) \\ (x + y) \cdot z &\approx (x \cdot z) + (y \cdot z). \end{aligned}$$

Let Δ be the set consisting of these equations together with $x \dot{-} y \approx x + (-y)$. So this line of reasoning applies to a class of equational theories wider than the theories of rings of the second kind.

To invoke Corollary 11 we must find a suitable term t and a term $p(x, y)$ in the two variables x and y which can be used to define $*$. We will take the term t to be the variable x. So what we need is a term $p(x, y)$ so that the following equations belong to T:

$$\begin{aligned} p(x, 0) &\approx 0 \\ p(0, x) &\approx 0 \\ p(x, x) &\approx x \end{aligned}$$

A **monomial** is a term built from the product $\cdot$ and variables. A **proper** monomial is one that is not a variable. A **sum** is a term of the form $s_0 + s_1 + \cdots + s_{m-1}$ where each s_i is a monomial or the negation of a monomial. This sum is **proper** if all the monomials involved are proper. The term s is a **sum in** x provided x is the only variable to occur in s. For each integer k we use kx to abbreviate $\underbrace{x + \cdots + x}_{k-\text{times}}$ when $k \geq 0$ and to abbreviate $-k(-x)$ when $k < 0$.

Let $I = \{k \mid 0 \approx kx + s \in T \text{ for some proper sum } s \text{ in } x\}$. I is an ideal of the ring of integers. Since such ideals are principal, let $d \geq 0$ be a generator of I.

On the basis of Δ every equation is equivalent to one of the form

$$0 \approx k_0 x_0 + \cdots + k_{m-1} x_{m-1} + s$$

where s is a proper sum and $x_0, \ldots, x_{m-1}$ are the variables appearing in the original equation. We say equations in this form are **normal** and we refer to $k_0, \ldots, k_{m-1}$ as **coefficients**. Because T is of the second kind, it cannot happen that the coefficients of normal equations in T are always 0. In particular, $I \neq \{0\}$ and so $d > 0$.

Suppose $0 \approx k_0 x_0 + \cdots + k_{m-1} x_{m-1} + s$ is a normal equation belonging to T. Let $i < m$. By substituting 0 for x_j whenever $i \neq j$, we see that $k_i \in I$. Hence,

$d \mid k_i$ for all $i < m$. This means that the cyclic group of order d equipped with the constantly 0 product is a model of T. Because T is a theory of the second kind, this model can have only one element. Hence $d = 1$.

This means that $0 \approx -x+s \in T$ for some proper sum s in x. Hence $s \approx x \in T$. Let $p(x, y)$ result from s changing the leftmost occurrence of x to y in each of the monomials in s. This term $p(x, y)$ has the required properties. □

Theorem 5 has another corollary that has particular applications to equational theories of groups as well as to rings.

Corollary 12 (Tarski) *Let T be a finitely based equational theory such that $\varepsilon \in T$ and every model of T has a one-element subalgebra. Under these assumptions, T has a base with no more than* $\max\{1, n-1\}$ *where n is the number of operation symbols occurring in T.*

PROOF. Because Lemma 7 tells us that $x \dashv x \approx y \dashv y \in T$ we see that each model of T has a unique one-element subalgebra whose single element is denoted by the term $x \dashv x$. If $n > 1$, let Γ be the set consisting of the $n-1$ equations of the form

$$Q(x \dashv x)(x \dashv x) \dots (x \dashv x) \approx x \dashv x$$

where Q is an operation symbol other than $\dashv$. If $n = 1$ (that is, $\dashv$ is the only operation symbol in T) let $\Gamma = \{x \approx x\}$. Then $\Gamma \subseteq T$. Now let Σ be any finite base for T. For each equation $u \approx r \in \Sigma$ let $u^* \approx r^*$ be the result of substituting $x \dashv x$ for each variable in $u \approx r$. Evidently $\Gamma \cup \{\varepsilon\} \vdash u^* \approx r^*$ for every $u \approx r \in \Sigma$. So according to Theorem 5, T has a base with no more than $\max\{1, n-1\}$ equations. □

This corollary includes the Theorem of Higman and Neumann according to which each finitely based equational theory of groups, construed as algebras with the sole basic operation of right division, is one-based. Moreover, every finitely based equational theory of groups is one-based provided it is framed using two operation symbols, one of which is $\dashv$. Of course, every finitely based equational theory of groups has a definitional extension including the operation symbol $\dashv$. So every finitely based theory of groups with n operation symbols has a base with no more than $\max\{1, n\}$ equations. In particular, we see that each finitely based equational theory of groups using the customary symbols $\cdot$ and $^{-1}$ has a base with no more than two equations. This result is improved by the following theorem due to Thomas C. Green.

Theorem 13 *A finitely based equational theory of groups has a base with no more than* $\max\{1, n-1\}$ *equations, where* n *is the number of operation symbols.*

PROOF. It does no harm to restrict our attention to a finitely based equational theory T of groups such that T has at least 2 operation symbols and T has models with more than one element. We also suppose that $\dashv$ and the constant symbol e do not occur in T.

Let $Q_0, \ldots, Q_{n-1}$ be the operation symbols occurring in T. Now T is term equivalent to a theory T_0 in just the symbols $\dashv$ and e such that $\varepsilon, e \approx z \dashv z \in T_0$. We let T_1 be a common definitional extension of T and T_0. We select terms $t_0, \ldots, t_{n-1}$ built from variables $x_0, x_1, \ldots$ and the operation symbols $\dashv$ and e, and the term $d(x, y)$ built using only the variables x and y and the operation symbols $Q_0, \ldots, Q_{n-1}$ has appropriate definitions. In particular, the following equations belong to T_1:

$$\begin{aligned} x \dashv y &\approx d(x, y) \\ Q_0 x_0 x_1 \ldots x_{r_0-1} &\approx t_0(x_0, x_1, \ldots, x_{r_0-1}) \\ &\vdots \\ Q_{n-1} x_0 x_1 \ldots x_{r_{n-1}-1} &\approx t_{n-1}(x_0, x_1, \ldots, x_{r_{n-1}-1}) \end{aligned}$$

Here r_i denotes the rank of the operation symbol Q_i. The only variables to occur in t_i are among $x_0, \ldots, x_{r_i-1}$.

It is convenient to abbreviate various terms. When s and t are terms in which the variable z does not occur and r is an integer, then

$$\begin{aligned} s^{-1} &\text{ abbreviates } (z \dashv z) \dashv s \\ s \cdot t &\text{ abbreviates } s \dashv (t^{-1}) = s \dashv ((z \dashv z) \dashv t) \\ e_i &\text{ abbreviates } Q_i(z \dashv z) \ldots (z \dashv z) \\ s^r &\text{ abbreviates } \begin{cases} \underbrace{s \cdot \cdots \cdot s}_{r\text{-times}} & \text{if } r > 0 \\ (\underbrace{s \cdot \cdots \cdot s}_{r\text{-times}})^{-1} & \text{if } r < 0 \\ z \dashv z & \text{if } r = 0 \end{cases} \end{aligned}$$

Let $\Sigma = T_0 \cup \{e_i \cdot x \approx x \cdot e_i \mid i < n\} \cup \{Q_i x_0 \ldots x_{r_i-1} \approx t_i \cdot e_i \mid i < n\}$. Let $\Gamma = \{e \approx z \dashv z\}$. It is not hard to see that every equation in Σ has a substitution instance which is a consequence of $\Gamma \cup \{\varepsilon\}$. For equations from T_0 we can substitute e for every variable. For equations of the form $e_i \cdot x \approx x \cdot e_i$ we substitute e_i for x, while for equations of the last sort we can substitute $z \dashv z$ for all the x_i's. Lemma 7 helps secure these substitution instances. So it

follows from Theorem 5 that we can pick a single equation η which is logicially equivalent to Σ.

By reasoning familiar from elementary group theory, for any term t built from the variable x and the operation symbols $\dashv$ and e there is an integer p such that $t \approx x^p \in T_0$. Pick integers $p_0, p_1, \ldots, p_{n-1}$ so that $t_j(x, x, \ldots, x) \approx x^{p_j} \in T_0$.

Recall that the term $d(x, y)$ built from the variables x and y and the operations symbols Q_i for $i < n$ serves as a definition of $x \dashv y$. Using the equations $Q_i x_0 \ldots x_{r_i - 1} \approx t_i \cdot e_i$ we can recursively eliminate the Q_i's from $d(x, x)$ in favor of $t_i \cdot e_i$ to obtain a term $\bar{d}$ built from the variable x, the operation symbols $\dashv$ and e, and the terms e_i. In fact $\eta \vdash d(x, x) \approx \bar{d}$. A careful accounting shows

$$\eta \vdash d(x, x) \approx x^{1-m} \cdot e_0^{r_{0,0}} \cdot e_1^{r_{0,1}} \cdot \cdots \cdot e_{n-1}^{r_{0,n-1}}$$

where each $r_{0,j}$ is an integer and

$$m = -\sum_{j<n} r_{0,j}(p_j - 1).$$

Now $x \dashv x \approx x^{1-m} \in T_0$ by the term equivalence between T and T_0. Consequently, $x^m \approx x \in T_0$. This means

$$\eta \vdash x^m \approx x.$$

Since T has a nontrivial model, we know that $m \neq 0$.

The following lemma from linear algebra was attributed by Thomas Green to Andrew Ogg.

Lemma 14 *Let $\langle r_{0,0}, r_{0,1}, \ldots, r_{0,n-1} \rangle$ and $\langle c_0, c_1, \ldots, c_{n-1} \rangle$ be n-tuples of integers. There are $n \times n$ integer matrices A and R so that*

(1) $AR = mI$, where $m = \sum_{j<n} c_j r_{0,j}$, and
(2) $\langle r_{0,0}, r_{0,1}, \ldots, r_{0,n-1} \rangle$ is the top row of the matrix R.

The proof uses the Euclidean algorithm for computing greatest common divisors and some elementary matrix theory. Roughly speaking, the matrix R is constructed by starting with the following matrix

$$\begin{pmatrix} g & 0 & 0 \ldots & 0 \\ 0 & m/g & 0 \ldots & 0 \\ 0 & 0 & 1 \ldots & 0 \\ \vdots & \vdots & \vdots \ddots & \vdots \\ 0 & 0 & 0 \ldots & 1 \end{pmatrix}$$

where g is the greatest common divisor of $\langle r_{0,0}, r_{0,1}, \ldots, r_{0,n-1}\rangle$, and reversing the steps in the Euclidean algorithm, treated as elementary column operations, to obtain the desired top row. In this way, R will be an integer matrix with determinant m. The matrix A is the adjoint of R.

The integer $r_{i,j}$ is the (i,j) entry in the matrix R. Let

$$\Psi = \{\eta\} \cup \{e \approx e_0^{r_{i,0}} \cdot e_1^{r_{i,1}} \cdot \cdots \cdot e_{n-1}^{r_{i,n-1}} \mid i < n\}$$

Now let $\langle a_{i,0}, \ldots, a_{i,n-1}\rangle$ be the i-th row of A and let $j < n$. Then

$$a_{i,0}r_{0,j} + a_{i,1}r_{1,j} + \cdots + a_{i,n-1}r_{n-1,j} = \begin{cases} m & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Relying on the equations $e_i \cdot x \approx x \cdot e_i$ and familiar group theory we obtain the following consequences of Ψ:

$$\begin{aligned} e^{a_{i,0}} &\approx e_0^{a_{i,0}r_{0,0}} \cdot e_1^{a_{i,0}r_{0,1}} \cdot \cdots \cdot e_{n-1}^{a_{i,0}r_{0,n-1}} \\ e^{a_{i,1}} &\approx e_0^{a_{i,1}r_{1,0}} \cdot e_1^{a_{i,1}r_{1,1}} \cdot \cdots \cdot e_{n-1}^{a_{i,1}r_{1,n-1}} \\ &\vdots \\ e^{a_{i,n-1}} &\approx e_0^{a_{i,n-1}r_{n-1,0}} \cdot e_1^{a_{i,n-1}r_{n-1,1}} \cdot \cdots \cdot e_{n-1}^{a_{i,n-1}r_{n-1,n-1}} \end{aligned}$$

Again with the help of the equations $e_i \cdot x \approx x \cdot e_i$ and $e^k \approx e$ for all k we can in essence multiply these equations vertically to obtain the following consequence of Ψ:

$$e \approx e_i^m$$

for each $i < n$. But $\Psi \vdash x^m \approx x$. This means that $\Psi \vdash e \approx e_i$ for each $i < n$. Since $\Psi \vdash Q_i x_0 x_1 \ldots x_{r_i - 1} \approx t_i \cdot e_i$, consequently,

$$\Psi \vdash T_0 \cup \{Q_i x_0 x_1 \ldots x_{r_i - 1} \approx t_i \mid i < n\}.$$

This means that Ψ is a base for T_1.

Now we make a small adjustment in Ψ. Let

$$\Delta = \{\eta\} \cup \{d(x,y) \approx x \dot{-} y\} \cup \{e \approx e_0^{r_{i,0}} \cdot e_1^{r_{i,1}} \cdot \cdots \cdot e_{n-1}^{r_{i,n-1}} \mid 1 \leq i < n\}$$

To see that Δ is also a base for T_1 we need only derive the equation

$$e \approx e_0^{r_{0,0}} \cdot e_1^{r_{0,1}} \cdot \cdots \cdot e_{n-1}^{r_{0,n-1}}.$$

We have already noted the following consequences of η

$$d(x,x) \approx x^{1-m} \cdot e_0^{r_{0,0}} \cdot e_1^{r_{0,1}} \cdot \cdots \cdot e_{n-1}^{r_{0,n-1}}$$
$$x \dot{-} x \approx x^{1-m}$$

So using $d(x,x) \approx x \dot{-} x$ we get

$$x \dot{-} x \approx (x \dot{-} x) \cdot e_0^{r_{0,0}} \cdot e_1^{r_{0,1}} \cdot \cdots \cdot e_{n-1}^{r_{0,n-1}}$$

this gives immediately

$$e \approx e_0^{r_{0,0}} \cdot e_1^{r_{0,1}} \cdot \cdots \cdot e_{n-1}^{r_{0,n-1}}$$

So Δ is a base for T_1 and Δ has cardinality $n+1$. Now using the equation $d(x,y) \approx x \dot{-} y$ we can eliminate $\dot{-}$ from Δ to obtain the set Δ', which will have just n equations (since we can drop the tautology $d(x,y) \approx d(x,y)$). Now we use one of the equations in Δ' as a definition of e. Eliminating e from Δ' results in the desired base of T with $n-1$ equations (since we can drop another tautology). $\square$

To tackle the proof of Theorem 4 (i) we need to prove an analog of Theorem 13.

Theorem 15 *A finitely based equational theory of rings has a base with no more than* $\max\{1, n-1\}$ *equations, where* n *is the number of operation symbols.*

PROOF. The proof given above for Theorem 13 needs to be modified at several points. This time we assume that $\dot{-}$ and $\cdot$ do not occur in T and we take T_0 to be term equivalent to T in just the symbols $\dot{-}$ and $\cdot$. As before the equational theory T_1 is the common definitional extension of T and T_0. We select terms $t_0, \ldots, t_{n-1}$ built from variables $x_0, x_1, \ldots$ and the operation symbols $\dot{-}$ and $\cdot$, and terms $d(x,y)$ and $p(x,y)$ built using only the variables x and y and the operation symbols $Q_0, \ldots, Q_{n-1}$ has appropriate definitions. In particular, the following equations belong to T_1:

$$\begin{aligned}
x \dot{-} y &\approx d(x,y) \\
x \cdot y &\approx p(x,y) \\
Q_0 x_0 x_1 \ldots x_{r_0-1} &\approx t_0(x_0, x_1, \ldots, x_{r_0-1}) \\
&\vdots \\
Q_{n-1} x_0 x_1 \ldots x_{r_{n-1}-1} &\approx t_{n-1}(x_0, x_1, \ldots, x_{r_{n-1}-1})
\end{aligned}$$

Here r_i denotes the rank of the operation symbol Q_i. The only variables to occur in t_i are among $x_0, \ldots, x_{r_i-1}$.

It is convenient to abbreviate various terms. When s and t are terms in which the variable z does not occur and r is an integer, then

$$\begin{aligned}
-s &\text{ abbreviates } (z \dashv z) \dashv s \\
s + t &\text{ abbreviates } s \dashv (t^{-1}) = s \dashv ((z \dashv z) \dashv t) \\
e_i &\text{ abbreviates } Q_i(z \dashv z) \dots (z \dashv z) \\
rs &\text{ abbreviates } \begin{cases} \underbrace{s + \dots + s}_{r\text{-times}} & \text{if } r > 0 \\ -(\underbrace{s + \dots + s}_{r\text{-times}}) & \text{if } r < 0 \\ z \dashv z & \text{if } r = 0 \end{cases}
\end{aligned}$$

Let $\Sigma = T_0 \cup \{e_i \cdot x \approx z \dashv z \mid i < n\} \cup \{x \cdot e_i \approx z \dashv z \mid i < n\} \cup \{Q_i x_0 \dots x_{r_i - 1} \approx t_i + e_i \mid i < n\}$. Now T_0 fulfills the conditions of Corollary 12 so we can pick a single equation γ which is a base for T_0. Among the consequences of γ we find:

$$\begin{gathered}
\varepsilon \\
x \cdot (z \dashv z) \approx z \dashv z \\
(z \dashv z) \cdot z \approx z \dashv z
\end{gathered}$$

As a consequence of Theorem 5, taking $\Gamma = \{\gamma\}$ we find that there is a single equation η which is logically equivalent with Σ.

Observe that any term built using just $\dashv$ and $\cdot$ and variables can be recast, using η, in the form $a_0 x_0 + \dots + a_{\ell-1} x_{\ell-1} + s$ where s is a proper sum and $a_0, \dots, a_{\ell-1}$ are integers. In particular, we pick integers p_i and proper sums s_i so that

$$\eta \vdash t_i(x, x, \dots, x) \approx p_i x + s_i$$

As in the proof of Theorem 13 this leads to integers $r_{0,0}, \dots, r_{0,n-1}$ and a proper sum $s(x)$ so that

$$\begin{aligned}
&\eta \vdash d(x, x) \approx (1 - m)x + s + r_{0,0} e_0 + \dots + r_{0,n-1} e_{n-1} \\
&m = -\sum_{j<n} r_{0,j}(p_j - 1) \\
&\eta \vdash mx \approx x + s
\end{aligned}$$

We can obtain integers $r_{i,j}$ as in the previous proof. We take

$$\Psi = \{\eta\} \cup \{z \dashv z \approx r_{i,0} e_0 + \dots + r_{i,n-1} e_{n-1} \mid i < n\}$$

Among the consequences of Ψ we find $z \dashv z \approx m e_i$ for each i as well as $mx \approx x + s(x)$ and $z \dashv z \approx s(e_i)$. This means $\Psi \vdash z \dashv z \approx e_i$ for each $i < n$.

Hence Ψ is a base for T_1. Finally, we take Δ to be the union of the following three sets of equations:

$$\{\eta\}$$
$$\{x \dot{-} y \approx d(x,y)\}$$
$$\{x \cdot y \approx p(x,y) \dot{-} p(z \dot{-} z, z \dot{-} z) + r_{i,0}e_0 + \cdots + r_{i,n-1}e_{n-1} \mid 1 \leq i < n\}$$

Notice that each equation in the latter set, with the help of η, entails $z \dot{-} z \approx r_{i,0}e_0 + \cdots + r_{i,n-1}e_{n-1}$.

The rest of this proof is accomplished in the same manner that the proof of Theorem 13 is concluded. $\Box$

At this point we have in hand bases of all the cardinalities required in Theorems 3 and 4. So we turn to the task of demonstrating the nonexistence of smaller bases, has required by Theorem 3 and Theorem 4 (i).

Lemma 16 (Thomas Green) *Let T be a finitely based equational theory of groups such that T has a model with more than one element. Assume that the operation symbols of T are $\dot{-}$ and the n constant symbols $e_0, \ldots, e_{n-1}$ all denoting the identity element. Under these assumptions, T has no base with fewer than n equations.*

PROOF. Let $\mathbf{G}$ be a finite Abelian group with more than one element which is a model of T. (In fact, we could even choose $\mathbf{G}$ to be cyclic.) So $\mathbf{G} = \langle G, \dot{-}, 0, \ldots, 0\rangle$ where we can construe $\dot{-}$ as subtraction. As above we take $x + y$ as an abbreviation from the term $x \dot{-} ((z \dot{-} z) \dot{-} y)$.

Let $\mathbf{r} = \langle r_0, \ldots, r_{n-1}\rangle$ be an n-tuple of elements of G. Let $\mathbf{G_r}$ be the algebra $\langle G, \dot{-}, r_0, \ldots, r_{n-1}\rangle$. So $\mathbf{G_r}$ is a model of T if and only if $\mathbf{r} = \langle 0, \ldots, 0\rangle$.

Just using the properties of Abelian groups we know that every equation $s \approx t$ in the operation symbols $\dot{-}, e_0, \ldots, e_{n-1}$ is equivalent to one of the form

$$z \dot{-} z \approx a_0e_0 + \cdots + a_{n-1}e_{n-1} + b_0x_0 + \cdots + b_{m-1}x_{k-1}$$

where $a_0, \ldots, a_{n-1}, b_0, \ldots, b_{k-1}$ are certain integers and $x_0, \ldots, x_{m-1}$ are the variables that occur in $s \approx t$.

Now suppose $z \dot{-} z \approx a_0e_0 + \cdots + a_{n-1}e_{n-1} + b_0x_0 + \cdots + b_{k-1}x_{k-1} \in T$. Then $z \dot{-} z \approx b_0x_0 + \cdots + b_{k-1}x_{k-1} \in T$. Since none of the e_i's occurs in this last equation, we see that $\mathbf{G_r}$ is a model of $z \dot{-} z \approx b_0x_0 + \cdots + b_{k-1}x_{k-1} \in T$, regardlesss of the choice of $\mathbf{r}$.

So for $s \approx t \in T$, we find that $\mathbf{G_r}$ is a model of $s \approx t$ if and only if $0 = a_0r_0 + a_1r_1 + \cdots + a_{n-1}r_{n-1}$ where integer multiples stand for repeated additions and all the additions are carried out in the sense of $\langle G, \dot{-} \rangle$.

Now, for the sake of contradiction, suppose that Σ is a base for T and that $|\Sigma| = m < n$. As above, each equation in Σ is associated wtih an n-tuple of integers. These m n-tuples can be organized into an $m \times n$ matrix

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m-1,0} & a_{m-1,1} & \dots & a_{m-1,n-1} \end{pmatrix}$$

Taking $\mathbf{r}$ as a column vector and letting $\mathbf{0}$ denote the column vector of 0's we find

$$\mathbf{G_r} \text{ is a model of } \Sigma \text{ if and only if } A\mathbf{r} = \mathbf{0}$$

where the additions in the matrix multiplication are carried out in $\langle G, \dot{-} \rangle$. But since Σ is a base for T this means

$$\mathbf{r} = \mathbf{0} \text{ if and only if } A\mathbf{r} = \mathbf{0}.$$

But multplication by A gives a function from G^n into G^m. Since the finite set G^n is larger than the finite set G^m, this function must fail to be one-to-one. So there must be an n-tuple $\mathbf{r} \neq \mathbf{0}$ so that $A\mathbf{r} = \mathbf{0}$. This is a contradiction. $\square$

With the help of this lemma and Theorem 13 we can complete the proof of Theorem 3.

PROOF. [Proof of Theorem 3] Let T be a finitely based theory of groups in n operation symbols such that T has a model with more than one element. All that remains in to prove that if $n > 1$, then T has no base with fewer than $n - 1$ elements. Without loss of generality, we assume that the symbol $\dot{-}$ does not occur in T. Let $d(x, y)$ be a term in the symbols of T that can define difference/division. Let T' be the definitional extension of T based on $T \cup \{x \dot{-} y \approx d(x, y)\}$. Then T' is an equational theory using $n + 1$ operation symbols. Suppose for the moment that we know that every base of T' has cardinality at least n. Let Σ we any base for T. Then $\Sigma \cup \{x \dot{-} y \approx d(x, y)\}$ is a base for T'. Hence $n \leq |\Sigma| + 1$. This means that $n - 1 \leq |\Sigma|$, as desired. It remains to argue that every base of T' has cardinality at least n.

Take $Q_0, Q_1, \dots, Q_{n-1}$ to be the operation symbols that occur in T and take $e_0, e_1, \dots, e_{n-1}$ to be constant symbols not occuring in T. We consider three

equational theories which are term equivalent to T:

T'	with operation symbols $\dot{-}, Q_0, \ldots, Q_{n-1}$
T'_1	with operation symbols $\dot{-}$
T'_2	with operation symbols $\dot{-}, Q_0, \ldots, Q_{n-1}, e_0, \ldots, e_{n-1}$
T'_3	with operation symbols $\dot{-}, e_0, \ldots, e_{n-1}$

Observe that T'_2 is based on $T' \cup \{z \dot{-} z \approx e_i \mid i < n\}$. Now pick n terms $t_0, \ldots, t_{n-1}$ built up from variables with only the help of $\dot{-}$ and such that $Q_i x_0 x_1 \ldots x_{r_i - 1} \approx t_i \in T'$ for each $i < n$. Let

$$\Delta = \{Q_i x_0 \ldots x_{r_i - 1} \approx t_i \dot{-} e_i \mid i < n\}.$$

To see that $T' \cup \Delta$ is a base for T'_2 we only need to derive $z \dot{-} z \approx e_i$ for each $i < n$. Here is how. The equation $t_i \approx t_i \dot{-} e_i$ is immediate from the last few lines. Next ε holds since T'_1 is a theory of groups. So $t_i \dot{-} (z \dot{-} z) \approx t_i \dot{-} e_i$ follows from Lemma 7. So from the Cancellation Lemma we obtain $z \dot{-} z \approx e_i$ as desried.

Now let Σ be any base for T'. Then $\Sigma \cup \Delta$ is a base for T'_2. Now use the equations in Δ to eliminate the symbols $Q_0, \ldots, Q_{n-1}$, obtaining a set Σ' of equations in $\dot{-}$ and $e_0, \ldots, e_{n-1}$. This set Σ' is a base of T'_3 and $|\Sigma'| \leq |\Sigma|$. But by Lemma 16 we know that $n \leq |\Sigma'|$. So Σ has no fewer than n equations. $\square$

We are now in a position to complete the proof of Theorem 4 (i).

PROOF. [The proof of Theorem 4 (i)] Suppose that T is a finitely based equational theory of rings so that T is of the first kind. Let n be the number of operation symbols occurring in T. We already know that T has a base consisting of $n - 1$ equations according to Theorem 15. We argue here that every base of T has at least $n - 1$ equations.

Without loss of generality, we suppose that the operation symbols $\dot{-}$ and $\cdot$ do not occur in T.

Let Σ be a base for T. Pick terms $d(x, y)$ and $p(x, y)$ built from the operations of T and the variables x and y so that $d(x, y)$ denotes the ring difference and $p(x, y)$ denotes the ring product in every model of T. Let T' be the equational theory based on $\Sigma \cup \{x \dot{-} y \approx d(x, y), x \cdot y \approx p(x, y)\}$. So T' is a definitional extension of T and all the operations of T can be defined by terms in $\dot{-}$ and $\cdot$. Also T' is an equational theory of the first kind. Let T'' be the equational theory based on $\Sigma \cup \{x \dot{-} y \approx d(x, y), x \cdot y \approx p(x, y), x \cdot y \approx z \dot{-} z\}$. Since T' is an equational theory of the first kind, we know that T'' has a model with more

than one element. But observe that T'' is a finitely based equational theory of groups since all the operations can be defined using only terms built from the group operation $\dot{-}$ and variables.

Claim 17 *The set* $\Sigma \cup \{x \dot{-} y \approx d(x,y), x \cdot z \approx p(x,y)\}$ *is a base for* T''.

PROOF. The equation $x \cdot z \approx p(x,y)$ is a consequence of $x \cdot y \approx p(x,y)$ and $x \cdot y \approx z \dot{-} z$. It is also evident that $x \cdot y \approx p(x,y)$ is a consequence of $x \cdot z \approx p(x,y)$. It remains only to show that $x \cdot y \approx z \dot{-} z$ is a consequence of $\Sigma \cup \{x \dot{-} y \approx d(x,y), x \cdot z \approx p(x,y)\}$. Now in any ring $x \cdot (z \dot{-} z) \approx z \dot{-} z$. Using this equation and $x \cdot (z \dot{-} z) \approx p(x,y)$, which is a substitution instance of $x \cdot z \approx p(x,y)$, we obtain $z \dot{-} z \approx p(x,y)$. But we have the consequence $x \cdot y \approx p(x,y)$. So symmetry and transitivity yield $x \cdot y \approx z \dot{-} z$. □

T'' is an equational theory of groups in $n+2$ operation symbols and it has a model with more than one element. By Theorem 3 any base of T'' must have at least $n+1$ equations. This means $\Sigma \cup \{d(x,y) \approx x \dot{-} y, x \cdot z \approx p(x,y)\}$ has at least $n+1$ elements. Therefore $|\Sigma| \geq n-1$, as desired. □

References

Garrett Birkhoff

(1935) On the structure of abstract algebras. *Proc. Cambridge Philos. Soc.*, **31**, 433–454.

Steven Givant

(1975) Possible cardinalities of irredundant bases for finite closure structures. *Discrete Math.*, **12**, 201–204.

George Grätzer and Ralph N. McKenzie

(1967) Equational spectra and reduction of identities. *Notices Amer. Math. Soc.*, **14**, 697.

George Grätzer and R. Padmanabhan

(1978) Symmetric difference in abelian groups. *Pacific J. Math.*, **74**, 339–347.

Thomas C. Green and Alfred Tarski

(1970a) The minimum cardinality of equational bases for varieties of groups and rings. *Notices Amer. Math. Soc.*, **17**, 429–430.

(1970b) The least cardinality of equational bases for varieties of groups and rings. *Congrès International des Mathématiciens, Nice 1970: les 265 Communications Individuelles*, 13.

Marshall Hall, Jr.

(1976) *The theory of groups*. Chelsea Publishing Co., New York. Reprinting of the 1968 edition.

Graham Higman and B. H. Neumann
(1952) Groups as groupoids with one law. *Publ. Math. Debrecen*, **2**, 215–221.
Jan Kalicki
(1955) The number of equationally complete classes of equations. *Nederl. Akad. Wetensch. Proc. Ser. A.* **58** = *Indag. Math.*, **17**, 660–662.
William McCune and R. Padmanabhan
(1996) Single identities for lattice theory and for weakly associative lattices. *Algebra Universalis*, **36**, 436–449.
Ralph N. McKenzie
(1970) Equational bases for lattice theories. *Math. Scand.*, **27**, 24–38.
Ralph N. McKenzie, George F. McNulty, and Walter F. Taylor
(1987) *Algebras, lattices, varieties. Vol. I*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA. ISBN 0-534-07651-3.
George F. McNulty
(1976) The decision problem for equational bases of algebras. *Ann. Math. Logic*, **10**, 193–259.
George F. McNulty and Walter F. Taylor
(1975) Combinatory interpolation theorems. *Discrete Math.*, **12**, 193–200.
Nathan S. Mendelsohn and R. Padmanabhan
(1972) A single identity for Boolean groups and Boolean rings. *J. Algebra*, **20**, 78–82.
(1975) Minimal identities for Boolean groups. *J. Algebra*, **34**, 451–457.
R. Padmanabhan
(1969a) Inverse loops as groupoids with one law. *J. London Math. Soc. (2)*, **1**, 203–206.
(1969b) On single equational-axiom systems for abelian groups. *J. Austral. Math. Soc.*, **9**, 143–152.
(1969c) Two identities for lattices. *Proc. Amer. Math. Soc.*, **20**, 409–412.
(1972) On identities defining lattices. *Algebra Universalis*, **1**, 359–361.
R. Padmanabhan and Robert W. Quackenbush
(1973) Equational theories of algebras with distributive congruences. *Proc. Amer. Math. Soc.*, **41**, 373–377.
R. Padmanabhan and B. Wolk
(1981) Equational theories with a minority polynomial. *Proc. Amer. Math. Soc.*, **83**, 238–242.
Alfred Tarski
(1938) Ein Beitrag zur Axiomatik der Abelschen Gruppen. *Fund. Math.*, **30**, 253–256.
(1968) Equational logic and equational theories of algebras. In *Contributions to Math. Logic (Colloquium, Hannover, 1966)*. North-Holland, Amsterdam, pages 275–288.
(1975) An interpolation theorem for irredundant bases of closure structures. *Discrete Math.*, **12**, 185–192.