Can. J. Math., Vol. XXX, No. 2, 1978, pp. 231–237

AFFINE PARTS OF ALGEBRAIC THEORIES II

J. R. ISBELL, M. I. KLUN, AND S. H. SCHANUEL

Introduction. This paper concerns relative complexity of an algebraic theory $\mathbf{T}$ and its affine part $\mathbf{A}$, primarily for theories $\mathbf{T}_R$ of modules over a ring R. $\mathbf{T}_R$, $\mathbf{A}_R$ and R itself are all, or none, finitely generated or finitely related. The minimum number of relations is the same for $\mathbf{T}_R$ and $\mathbf{A}_R$. The minimum number of generators is a very crude invariant for these theories, being 1 for $\mathbf{A}_R$ if it is finite, and 2 for $\mathbf{T}_R$ if it is finite (and $1 \neq 0$ in R). The minimum arity of generators is barely less crude: 2 for $\mathbf{T}_R$, and 2 or 3 for $\mathbf{A}_R$ $(1 \neq 0)$. $\mathbf{A}_R$ is generated by binary operations if and only if R admits no homomorphism onto $\mathbf{Z}_2$.

For arbitrary algebraic theories $\mathbf{T}$, of course $\mathbf{A}$ is a subtheory, so that the cardinality of $\mathbf{T}$ bounds that of $\mathbf{A}$. However, we find that a finitely presented theory can have an affine part that is not finitely generated, or a finitely generated affine part that is not finitely related.

In general, $\mathbf{T}$ is not determined by $\mathbf{A}$; the close relationship between $\mathbf{T}_R$ and $\mathbf{A}_R$ holds because $\mathbf{T}_R$ is constructible from $\mathbf{A}_R$ by the "trivial" process of adjoining a constant symbol. We find that for general affine theories this process is not reversible; there are non-isomorphic affine $\mathbf{A}$, $\mathbf{B}$, which produce isomorphic theories $\mathbf{A}^*$, $\mathbf{B}^*$ when a constant is adjoined. The theories $\mathbf{A}$, $\mathbf{B}$ are not even mutually interpretable. We have not been able to find finitely presented affine examples for this; however, we have a similarly related pair of theories $\mathbf{V}$, $\mathbf{W}$, which are finitely presented but not affine.

1. Simplicity. We define the *pointed cover* $\mathbf{A}^*$ of an algebraic theory $\mathbf{A}$ as the coproduct of $\mathbf{A}$ and the theory $\mathbf{P}$ of pointed sets. Thus a model of $\mathbf{A}^*$ is a (non-empty) model A of $\mathbf{A}$ with a distinguished element, p, and morphisms are $\mathbf{A}$-morphisms preserving distinguished elements. The n-ary operations $w(x_1, \ldots, x_n)$ of $\mathbf{A}^*$ may be described as the $(n+1)$-ary operations $v(x_0, \ldots, x_n)$ of $\mathbf{A}$, with the understanding that x_0 is treated as a constant denoting p. (For the coproduct must have these operations; these $w(x_1, \ldots, x_n) = v(p, x_1, \ldots, x_n)$ are all different since p is arbitrary; and these operations are closed under composing with operations of $\mathbf{A}$ or of $\mathbf{P}$).

Note the known result about theories $\mathbf{T}_R$ of modules and their affine parts $\mathbf{A}_R$:

$$\mathbf{T}_R = \mathbf{A}_R^*.$$

Received April 27, 1976 and in revised form, November 4, 1977. This research was supported in part by NSF Grant PO28491-002.

(For models of $\mathbf{A}_R$ are cosets in R-modules, isomorphic by translation with submodules. In view of translations, the category of affine modules with distinguished points is isomorphic, by an isomorphism preserving ground sets, with the category of modules. Hence the theories are isomorphic.)

PROPOSITION. *For any theory* $\mathbf{A}$, *both or neither of* $\mathbf{A}$ *and* $\mathbf{A}^*$ *are finitely generated; both or neither are finitely related; the minimum number of operations generating* A^* *is either the same as for* A *or one greater; the minimum number of defining relations (generators varying) is the same for both.*

Proof. Of course, generators of $\mathbf{A}$ and one generator p of $\mathbf{P}$ generate $\mathbf{A} + \mathbf{P} = \mathbf{A}^*$. Conversely, generators for $\mathbf{A}^*$ may be written $y_i(p, x_1, \ldots, x_{ni})$; if a general operation $\nu(p, x_1, \ldots, x_n)$ is expressed as a composite $K((\nu_i(p, \ldots)))$, it is true in $\mathbf{A}$ that $\nu(x_0, \ldots, x_n) = K((\nu_i(x_0, \ldots)))$. (Legally, we just suppressed an induction.) Similarly, generators and relations for $\mathbf{A}$, plus p and no relations, define $\mathbf{A}^*$. Generators and relations for $\mathbf{A}^*$ written in terms of $\nu(p, \ldots)$ work ("for general p") for $\mathbf{A}$.

A theory $\mathbf{T}_R$, one may readily check, is finitely generated or finitely related if and only if R is. For theories like these, the minimum number of generators is not very sensitive. If R is finitely generated by g_i's, $\mathbf{A}_R$ has the single generator $x_1 + x_2 - x_3 + \sum g_i(y_i - z_i)$. So $\mathbf{T}_R$ has two generators; not one, since 0-ary and non-0-ary operations never generate each other.

To amplify the last remark: if $\mathbf{A}$ has no constants, the minimum number of generators for $\mathbf{A}^*$ is one greater. For the generators $\nu_i(p, \ldots)$ of $\mathbf{A}^*$ must include a constant, which can be omitted for $\mathbf{A}$. It seems likely that one can extend this equation to reasonable theories with constants, but there are such counterexamples as the theory $\mathbf{A}$ of pointed semigroups with $x^2 = 0$. Two generators x_1x_2 and 0 are needed for $\mathbf{A}$; two generators x_1x_2 and p suffice for $\mathbf{A}^*$.

On the minimum number of relations for $\mathbf{T}_R$ (if finite), there is a small finite bound because, once we know we have modules, we can combine laws as generators g_i were combined above. Tarski announced in [**2**] a theorem implying that $\mathbf{T}_R$ is definable by $x_1 - x_2$ and 0 with one relation if $R = \mathbf{Z}$ or $\mathbf{Z}_n$.

Though the number of generators required for $\mathbf{A}_R$ is always 1 or infinite, the generators for various R are not equally simple. $\mathbf{A}_R$ is always generated by binary and ternary operations (viz. $x + y - z$ and all $gx + (1 - g)y$). It is generated by binary operations if and only if R admits no homomorphism onto $\mathbf{Z}_2$. We prove a bit more. Define the *Boolean radical* B of R to be the intersection of the kernels of all (1-preserving) ring homomorphisms from R to $\mathbf{Z}_2$. If there are no such, then of course $B = R$.

PROPOSITION. *The affine operation*

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} r_i x_i \quad (\textit{with } \textstyle\sum r_i = 1)$$

is expressible in terms of binary affine operations if and only if $r_i r_j \in B$ *for all pairs* $i \neq j$.

https://doi.org/10.4153/CJM-1978-021-3 Published online by Cambridge University Press

Proof. If $R = \mathbf{Z}_2$, the only affine operations f expressible in terms of binary affines are those with one $r_k = 1$ and the rest zero; that is, those with $r_i r_j = 0$ for all $i \neq j$. From the definition of B it follows for general R that if f is expressible in terms of binary affines, then $r_i r_j \in B$ for $i \neq j$.

Now let I be the set of all $b \in R$ for which $X + bY - bZ$ is expressible in terms of binary affines. The preceding paragraph shows $I \subset B$. To prove the reverse inclusion, observe that the equations

$$X + rbY - rbZ = (1 - r)X + r(X + bY - bZ)$$
$$X + brY - brZ = X + b(rY + (1 - r)X) - b(rZ + (1 - r)X)$$
$$X + (b + b')Y - (b + b')Z$$
$$= (X + bY - bZ) + b'(Y + bX - bZ) - b'(bX + (1 - b)Z)$$
$$X + (r^2 - r)Y - (r^2 - r)Z$$
$$= r((r - 1)Y + (2 - r)X) + (1 - r)(rZ + (1 - r)X)$$

demonstrate that if $r \in R$ and $b \in I$, $b' \in I$ then $rb \in I$, $br \in I$, $b + b' \in I$, and $r^2 - r \in I$. Since $0 \in I$, we conclude that I is a two-sided ideal, and R/I is a ring in which every element is idempotent. Thus $B \subset I$ as desired.

Now by induction on n, for $n \geqq 3$, we show that $r_i r_j \in B$ for all pairs $i \neq j$ implies f is expressible in terms of binary affine operations. The initial and inductive steps both follow by writing

$$\sum_{i=1}^{n} r_i X_i = X + bX_1 - bX_2,$$

where

$$X = (r_1 + r_2)(r_1 X_1 + (1 - r_1)X_2) + \sum_{i=3}^{n} r_i X_i,$$

and

$$b = \left(\sum_{i=3}^{n} r_i\right) r_1.$$

This concludes the proof.

Leaving modules, what can one say about models of the affine part of $\mathbf{T}$? By Part I [**1**], they are representable as subsets of $\mathbf{T}$-algebras, closed under affine operations. Functorial semantics provides a distinguished representation and a description of the (affine) morphisms. If $I: \mathbf{A} \rightarrow \mathbf{T}$ inserts the affine part, then $(\ \ ) \otimes_{\mathbf{A}} I$, or $\otimes\ I$ for short, is the adjoint of the forgetful functor. For any $\mathbf{A}$-algebra X, $X \otimes I$ is generated by the subset which is (under) the image of the adjunction map $X \rightarrow \mathrm{Hom}(I, X \otimes I)$ (because of the universal problem which $X \otimes I$ solves). The free affine algebra P on one generator is a singleton, so there is a unique morphism $h: X \rightarrow P$. $P \otimes I$ is a free $\mathbf{T}$-algebra on one generator. $h \otimes I$ extends (lies over) h, so it takes the subset X to the generator p; but a $\mathbf{T}$-word in the elements of X taken by $h \otimes I$ to the generator is precisely the result of an idempotent operation, so we have $X = (h \otimes I)^{-1}(p)$.

We write $S_{\mathbf{T}}$ for the category of $\mathbf{T}$-algebras; and $S_{\mathbf{T}}/F(p)$ for the category of objects over $F(p)$, the free $\mathbf{T}$-algebra on one generator p. We have thus obtained an equivalence of categories between $S_{\mathbf{A}}$ and a full subcategory $\mathscr{C}$ of $S_{\mathbf{T}}/F(p)$, given by $X \mapsto (X \otimes I \to p \otimes I)$ for $S_{\mathbf{A}} \to \mathscr{C}$ and

$$\left(T \xrightarrow{\pi} F(p)\right) \mapsto \pi^{-1}(p)$$

for $\mathscr{C} \to S_{\mathbf{A}}$. What is the subcategory $\mathscr{C}$ (the essential image of $S_{\mathbf{A}} \to S_{\mathbf{T}}/F(p)$)? If T is a $\mathbf{T}$-algebra and $X \subset$ (the underlying set of) T, T is *relatively free* on X if for every $\mathbf{T}$-homomorphism φ: $T' \to T$ such that $\varphi^{-1}(X)$ generates T' and φ restricts to a bijection $\varphi^{-1}(X) \to X$, φ is necessarily an isomorphism. Then we can sum up the elementary facts:

The functors $S_{\mathbf{A}} \rightleftarrows S_{\mathbf{T}}/F(p)$ induce an equivalence of categories between $S_{\mathbf{A}}$ and the full subcategory of $S_{\mathbf{T}}/F(p)$ with objects those π: $T \to F(p)$ such that T is relatively free on $\pi^{-1}(p)$.

"Relatively free on" can be replaced by "generated by" when a congruence on a $\mathbf{T}$-algebra is determined by one of its cosets.

For the usual (clarified as we continue) varieties $S_{\mathbf{T}}$ in which that happens, a surjective π: $Y \to F$ will be generated by $\pi^{-1}(p)$, because there is a zero $0 \in F$, and Y is generated by the kernel $\pi^{-1}(0)$ and any one element of $\pi^{-1}(p)$ (an extension of F by $\pi^{-1}(0)$), *and* any element of $\pi^{-1}(0)$ is some sort of difference of two elements of $\pi^{-1}(p)$. It is not true that all $h \otimes I$: $X \otimes I \to F$ are surjective, but the only way to fail is for $X \to P$ to be non-surjective, i.e. for X to be the empty affine algebra.

We conclude with examples showing:

A. *Non-mutually interpretable affine theories can have isomorphic pointed covers.*

A′. *Finitely presented theories which are not mutually interpretable can have isomorphic pointed covers.*

B. *The affine part of a finitely presented theory need not be finitely generated.*

C. *The affine part of a finitely presented theory can be finitely generated and still not finitely related.*

For examples A, A′, we want the following observation. Let $\mathbf{T}_1$, $\mathbf{T}_2$ be elementary theories, M a model of $\mathbf{T}_2$, and f: $M \to M$ an automorphism of M. If there is an interpretation of $\mathbf{T}_1$ in $\mathbf{T}_2$, then $\mathbf{T}_1$ has a model on the same underlying set as M with an automorphism given by the same function as f.

Example A. Each of the theories $\mathbf{A}, \mathbf{B}$, is generated by infinitely many operations α^k, where α^k is k-ary. $\mathbf{A}$ has such an operation for each even $k \geqq 2$, $\mathbf{B}$ for odd $k \geqq 3$. The laws of both theories are

$$\alpha^k(x_1, x_2, \ldots, x_{k-1}, x_{k-1}) = x_{k-1};$$

$$\alpha^k(x_1, \ldots, x_{k-2}, x_k, x_{k-1}) = \alpha^k(x_1, \ldots, x_{k-2}, x_{k-1}, x_k);$$

$$\alpha^k(x_1, \ldots, x_i, x_i, x_{i+2}, \ldots, x_k) = \alpha^{k-2}(x_1, \ldots, x_{i-1}, x_{i+2}, \ldots, x_k)$$

for $1 \leqq i \leqq k - 3$.

To distinguish $\mathbf{A}$ from $\mathbf{B}$, define a two-element $\mathbf{B}$-algebra $\{0, 1\}$ as follows. $\alpha^k(x_1, \ldots, x_{k-1}, x_{k-1})$ is as required, x_{k-1}. For $x_{k-1} \neq x_k$, note that the odd number of occurrences $x_1, \ldots, x_k$ of 0 and 1 include an odd number of one of them, u, and an even number of the other; define $\alpha^k(x_1, \ldots, x_k) = u$. The laws are obvious, and symmetry is obvious. However, any $\mathbf{A}$-structure on $\{0, 1\}$ involves $\alpha^2(0, 1) = \alpha^2(1, 0)$, a definable element. Hence there is no interpretation of $\mathbf{A}$ in $\mathbf{B}$.

$\mathbf{A}^*$ and $\mathbf{B}^*$ are isomorphic. Interpret $\mathbf{A}^*$ in $\mathbf{B}^*$ by taking p to p and $\alpha^k(x_1, \ldots, x_k)$ to $\alpha^{k+1}(p, x_1, \ldots, x_k)$. Laws go to laws, so we have an interpretation. The same formulas for odd k define an interpretation of $\mathbf{B}^*$ in $\mathbf{A}^*$. Since $\alpha^{k+2}(p, p, x_1, \ldots, x_k) = \alpha^k(x_1, \ldots, x_k)$, the composites are identities.

Example A′. One of the theories $\mathbf{V}$ is nearly the theory of $\mathbf{Z}_2$-modules; but we omit the 0-ary operation. So $\mathbf{V}$ is given by a commutative associative binary operation $+$ satisfying the further (torsion) law $x + x + y = y$. The theory $\mathbf{W}$ is generated by a symmetric associative ternary operation which will be written $x \oplus y \oplus z$; symmetry means $x \oplus y \oplus z = y \oplus x \oplus z = x \oplus z \oplus y$, and associativity $(a \oplus b \oplus c) \oplus d \oplus e = a \oplus b \oplus (c \oplus d \oplus e)$. It follows that parentheses can be omitted, for there is only one operation on a, b, c, d, e which involves all of them. Finally, the torsion law of $\mathbf{W}$ is the weak one

$$x \oplus x \oplus y \oplus y \oplus z = z.$$

There is no interpretation $\varphi: \mathbf{V} \rightarrow \mathbf{W}$. For every $\mathbf{W}$-algebra has a transitive automorphism group; the functions h_{ab} taking x to $a \oplus a \oplus a \oplus b \oplus x$ are easily seen to be homomorphic and involutory and take a to b. There exist many $\mathbf{W}$-algebras, e.g. every Z_2-module (with $x \oplus y \oplus z$ defined as $x + y + z$). If φ existed it would convert those algebras into homogeneous $\mathbf{V}$-algebras, which is absurd; an automorphism of a non-empty $\mathbf{V}$-algebra fixes $x + x = (x + x + y + y) = y + y$.

However, $\mathbf{V}^*$ is easily interpreted in $\mathbf{W}^*$, p as p and $x + y$ as $p \oplus x \oplus y$. Among the interpretations of $\mathbf{W}^*$ in $\mathbf{V}^*$ use the one taking p to p and $x \oplus y \oplus z$ to $p + x + y + z$. There is no difficulty in checking that the composites are identities; so $\mathbf{V}^*$ and $\mathbf{W}^*$ are isomorphic.

Example B. Present $\mathbf{T}$ by a binary associative multiplication and a unary operation $^{(1)}$ subject to $xx^{(1)} = x$ and $(xy)^{(1)} = x^{(1)}$. We write $x^{(0)} = x, x^{(n+1)} = x^{(n)(1)}$. Observe that $x^{(n)}zx^{(n+k)} = x^{(n)}z$ for $k \geqq 1$, even for z empty (by induction; for $k = 1$, $x^{(n)}z\, x^{(n+1)} = (x^{(n)}z)(x^{(n)}z)^{(1)} = x^{(n)}z$; and if true for k, then

$x^{(n)}z\ x^{(n+k+1)} = x^{(n)}z\ x^{(n+k)}x^{(n+k+1)} = x^{(n)}z\ x^{(n+k)} = x^{(n)}z$). One readily verifies that the free $\mathbf{T}$-algebra on a set X consists of non-empty expressions $\prod x_i^{(n_i)}$ with $x_i \in X$, $n_i \geqq 0$, and when $x_i = x_j$ with $i < j$, then $n_j \leqq n_i$. For such an expression e, $e^{(1)}$ is $x_1^{(n_1+1)}$. The product $\prod x_i^{(n_i)} \prod y_j^{(m_j)}$ is the concatenation reduced by erasing any $y_j^{(m_j)}$ if y_j occurs in $\prod x_i^{(n_i)}$ with a smaller superscript.

$\prod x_i^{(n_i)}$ is affine provided $\prod x^{(n_i)} = x$, which means $n_1 = 0$ and $n_i > 0$ for all $i > 1$. So an affine word which is not a variable has the form $x_0 \prod x_i^{(n_i)}$ with the product non-empty and all $n_i > 0$. Suppose its value $w_0 \prod w_i^{(n_i)}$ on certain affine words is $x_1x_2^{(n)}$. The value is a product expression having w_0 as an initial segment; so $w_0 = x_1x_2^{(n)}$ or $w_0 = x_1$. We shall show that in the latter case as well as the former, one cannot get the superscript n unless one already had it, specifically some $n_i = n$. The concatenation of $w_0, w_1^{(n_1)}, \ldots, w_r^{(n_r)}$ is a monomial in x_1 and x_2 (because the product $x_1x_2^{(n)}$ has that form). Since $n_i > 0$, $w_i^{(n_i)}$ involves only one variable. The first time that variable is x_2, w_i is x_2 or an affine word $x_2 \prod z_j^{(m_j)}$. In either case, $n_i = n$, since $x_2^{(n_i)}$ must occur in the value $x_1x_2^{(n)}$. Thus the affine part is not finitely generated.

Example C. (The affine part to be finitely generated but not finitely related.) Present $\mathbf{T}$ by a (non-associative) idempotent multiplication and a unary operation $*$, with the laws $(xy)x = xy$, $x^*y = x^*$, and $xy^* = (xy)y$. It will be convenient to describe words inductively. $(w_1 \cdot w_2)$ will mean the concatenation of words w_1 and w_2 (not necessarily a word. Of course w_1w_2 is the product, a word whose appearance is not given by this notation.) Observe that there is no need for ordinary exponents since multiplication is idempotent. We define $w_1[w_2]^1$ to mean $(w_1 \cdot w_2)$, $w_1[w_2]^{n+1} = w_1[w_2]^n[w_2]^1$. Also (without brackets) $w^0 = w$, $w^{n+1} = (w^n)^*$.

The words in a free $\mathbf{T}$-algebra are (1) the variables, (2) w^* for any word w, and (3) expressions $(w_1 \cdot w_2)$ where w_1 and w_2 are distinct unstarred words and w_1 is not of the form $w_2[w_3]^r$ for a power of $2(\geqq 2^0)$, r. Operations are performed as follows. The star of w is w^*. The product w_1w_2 is $(w_1 \cdot w_2)$ if that is a word, w_1 if w_1 is starred or equal to w_2 or of the mentioned form $w_2[w_3]^r$. It remains to define w_1v^n, $n > 0$.

We may assume inductively that wv^{n-1} is defined for all words w; then $w_1v^n = (w_1v^{n-1})v^{n-1}$.

The last case overlaps the cases $w_1 = w_2$ and $w_1 = u^*$, but consistency is obvious. So are the laws except perhaps $(w_1w_2)w_1 = w_1w_2$. If w_1w_2 is $(w_1 \cdot w_2)$, this is $w_1[w_2]^1$, and the law holds; it certainly holds if the product is w_1; it holds by induction if w_2 is starred. Since the identifications made are clearly required by the laws, we have the free $\mathbf{T}$-algebra.

The affine part of $\mathbf{T}$ is the theory of multiplication, i.e. affine words are words written without $*$; for the free $\mathbf{T}$-algebra on one generator x consists of the words x^n, and equating the generators of any free algebra takes products to x and words w^n (w unstarred) to x^n. The rules for multiplying affine words are $ww = w$ and $w[v]^rw = w[v]^r$ for r a power of 2. It is obvious that there is a two-generator algebra satisfying any initial segment of these laws but not the rest.

REFERENCES

1. J. R. Isbell, M. I. Klun, and S. H. Schanuel, *Affine parts of algebraic theories I*, Journal of Algebra *44* (1977), 1–8.
2. A. Tarski, *Equational logic and equational theories of algebras*, Contributions to Math. Logic (Colloquium, Hanover, 1966), 275–288 (Amsterdam, 1968).

State University of New York at Buffalo
Amherst, New York

https://doi.org/10.4153/CJM-1978-021-3 Published online by Cambridge University Press