A Quasi-Affine Representation

Keith A. Kearnes

Abstract

We describe a quasi-affine representation for algebras having a binary central term.

1 Introduction

An algebra $\mathbf{A}$ is said to be **affine** if there is a module $\mathbf{M}$ with the same universe and the same polynomial operations as $\mathbf{A}$. $\mathbf{A}$ is said to be **quasi-affine** if it is a subreduct (= a subalgebra of a reduct) of an affine algebra. A homomorphism from $\mathbf{A}$ into a reduct of an affine algebra will be called a **quasi-affine representation** or just a **representation**. A representation is **faithful** if it is 1-1.

Not every algebra has a faithful quasi-affine representation, of course. A non-trivial semilattice is an example of an algebra with no non-constant quasi-affine representation. To see this one can use a direct calculation to show that it is impossible for a semilattice operation to be represented by a module polynomial. But let us get to the root of the problem: quasi-affine algebras satisfy the term condition and semilattices do not. ($\mathbf{A}$ is said to satisfy the **term condition** if whenever $s(x, \bar{y})$ is an $(n+1)$-ary term in the language of $\mathbf{A}$, $a, b \in A$ and $\bar{u}, \bar{v} \in A^n$, then the implication

$$s^{\mathbf{A}}(a, \bar{u}) = s^{\mathbf{A}}(a, \bar{v}) \to s^{\mathbf{A}}(b, \bar{u}) = s^{\mathbf{A}}(b, \bar{v})$$

holds. The groups which satisfy the term condition are precisely the abelian groups so it is common to call any algebra which satisfies the term condition an **abelian algebra** as is done in [2], for example.) Affine algebras satisfy the term condition and the term condition is inherited by subreducts. Hence the implication

$$\text{quasi-affine} \Rightarrow \text{abelian}$$

holds. To obtain a faithful quasi-affine representation of an algebra $\mathbf{A}$ it is necessary for $\mathbf{A}$ to be abelian. In general, this is not a sufficient condition. It is shown in [8] that there is a list of implication schemes generalizing and including the term condition which, when taken together, are necessary and sufficient conditions for an algebra to have a faithful quasi-affine representation. It is further shown in [8] that no finite sublist of these schemes suffices (in general). This paper is one among a few papers that describe commonly occurring situations where abelian algebras are quasi-affine. (Others are [3], [5], [6] and Chapter 13 of [4].)

If f is an m-ary operation on U and g is an n-ary operation on U, then f and g are said to **commute** if for all $m \times n$ arrays $[u_i^j]$ of elements of U:

$$\begin{bmatrix} u_1^1 & u_1^2 & \cdots & u_1^n \\ u_2^1 & u_2^2 & & \\ \vdots & & \ddots & \\ u_m^1 & & & u_m^n \end{bmatrix},$$

it is the case that $f(\overline{g(\bar{u}_i)}) = g(\overline{f(\bar{u}^j)})$ holds. If $\mathbf{A}$ is an algebra, then we define a **central** term operation for $\mathbf{A}$ to be a term operation which commutes with all the fundamental operations of $\mathbf{A}$. It is a theorem of H. -P. Gumm that an algebra is affine if and only if it has a central Mal'cev operation. A proof of this can be found in [2]. In this paper we show that central binary operations afford quasi-affine representations of abelian algebras. Our main result is the following.

THEOREM 1.1 *Let* $\mathbf{A}$ *be an abelian algebra with a central term* $t(x,y) = xy$*. Then* $\mathbf{A}$ *has a quasi-affine representation* Φ *with the property that* $\Phi(a) = \Phi(b)$ *if and only if for some* n *it is the case that*

$$(\cdots((aa)a)\cdots)a = (\cdots((ab)a)\cdots)a \quad \text{and} \quad a(\cdots(a(aa))\cdots) = a(\cdots(a(ba))\cdots)$$

where there are $n+1$ *occurrences of* a *on the left side of each equation and* n *occurrences of* a *on the right.*

We list a few of the consequences of this theorem. First, we need some definitions. The left cancellation law for a binary operation $t(x,y) = xy$ is the law:

$$au = av \rightarrow u = v$$

and the right cancellation law is defined dually. An operation satisfying the left, right or both of the cancellation laws will be called **left cancellative**, **right cancellative** or just **cancellative**, respectively.

COROLLARY 1.2 *If* $\mathbf{A}$ *is an abelian algebra with a cancellative binary central term, then* $\mathbf{A}$ *is quasi-affine.*

Proof: Using right cancellation one may reduce $(\cdots((aa)a)\cdots)a = (\cdots((ab)a)\cdots)a$ to $aa = ab$. Using left cancellation one may reduce $aa = ab$ to $a = b$. Hence the representation of Theorem 1.1 is faithful when the central term is cancellative. □

COROLLARY 1.3 *If* $\mathbf{A}$ *is an algebra which has a homomorphism* $t : \mathbf{A}^2 \rightarrow \mathbf{A}$ *where*

(i) t *commutes with itself on* $\mathbf{A}$ *and*

(ii) $\ker t$ *complements the coordinate projection kernels in* $\mathbf{Con}\mathbf{A}^2$,

then $\mathbf{A}$ *is quasi-affine.*

Proof: Expand $\mathbf{A}$ to $\mathbf{A}'$ by adding t as a new operation. If we show that $\mathbf{A}'$ is quasi-affine, then $\mathbf{A}$, as a reduct of $\mathbf{A}'$, will also be quasi-affine.

The hypotheses guarantee that t is a cancellative central binary operation. This result is a therefore a consequence of Corollary 1.2 if we can show that $\mathbf{A}'$ is abelian.

To see that $\mathbf{A}'$ is abelian assume otherwise that $f(x,\bar{y})$ is an $(n+1)$-ary term, $a, b \in A$ and $\bar{u}, \bar{v} \in A^n$ and $f^{\mathbf{A}}(a,\bar{u}) = f^{\mathbf{A}}(a,\bar{v})$ while $f^{\mathbf{A}}(b,\bar{u}) \neq f^{\mathbf{A}}(b,\bar{v})$. Set $\theta = \ker t$ and choose $r_0, \ldots, r_n \in A^2$ such that $r_0 = (a,a)$, $r_n = (b,b)$, $(r_{2i}, r_{2i+1}) \in \theta$ and $(r_{2i+1}, r_{2i+2}) \in \eta_1$. This is possible since $\theta \vee \eta_1 = 1$ in $\mathbf{Con}\mathbf{A}^2$. Now in $\mathbf{A}^2$ we have

$$f(r_0, \overline{(u_i,u_i)}) = f(r_0, \overline{(u_i,v_i)}), \quad \text{but} \quad f(r_n, \overline{(u_i,u_i)}) \neq f(r_n, \overline{(u_i,v_i)}).$$

Hence there is some j such that

$$f(r_j, \overline{(u_i,u_i)}) = f(r_j, \overline{(u_i,v_i)}) \quad \text{and} \quad f(r_{j+1}, \overline{(u_i,u_i)}) \neq f(r_{j+1}, \overline{(u_i,v_i)}).$$

If j is even, then $(r_j, r_{j+1}) \in \theta$, so in this case

$$c = f(r_{j+1}, \overline{(u_i,u_i)}) \;\theta\; f(r_j, \overline{(u_i,u_i)}) = f(r_j, \overline{(u_i,v_i)}) \;\theta\; f(r_{j+1}, \overline{(u_i,v_i)}) = d.$$

Also, we clearly have

$$c = f(r_{j+1}, \overline{(u_i,u_i)}) \;\eta_0\; f(r_{j+1}, \overline{(u_i,v_i)}) = d.$$

Hence $(c,d) \in \theta \wedge \eta_0 = 0$. This is impossible since $c \neq d$. The case where j is odd is handled the same way. One argues that $(c,d) \in \eta_1$ and $(c,d) \in \eta_0$, so $(c,d) \in \eta_0 \wedge \eta_1 = 0$; a contradiction. This finishes the proof. □

The hypotheses of Corollary 1.3 are not at all artificial. If $\mathbf{A}$ is a reduct of a module $\mathbf{M}$, then there will be homomorphisms $t : \mathbf{A}^2 \rightarrow \mathbf{A}$ satisfying hypotheses (i) & (ii) of Corollary 1.3. For example, one can take $t(x,y) = x + y$ or $t(x,y) = x - y$ where $+$ and $-$ are the addition and subtraction of $\mathbf{M}$ (they need not be operations of $\mathbf{A}$).

If $\mathbf{A}$ is an algebra defined with a single binary fundamental operation, then we call $\mathbf{A}$ a **binar**. An m-ary operation on a set U is called **idempotent** if $f(u,\ldots,u) = u$ holds for all $u \in U$. Thus a binar is idempotent if it satisfies the equation $xx = x$.

COROLLARY 1.4 *If* $\mathbf{A}$ *is an idempotent abelian binar, then* $\mathbf{A}$ *has a strongly abelian congruence* θ *such that* $\mathbf{A}/\theta$ *is quasi-affine.*

Definition 2.10 explains what the phrase "strongly abelian" means. In Corollary 1.4 the statement that θ is strongly abelian implies that the subalgebra structure on each θ-class is that of a rectangular band, but the strongly abelian property says more than this. We will not give the proof of Corollary 1.4 now, since we have deferred the explanation of what it means to be strongly abelian. (For economy, we later combine the parts of Corollary 1.4 and Theorem 1.5 which refer to the strongly abelian congruence θ into Theorem 2.11.)

If $\mathbf{A}$ is an idempotent, abelian binar, it follows that

$$(\underline{y}y)(z\underline{z}) = yz = (\underline{y}z)(y\underline{z})$$

holds. Applying the term condition to $\underline{y}$ and also to $\underline{z}$ we deduce

$$(xy)(zw) = (xz)(yw)$$

which is the statement that the term xy commutes with itself. (Change $\underline{y}$ to x with one application of the term condition and with a second application change $\underline{z}$ to w.) Thus, any idempotent binar for which a faithful quasi-affine representation exists satisfies $(xy)(zw) = (xz)(yw)$. This shows that the fundamental operation of an idempotent abelian binar is central.

THEOREM 1.5 *If* $\mathbf{A}$ *is a right cancellative binar which satisfies*

(i) $xx = x$ *(Idempotent Law),*

(ii) $(xy)(zu) = (xz)(yu)$ *(Entropic Law),*

then $\mathbf{A}$ *has a strongly abelian congruence* θ *such that* $\mathbf{A}/\theta$ *is quasi-affine and each* θ*-class is the universe of a left-zero subsemigroup of* $\mathbf{A}$.

The important point to notice in Theorem 1.5 is that it applies to non-abelian algebras. (In the case that $\mathbf{A}$ is abelian we can drop the assumption that the entropic law holds, since this is a consequence of the idempotent law. In this situation Theorem 1.5 is a consequence of Theorem 1.1.) The representation in Theorem 1.5 is a slight modification of representation in Theorem 1.1. To see that the hypotheses of Theorem 1.5 can be satisfied by non-abelian algebras, we describe a non-abelian, right cancellative binar satisfying Theorem 1.5 (i)&(ii) (the smallest example, in fact). Let $\mathbf{A}$ be the binar with operation table

	0	1	2
0	0	0	1
1	1	1	0
2	2	2	2

$\mathbf{A}$ is a non-abelian binar ($\underline{2}0 = \underline{2}2 \rightarrow 00 = 02$ is false) which is right cancellative and satisfies the idempotent and entropic laws. It has no faithful quasi-affine representation, since it is non-abelian, but it has only one non-zero congruence whose classes are left-zero semigroups: the congruence $\mathrm{Cg}(0,1)$. It follows from Theorem 1.5 that $\theta = \mathrm{Cg}(0,1)$ is strongly abelian and that $\mathbf{A}/\theta$ is quasi-affine.

Since Theorem 1.5 applies to non-abelian algebras, it is clear that the cancellativity hypothesis is not extraneous. Any semilattice is an idempotent, entropic binar, but only the 1-element semilattice satisfies the conclusion of Theorem 1.5.

One unfortunate feature of Theorem 1.1 is that the quasi-affine representation obtained may not be faithful even when a faithful representation exists. In Corollary 1.4, for example, we do not always obtain a faithful representation; and yet, we are not aware of a single idempotent abelian algebra which is not quasi-affine. ([5] contains a proof that every *simple* idempotent abelian algebra *is* quasi-affine.) It is easily seen that a rectangular band is quasi-affine, but our representation theorem discovers only the constant representation for any rectangular band. It seems that this is a difficulty that cannot be avoided by any general affinization method. Rectangular bands (for example) are quasi-affine, but they

do not have enough structure to recover even a small part of any representing affine algebra. One way to improve the fidelity of the representation we describe in this paper is to choose a collection of binary central terms, $\{t_i(x,y) \mid i \in I\}$, and obtain a representation ϕ_i for each one. Then the product of these representations leads to a single representation whose kernel is $\bigcap_{i\in I} \theta_i$ where θ_i is the kernel of ϕ_i.

The theorems we prove belong to the classical genre of embeddability theorems. E.g., Ore domains can be embedded into division rings, cancellative semigroups satisfying a non-trivial equation can be embedded into groups, cancellative binars can be embedded into quasigroups. What motivates these results is the desire to extend a given algebra to a "nicer sort" of algebra; one in which calculation is easier. For example, problems in number theory are ostensibly problems about the rational integers. But solutions of these problems often require the fact that the rational integers can be extended to number rings where certain polynomial equations have solutions. Similarly, some problems concerning semigroups or rings are solved by embedding the algebra into a group or division ring respectively and then using the inverse operation in calculations. (This is not possible for every semigroup or ring, of course.) What motivates us is the desire to express the term and polynomial operations of an abelian algebra as affine combinations of variables to facilitate calculation. For an excellent example of the usefulness of this, the reader should peruse [3]. In this fundamental paper, abelian algebras in congruence modular varieties are shown to be affine in two steps. First, they are shown to be quasi-affine. Then, expressing Day terms as linear combinations of variables, it is shown that a central Mal'cev operation can be constructed by composition from these terms.

2 The Representation

Our task is to construct a quasi-affine representation for $\mathbf{A}$ from the properties of a fixed binary central term. We shall do this as follows. We shall show how to

(i) Construct an abelian group $\mathbf{H}$.

(ii) Define a collection P of endomorphisms of $\mathbf{H}$.

(iii) Define a function $\phi : A \to H$.

(iv) For each fundamental operation f of $\mathbf{A}$, define an affine function $(\Sigma_{i=1}^n \hat{f}_i(x_i)) + h$ on $\mathbf{H}$ where each $\hat{f}_i \in P$ and $h \in H$.

(v) Show that

$$\phi(f^{\mathbf{A}}(x_1,\ldots,x_n)) = (\Sigma_{i=1}^n \hat{f}_i(\phi(x_i))) + h$$

Thus if we let $\widehat{\mathbf{H}}$ denote the algebra with universe H and operations $(\Sigma_{i=1}^n \hat{f}_i(\phi(x_i))) + h$, one such operation for each fundamental operation f of $\mathbf{A}$, then $\widehat{\mathbf{H}}$ is a reduct of the affine algebra obtained from adjoining to $\mathbf{H}$ all constant operations and all members of P as new unary operations. By (v), the function ϕ is a homomorphism of $\mathbf{A}$ into $\widehat{\mathbf{H}}$. This will complete the representation. After this, all that remains is to analyze the kernel of ϕ in order to establish Theorem 1.1. Throughout this section we will assume that $\mathbf{A}$ is an algebra which has a binary central term $t(x,y)$. We will not assume that $\mathbf{A}$ is abelian without stating this assumption.

Before getting into the proof we prove a necessary lemma concerning binary term operations that commute with themselves.

LEMMA 2.1 *Let xy be a term operation of $\mathbf{A}$ which commutes with itself. The following equation is true.*

$$\{[(xb)(da)][(ac)(aa)]\}\{[(cb)(cc)][(aa)(aa)]\} = \{[(xd)(bb)][(cc)(cc)]\}\{[(aa)(aa)][(aa)(aa)]\}.$$

If $\mathbf{A}$ is abelian and $f(x_1,\ldots,x_n)$ is an n-ary term operation of $\mathbf{A}$, then the following equation is true.

$$\begin{aligned}[(xf(a_1,\ldots,a_k,a_{k+1},u,\ldots,u))(f(u,\ldots,u)f(u,\ldots,u))] = \\ [(xf(a_1,\ldots,a_k,u,\ldots,u))(f(u,\ldots,u,a_{k+1},u,\ldots,u)f(u,\ldots,u))].\end{aligned}$$

Proof: In the following derivation we shall write $(p\underline{q})(\underline{r}s)$ to indicate that we plan to use the entropic law in the next step to infer from $(pq)(rs)$ that $(pr)\overline{(qs)}$ holds. If E is the expression

$$\{[(xb)(da)][(ac)(aa)]\}\{[(cb)(cc)][(aa)(aa)]\}$$

and F is the expression

$$\{[(xd)(bb)][(cc)(cc)]\}\{[(aa)(aa)][(aa)(aa)]\},$$

then we calculate that

$$\begin{aligned}
E &= \{[(xb)(da)][(ac)(aa)]\}\{[(cb)(cc)][(aa)(aa)]\} \\
&= \{[(x\underline{b})(\underline{d}a)][(a\underline{c})(\underline{a}a)]\}\{[(c\underline{b})(\underline{c}c)][(aa)(aa)]\} \\
&= \{\overline{[(xd)(ba)][(aa)(ca)]}\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{[(xd)\underline{(ba)}][(aa)\underline{(ca)}]\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{\overline{[(xd)(aa)][(ba)(ca)]}\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{[(xd)(aa)][(b\underline{a})(\underline{c}a)]\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{[(xd)(aa)]\overline{[(bc)(aa)]}\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{[(xd)\underline{(aa)}][\underline{(bc)}(aa)]\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{\overline{[(xd)(bc)][(aa)(aa)]}\}\{[(cc)(bc)][(aa)(aa)]\} \\
&= \{[(xd)(bc)]\underline{[(aa)(aa)]}\}\{\underline{[(cc)(bc)]}[(aa)(aa)]\} \\
&= \overline{\{[(xd)(bc)][(cc)(bc)]\}\{[(aa)(aa)][(aa)(aa)]\}} \\
&= \{[(xd)\underline{(bc)}]\underline{[(cc)}(bc)]\}\{[(aa)(aa)][(aa)(aa)]\} \\
&= \{\overline{[(xd)(cc)][(bc)(bc)]}\}\{[(aa)(aa)][(aa)(aa)]\} \\
&= \{[(xd)(cc)][(b\underline{c})(\underline{b}c)]\}\{[(aa)(aa)][(aa)(aa)]\} \\
&= \{[(xd)(cc)]\overline{[(bb)(cc)]}\}\{[(aa)(aa)][(aa)(aa)]\} \\
&= \{[(xd)\underline{(cc)}]\underline{[(bb)}(cc)]\}\{[(aa)(aa)][(aa)(aa)]\} \\
&= \{\overline{[(xd)(bb)][(cc)(cc)]}\}\{[(aa)(aa)][(aa)(aa)]\} \\
&= F.
\end{aligned}$$

For the second equation of the lemma, note that

$$\begin{aligned}
&[(xf(u,\ldots,u,a_{k+1},u,\ldots,u))(f(u,\ldots,u,u,u,\ldots,u)f(u,\ldots,u))] = \\
&\qquad [(xf(u,\ldots,u,u,u,\ldots,u))(f(u,\ldots,u,a_{k+1},u,\ldots,u)f(u,\ldots,u))]
\end{aligned}$$

is a consequence of the fact that xy commutes with itself. Now we use the term condition to change $\underline{u}$ to a_i in the i^{th} position of the left-most $f(-)$. We change

$$\begin{aligned}
&[(xf(\underline{u},\ldots,\underline{u},a_{k+1},u,\ldots,u))(f(u,\ldots,u,u,u,\ldots,u)f(u,\ldots,u))] = \\
&\qquad [(xf(\underline{u},\ldots,\underline{u},u,\ldots,u))(f(u,\ldots,u,a_{k+1},u,\ldots,u)f(u,\ldots,u))]
\end{aligned}$$

to

$$\begin{aligned}
&[(xf(a_1,\ldots,a_k,a_{k+1},u,\ldots,u))(f(u,\ldots,u)f(u,\ldots,u))] = \\
&\qquad [(xf(a_1,\ldots,a_k,u,u,\ldots,u))(f(u,\ldots,u,a_{k+1},u\ldots,u)f(u,\ldots,u))]
\end{aligned}$$

which was the equality to be established. This finishes the proof. □

If $\mathbf{A}$ is an algebra with a binary central term $t(x,y) = xy$, then we define a semigroup $\mathbf{S}$ associated to $\mathbf{A}$ as follows. For each $a \in A$ let $\rho_a : A \to A$ be the function $x \mapsto xa$. We shall refer to ρ_a as "right translation by a". Let $\mathbf{S}$ be the semigroup of self-maps of A generated by $R = \{\rho_a | a \in A\}$. We compose right translations on the right, so $\rho_a\rho_b$ denotes the polynomial $(xa)b$.

If $\mathbf{T}$ is a semigroup and $a, b \in T$, then we define $a < b$ if a is a proper left divisor of b. This means that $a < b$ if and only if $b = ac$ for some $c \in T$. We call $\mathbf{T}$ a **directed** semigroup if for any $a, b \in T$ there is a $c \in T$ such that $a < c$ and $b < c$. ($a < a$ is allowed if $a = ac$ for some $c \in T$.)

LEMMA 2.2 $\mathbf{S}$ *is directed.*

Proof: First let's examine multiplication in $\mathbf{S}$. The fact that xy commutes with all fundamental operation implies that it commutes with itself. This may be written as

$$(xu)(vw) = (xv)(uw)$$

or more suggestively as

$$\rho_u\rho_{vw} = \rho_v\rho_{uw}.$$

When $v = w$ this may be written as

$$\rho_u\rho_{vv} = \rho_v\rho_{uv}.$$

Hence for any $\rho_u, \rho_v \in R$ we have $\rho_u < \rho_u\rho_{vv}$ and $\rho_v < \rho_u\rho_{vv}$. Now suppose that $w = \rho_{v_1}\rho_{v_2}\cdots\rho_{v_k} \in S$. An easy induction on k, using equalities of the type $\rho_u\rho_{vv} = \rho_v\rho_{uv}$, establishes that for any $\rho_u \in R$ we have

$$\rho_u\rho_{v_1v_1}\rho_{v_2v_2}\cdots\rho_{v_kv_k} = \rho_{v_1}\rho_{v_2}\cdots\rho_{v_k}\rho_c = w\rho_c$$

where $c = (\cdots((uv_1)v_2)\cdots)v_k$. Hence, for any $\rho_u \in R$ and $w \in S$ we have $\rho_u < \rho_u\rho_{v_1v_1}\rho_{v_2v_2}\cdots\rho_{v_kv_k}$ and $w < \rho_u\rho_{v_1v_1}\rho_{v_2v_2}\cdots\rho_{v_kv_k}$.

If we verify the following statement, then invoking induction finishes the proof.

Claim. For $w_1, w_2 \in S$ and $\rho_v \in R$, if there is a $y \in S$ such that $w_1 < y$ and $w_2 < y$, then there is a $z \in S$ such that $w_1 < z$ and $w_2\rho_v < z$.

(If we let $\ell(w)$ denote the length of a minimal representation of w as a product of elements of R, then the argument can be completed using the above claim by induction on $\ell(w_1) + \ell(w_2)$.) In the last paragraph we proved the claim for $w_1 \in R$. Now suppose that $w_1 < y$ and $w_2 < y$. This means that $w_1r = w_2s$ for some $r, s \in S$. Also, by the first paragraph of the proof, $\rho_v < x$ and $s < x$ for some $x \in S$. Hence there exist $p, q \in S$ such that $\rho_v p = sq$. Now we have

$$\begin{aligned} w_1(rq) &= (w_1r)q \\ &= (w_2s)q \\ &= w_2(sq) \\ &= w_2(\rho_v p) \\ &= (w_2\rho_v)p \end{aligned}$$

which proves that $w_1 < w_1rq$ and $w_2\rho_v < w_1rq$. Hence, the claim holds. The proof of the lemma is complete. □

One usually says that a group $\mathbf{G}$ is a **group of fractions** for a semigroup $\mathbf{T}$ if $\mathbf{G}$ contains $\mathbf{T}$ as a subsemigroup and $\mathbf{G}$ is generated as a group by T. We shall say that a group $\mathbf{G}$ is a **group of simple fractions** for $\mathbf{T}$ if $\mathbf{G}$ contains $\mathbf{T}$ and every $g \in G$ is expressible as $g = ab^{-1}$ for some $a, b \in T$. For $\mathbf{T}$ to have a group of simple fractions it is clearly necessary for $\mathbf{T}$ to have a group of fractions, but this condition is not sufficient. The characterization of semigroups which have a group of simple fractions is the following classical result:

THEOREM 2.3 [1] *A semigroup* $\mathbf{T}$ *has a group of simple fractions if and only if* $\mathbf{T}$ *is cancellative and directed.* □

Incidentally, our definition of *directed* is not left-right symmetric; neither is our definition of a simple fraction ab^{-1}. If we define directedness in terms of right divisibility and write simple fractions as $b^{-1}a$, then we get a theorem dual to Theorem 2.3 but not equivalent to it.

The semigroup $\mathbf{S}$ is directed, but it may not cancellative. But any cancellative homomorphic image of $\mathbf{S}$ will be both cancellative and directed.

LEMMA 2.4 *The relation* $\equiv$ *on* $\mathbf{S}$ *defined by* $a \equiv b$ *if and only if* $aw = bw$ *for some* $w \in S$ *is a congruence.* $\mathbf{S}/\equiv$ *is right cancellative and directed. If* $\mathbf{A}$ *is abelian or* $t^{\mathbf{A}}(x, y)$ *is idempotent, then* $\mathbf{S}/\equiv$ *is also left cancellative. If, in addition,* $t^{\mathbf{A}}(x, y)$ *is right cancellative, then* $\equiv$ *is equality and* $\mathbf{S}$ *is cancellative and directed.*

Proof: The relation $\equiv$ is clearly reflexive and symmetric. To see that it is transitive, assume that $a \equiv b$ and $b \equiv c$. Then $au = bu$ and $bv = cv$ for some $u, v \in S$. Since $\mathbf{S}$ is directed, there is a $w \in S$ such that $u < w$ and $v < w$. Say that $ur = w = vs$. Then

$$aw = aur = bur = bvs = cvs = cw,$$

so $a \equiv c$. Thus $\equiv$ is an equivalence relation on $\mathbf{S}$. Whenever $a, b, c \in S$ we have

$$\begin{aligned} a \equiv b &\leftrightarrow (\exists w) aw = bw \\ &\rightarrow (\exists w) caw = cbw \\ &\leftrightarrow ca \equiv cb. \end{aligned}$$

That is, $a \equiv b \rightarrow ca \equiv cb$. This proves that $\equiv$ is a left congruence on $\mathbf{S}$. To see that it is a right congruence, assume that $a, b, c \in S$ and $a \equiv b$. Then there is some $u \in S$ such that $au = bu$. Choose $w \in S$ such that $c < w$ and $u < w$. Say that $cr = w = us$. Then

$$(ac)r = aus = bus = (bc)r,$$

so $ac \equiv bc$. Thus $\equiv$ is a congruence on $\mathbf{S}$.

We can prove that $\mathbf{S}/\equiv$ is right cancellative by showing that

$$ac \equiv bc \rightarrow a \equiv b.$$

Now if $ac \equiv bc$, then there is a $u \in S$ such that $acu = bcu$. But now with $w = cu$ we have $aw = bw$, so $a \equiv b$. The proof of the lemma will be complete if we show that $\mathbf{S}/\equiv$ is left cancellative. This part of the argument depends on more than just the fact that $\mathbf{S}$ is directed. We will argue separately the cases (i) $\mathbf{A}$ is abelian and (ii) $t^{\mathbf{A}}(x, y)$ is idempotent.

We must show that for $a, b, c \in S$ we have

$$ca \equiv cb \rightarrow a \equiv b.$$

Assume that $\mathbf{A}$ is abelian. Since $ca \equiv cb$ we have that $caw = cbw$ for some $w \in S$. Hence $(caw)(cbw) = (cbw)(caw)$. Write each of the elements a, b, c and w as products of elements of R. Say $a = \rho_{a_1} \cdots \rho_{a_k}$, $b = \rho_{b_1} \cdots \rho_{b_\ell}$, $c = \rho_{c_1} \cdots \rho_{c_m}$ and $w = \rho_{w_1} \cdots \rho_{w_n}$ where $a_i, b_i, c_i, w_i \in A$. We have that

$$caw = \rho_{c_1} \cdots \rho_{c_m} \rho_{a_1} \cdots \rho_{a_k} \rho_{w_1} \cdots \rho_{w_n}$$

and that cbw can be similarly expressed. The polynomial $(caw)(cbw)$ is of the form

$$(\cdots((\cdots((\cdots((\cdots((\cdots((\cdots(xc_1)c_2\cdots)a_1)a_2\cdots)w_1)w_2\cdots)c_1)c_2\cdots)b_1)b_2\cdots)w_1)w_2\cdots)w_n$$

and so is $(cbw)(caw)$. That is, $cawcbw$ and $cbwcaw$ are polynomials obtained from the same term with different parameters. We may apply the term condition to the equality $\underline{c}awcbw = \underline{c}bwcaw$ to change the subexpression $(((xc_1)\cdots)c_m)$ which occurs in the same position on both sides to the new expression x. This yields that $awcbw = bwcaw$. Now we use the fact that $caw = cbw$ to obtain

$$awcbw = bw(caw) = bw(cbw).$$

But for $v = wcbw \in S$ this is just the statement $av = bv$. Hence $a \equiv b$. We have shown that $ca \equiv cb \rightarrow a \equiv b$, so $\mathbf{S}/\equiv$ is left cancellative. (We remark that what makes this argument work is that $cawcbw$ and $cbwcaw$ are polynomials derived from the same term with different parameters. Hence the term condition may be applied in $\mathbf{A}$ to the equality $cawcbw = cbwcaw$. This may not be the case with the equality $caw = cbw$ when a and b happen to be different-length products of elements of R.)

Now assume that $t^{\mathbf{A}}(x, y) = xy$ is idempotent, but that $\mathbf{A}$ is not necessarily abelian. We would like to show that $ca \equiv cb \rightarrow a \equiv b$. It will be enough to show that

$$\rho_u a \equiv \rho_u b \rightarrow a \equiv b,$$

since if $c = \rho_{c_1} \cdots \rho_{c_m}$, then we may use the displayed rule to show for each i, $1 \leq i \leq m$, that

$$\rho_{c_i} \cdots \rho_{c_m} a \equiv \rho_{c_i} \cdots \rho_{c_m} b \rightarrow \rho_{c_{i+1}} \cdots \rho_{c_m} a \equiv \rho_{c_{i+1}} \cdots \rho_{c_m} b.$$

Taken together, these imply that $ca \equiv cb \rightarrow a \equiv b$. Assuming that $\rho_u a \equiv \rho_u b$ we get that $\rho_u aw = \rho_u bw$ for some $w \in S$. Let $f(x) = aw \in \mathrm{Pol}_1\mathbf{A}$ and let $g(x) = bw \in \mathrm{Pol}_1\mathbf{A}$. Notice that $\rho_u aw$ is just $f(xu)$ and that $\rho_u bw$ is just $g(xu)$. We have assumed that $f(xu) = g(xu)$ holds. Now $t^{\mathbf{A}}(x, y)$ is idempotent, so it commutes with every constant operation. $t^{\mathbf{A}}(x, y)$ is also central, so it commutes with every term operation. Together these facts show that $t^{\mathbf{A}}(x, y)$ commutes with every polynomial operation of $\mathbf{A}$. In particular,

$$aw\rho_{f(u)} = f(x)f(u) = f(xu) = g(xu) = g(x)g(u) = bw\rho_{g(u)}.$$

$f(xu) = g(xu)$ for all values of x and in particular for $x = u$, so $f(u) = f(uu) = g(uu) = g(u)$. Thus $\rho_{f(u)} = \rho_{g(u)}$ and we conclude that for $v = w\rho_{f(u)}$ we have $av = bv$. Hence $a \equiv b$ is a consequence of $\rho_u a \equiv \rho_u b$ as promised.

The final remark of the lemma is that if $t^{\mathbf{A}}(x, y)$ is right cancellative, then $\equiv$ is equality. This follows from the fact that

$$a \equiv b \leftrightarrow (\exists w)aw = bw \leftrightarrow a = b$$

where the second bi-implication is the statement of right cancellativity. □

The congruence $\equiv$ has an especially simple description when $\mathbf{A}$ is abelian. We have $a \equiv b$ if and only if there is some $w \in S$ such that $aw = bw$. Assume that $w = \rho_{w_1} \cdots \rho_{w_k}$ for some $w_1, \ldots, w_k \in A$. If $a = f(x) \in \mathrm{Pol}_1\mathbf{A}$ and $b = g(x) \in \mathrm{Pol}_1\mathbf{A}$, then the equality $aw = bw$ of polynomials may be written

$$(\cdots((f(x)\underline{w_1})\underline{w_2})\cdots)\underline{w_k} = (\cdots((g(x)\underline{w_1})\underline{w_2})\cdots)\underline{w_k}.$$

Using the term condition in $\mathbf{A}$ to change each $\underline{w_i}$ to a fixed $v \in A$, we obtain

$$(\cdots((f(x)v)v)\cdots)v = (\cdots((g(x)v)v)\cdots)v$$

or just $a\rho_v^n = b\rho_v^n$. Hence $a \equiv b$ is equivalent to $a\rho_v^n = b\rho_v^n$ for some n and any $v \in A$ (independent of n) when $\mathbf{A}$ is abelian.

We get from Theorem 2.3 that $\mathbf{S}/\equiv$ has a group of simple fractions which we shall label $\mathbf{G}$. For $w \in S$ we let $\tilde{w}$ denote $w/\equiv$. Let $\mathbf{H}$ denote the subgroup of $\mathbf{G}$ generated by $Q = \{\tilde{\rho}_a\tilde{\rho}_b^{-1} | a, b \in A\}$.

LEMMA 2.5 $\mathbf{H}$ *is abelian.*

Proof: Our argument will be based only on the fact that xy commutes with itself. Since $\mathbf{H}$ is the subgroup of $\mathbf{G}$ generated by $Q = \{\tilde{\rho}_a\tilde{\rho}_b^{-1} | a, b \in A\}$, we need only to prove that any element of the form $\tilde{\rho}_a\tilde{\rho}_b^{-1}$ commutes with any element of the form $\tilde{\rho}_c\tilde{\rho}_d^{-1}$. Our goal is to remove the '?' over the equality in

$$1 \stackrel{?}{=} [\tilde{\rho}_a\tilde{\rho}_b^{-1}, \tilde{\rho}_c\tilde{\rho}_d^{-1}] \quad (:= \tilde{\rho}_b\tilde{\rho}_a^{-1}\tilde{\rho}_d\tilde{\rho}_c^{-1}\tilde{\rho}_a\tilde{\rho}_b^{-1}\tilde{\rho}_c\tilde{\rho}_d^{-1}).$$

We shall simplify the expression which is farthest right by a series of reversible maneuvers. This will lead to an equivalent equality. Then we shall prove directly that the new equality holds. Our tool will be the fact that $\tilde{\rho}_v^{-1}\tilde{\rho}_u = \tilde{\rho}_{uv}\tilde{\rho}_{vv}^{-1}$, a consequence of $\rho_u\rho_{vv} = \rho_v\rho_{uv}$ which we established in the proof of Lemma 2.2.

$$\begin{aligned}
1 &\stackrel{?}{=} \tilde{\rho}_b(\tilde{\rho}_a^{-1}\tilde{\rho}_d)(\tilde{\rho}_c^{-1}\tilde{\rho}_a)(\tilde{\rho}_b^{-1}\tilde{\rho}_c)\tilde{\rho}_d^{-1} \\
&= \tilde{\rho}_b(\tilde{\rho}_{da}\tilde{\rho}_{aa}^{-1})(\tilde{\rho}_{ac}\tilde{\rho}_{cc}^{-1})(\tilde{\rho}_{cb}\tilde{\rho}_{bb}^{-1})\tilde{\rho}_d^{-1} \\
&= \tilde{\rho}_b\tilde{\rho}_{da}(\tilde{\rho}_{aa}^{-1}\tilde{\rho}_{ac})(\tilde{\rho}_{cc}^{-1}\tilde{\rho}_{cb})\tilde{\rho}_{bb}^{-1}\tilde{\rho}_d^{-1} \\
&= \tilde{\rho}_b\tilde{\rho}_{da}(\tilde{\rho}_{(ac)(aa)}\tilde{\rho}_{(aa)(aa)}^{-1})(\tilde{\rho}_{(cb)(cc)}\tilde{\rho}_{(cc)(cc)}^{-1})\tilde{\rho}_{bb}^{-1}\tilde{\rho}_d^{-1} \\
&= \tilde{\rho}_b\tilde{\rho}_{da}\tilde{\rho}_{(ac)(aa)}(\tilde{\rho}_{(aa)(aa)}^{-1}\tilde{\rho}_{(cb)(cc)})\tilde{\rho}_{(cc)(cc)}^{-1}\tilde{\rho}_{bb}^{-1}\tilde{\rho}_d^{-1} \\
&= \tilde{\rho}_b\tilde{\rho}_{da}\tilde{\rho}_{(ac)(aa)}(\tilde{\rho}_{[(cb)(cc)][(aa)(aa)]}\tilde{\rho}_{[(aa)(aa)][(aa)(aa)]}^{-1})\tilde{\rho}_{(cc)(cc)}^{-1}\tilde{\rho}_{bb}^{-1}\tilde{\rho}_d^{-1} \\
&= \tilde{\rho}_b\tilde{\rho}_{da}\tilde{\rho}_{(ac)(aa)}\tilde{\rho}_{[(cb)(cc)][(aa)(aa)]}\tilde{\rho}_{[(aa)(aa)][(aa)(aa)]}^{-1}\tilde{\rho}_{(cc)(cc)}^{-1}\tilde{\rho}_{bb}^{-1}\tilde{\rho}_d^{-1}.
\end{aligned}$$

The questionable equality is therefore equivalent to

$$\tilde{\rho}_b\tilde{\rho}_{da}\tilde{\rho}_{(ac)(aa)}\tilde{\rho}_{[(cb)(cc)][(aa)(aa)]} \stackrel{?}{=} \tilde{\rho}_d\tilde{\rho}_{bb}\tilde{\rho}_{(cc)(cc)}\tilde{\rho}_{[(aa)(aa)][(aa)(aa)]}$$

which may be written in terms of elements of $\mathbf{S}$ as

$$\rho_b\rho_{da}\rho_{(ac)(aa)}\rho_{[(cb)(cc)][(aa)(aa)]} \stackrel{?}{\equiv} \rho_d\rho_{bb}\rho_{(cc)(cc)}\rho_{[(aa)(aa)][(aa)(aa)]}.$$

But in fact it is true that in $\mathbf{S}$ we have

$$\rho_b\rho_{da}\rho_{(ac)(aa)}\rho_{[(cb)(cc)][(aa)(aa)]} = \rho_d\rho_{bb}\rho_{(cc)(cc)}\rho_{[(aa)(aa)][(aa)(aa)]}$$

since this is a restatement of the equation

$$\{[(xb)(da)][(ac)(aa)]\}\{[(cb)(cc)][(aa)(aa)]\} = \{[(xd)(bb)][(cc)(cc)]\}\{[(aa)(aa)][(aa)(aa)]\}$$

which we proved in Lemma 2.1. This proves that $\mathbf{H}$ is abelian. □

LEMMA 2.6 *Assume that* $\mathbf{A}$ *is abelian. If* $r(x,\bar{y}) \in \mathrm{Clo}_{n+1}\langle A; xy\rangle$, $g(x) \in \mathrm{Pol}_1\mathbf{A}$ *and* $\bar{a}, \bar{b} \in A^n$, *then the following implication holds for any* $x \in A$.

$$r^{\mathbf{A}}(x, a_1, \ldots, a_n) = r^{\mathbf{A}}(x, b_1, \ldots, b_n) \to r^{\mathbf{A}}(x, g(a_1), \ldots, g(a_n)) = r^{\mathbf{A}}(x, g(b_1), \ldots, g(b_n)).$$

Proof: Express $g(x)$ as $g(x) = s^{\mathbf{A}}(x, c_1, \ldots, c_k)$ for some $s \in \mathrm{Clo}_{k+1}\mathbf{A}$ and $\bar{c} \in A^k$. The fact that xy is central implies that r is central, so r and s commute on each of the matrices

$$\begin{bmatrix} x & c_1 & \cdots & c_k \\ a_1 & c_1 & \cdots & c_k \\ a_2 & c_1 & \cdots & c_k \\ \vdots & \vdots & \ddots & \vdots \\ a_n & c_1 & \cdots & c_k \end{bmatrix}, \begin{bmatrix} x & c_1 & \cdots & c_k \\ b_1 & c_1 & \cdots & c_k \\ b_2 & c_1 & \cdots & c_k \\ \vdots & \vdots & \ddots & \vdots \\ b_n & c_1 & \cdots & c_k \end{bmatrix}.$$

The fact that r and s commute on the left matrix may be expressed as

$$r^{\mathbf{A}}(g(x), g(a_1), \ldots, g(a_n)) = s^{\mathbf{A}}(r^{\mathbf{A}}(x, \bar{a}), r^{\mathbf{A}}(c_1, \bar{c_1}), \ldots, r^{\mathbf{A}}(c_k, \bar{c_k})).$$

Here $\bar{c_i}$ is the n-tuple $(c_i, c_i, \ldots, c_i)$. The fact that r and s commute on the right matrix from above may be expressed as

$$r^{\mathbf{A}}(g(x), g(b_1), \ldots, g(b_n)) = s^{\mathbf{A}}(r^{\mathbf{A}}(x, \bar{b}), r^{\mathbf{A}}(c_1, \bar{c_1}), \ldots, r^{\mathbf{A}}(c_k, \bar{c_k})).$$

If $r^{\mathbf{A}}(x, \bar{a}) = r^{\mathbf{A}}(x, \bar{b})$ for some specific value of x, then the right-hand sides of the last two displayed equations are equal. Hence for this value of x we have

$$r^{\mathbf{A}}(\underline{g(x)}, g(a_1), \ldots, g(a_n)) = r^{\mathbf{A}}(\underline{g(x)}, g(b_1), \ldots, g(b_n)).$$

Using the term condition in $\mathbf{A}$ we may change $\underline{g(x)}$ to x and obtain

$$r^{\mathbf{A}}(x, g(a_1), \ldots, g(a_n)) = r^{\mathbf{A}}(x, g(b_1), \ldots, g(b_n))$$

for this value of x. This proves the lemma. □

LEMMA 2.7 *Assume that* $\mathbf{A}$ *is abelian. If* $g \in \mathrm{Pol}_1\mathbf{A}$, *then the function*

$$\hat{g} : Q \to Q : \tilde{\rho}_a\tilde{\rho}_b^{-1} \mapsto \tilde{\rho}_{g(a)}\tilde{\rho}_{g(b)}^{-1}$$

is well-defined and extends uniquely to an endomorphism of $\mathbf{H}$.

Proof: First we prove that $\hat{g}$ is well-defined. For this we must show that if $\tilde{\rho}_a\tilde{\rho}_b^{-1} = \tilde{\rho}_c\tilde{\rho}_d^{-1}$, then $\tilde{\rho}_{g(a)}\tilde{\rho}_{g(b)}^{-1} = \tilde{\rho}_{g(c)}\tilde{\rho}_{g(d)}^{-1}$. For some n and any $v \in A$ (independent of n) the following is true.

$$\begin{aligned}
\tilde{\rho}_a\tilde{\rho}_b^{-1} = \tilde{\rho}_c\tilde{\rho}_d^{-1} \quad &\leftrightarrow \tilde{\rho}_a\tilde{\rho}_b^{-1}\tilde{\rho}_d\tilde{\rho}_c^{-1} = 1 \\
&\leftrightarrow \tilde{\rho}_a(\tilde{\rho}_b^{-1}\tilde{\rho}_d)\tilde{\rho}_c^{-1} = 1 \\
&\leftrightarrow \tilde{\rho}_a(\tilde{\rho}_{db}\tilde{\rho}_{bb}^{-1})\tilde{\rho}_c^{-1} = 1 \\
&\leftrightarrow \tilde{\rho}_a\tilde{\rho}_{db} = \tilde{\rho}_c\tilde{\rho}_{bb} \\
&\leftrightarrow \rho_a\rho_{db} \equiv \rho_c\rho_{bb} \\
&\leftrightarrow \rho_a\rho_{db}\rho_v^n = \rho_c\rho_{bb}\rho_v^n.
\end{aligned}$$

(We used $\tilde{\rho}_b^{-1}\tilde{\rho}_d = \tilde{\rho}_{db}\tilde{\rho}_{bb}^{-1}$ which follows from the fact, established in Lemma 2.2, that $\tilde{\rho}_d\tilde{\rho}_{bb} = \tilde{\rho}_b\tilde{\rho}_{db}$.) The last equality may be written as $(\cdots([(xa)(db)]v)\cdots)v = (\cdots([(xc)(bb)]v)\cdots)v$. The same argument with $g(u)$ substituted for u for each occurrence of $u \in \{a, b, c, d, v\}$ yields that $\tilde{\rho}_{g(a)}\tilde{\rho}_{g(b)}^{-1} = \tilde{\rho}_{g(c)}\tilde{\rho}_{g(d)}^{-1}$ is equivalent to $(\cdots([(xg(a))(g(d)g(b))]g(v))\cdots)g(v) = (\cdots([(xg(c))(g(b)g(b))]g(v))\cdots)g(v)$. Hence to show that $\hat{g}$ is well-defined, we must verify that

$$\begin{aligned}
&(\cdots([(xa)(db)]v)\cdots)v = (\cdots([(xc)(bb)]v)\cdots)v \rightarrow \\
&\quad (\cdots([(xg(a))(g(d)g(b))]g(v))\cdots)g(v) = (\cdots([(xg(c))(g(b)g(b))]g(v))\cdots)g(v).
\end{aligned}$$

holds for all $x \in A$. This implication is a special case of Lemma 2.6.

Now we must show that $\hat{g}$ extends uniquely to an endomorphism of $\mathbf{H}$. Let $\mathbf{F}$ be the free abelian group on the set Q. $\hat{g}$ extends uniquely to a homomorphism $\nu : \mathbf{F} \rightarrow \mathbf{H}$. Let $\eta : \mathbf{F} \rightarrow \mathbf{H}$ be the canonical homomorphism; i.e., η is the homomorphism that extends the identity map from Q to Q. Now, if $\hat{g} : Q \rightarrow Q$ has an extension $G : \mathbf{H} \rightarrow \mathbf{H}$, then $G \circ \eta$ is a map from $\mathbf{F}$ to $\mathbf{H}$ which agrees with ν on Q. It follows that $G \circ \eta = \nu$, since Q generates $\mathbf{F}$. In particular, $\ker \eta \subseteq \ker \nu$. But, conversely, if $\ker \eta \subseteq \ker \nu$, then the first isomorphism theorem guarantees the existence of a unique homomorphism $G : \eta(\mathbf{F}) \rightarrow \mathbf{H}$ such that $G \circ \eta = \nu$. The fact that Q generates $\mathbf{H}$ implies that $\eta(\mathbf{F}) = \mathbf{H}$. The fact that $G \circ \eta = \nu$ implies that $G|_Q = \hat{g}$. Hence G extends $\hat{g}$ in this case. This shows that $\hat{g}$ has a unique extension to $\mathbf{H}$ if and only if $\ker \eta \subseteq \ker \nu$. Let us establish this inclusion.

An element $q_1 \cdots q_n \in F$ belongs to $\ker \eta$ if and only if $q_1 \cdots q_n = 1$ in $\mathbf{H}$. The same element belongs to $\ker \nu$ if and only if $\hat{g}(q_1) \cdots \hat{g}(q_n) = 1$ in $\mathbf{H}$. Hence we need to establish the implication

$$q_1 \cdots q_n = 1 \rightarrow \hat{g}(q_1) \cdots \hat{g}(q_n) = 1.$$

For $n = 1$ the argument is as follows. $q_1 = \tilde{\rho}_a\tilde{\rho}_b^{-1}$ for some $a, b \in A$. Hence $q_1 = 1$ is equivalent to $\tilde{\rho}_a\tilde{\rho}_b^{-1} = 1$ or $\tilde{\rho}_a = \tilde{\rho}_b$. This may be written as $\rho_a \equiv \rho_b$. This is equivalent to the statement that for some m and any $v \in A$, $\rho_a\rho_v^m = \rho_b\rho_v^m$ or just $(\cdots([xa]v)\cdots)v = (\cdots([xb]v)\cdots)v$. Similarly, $\hat{g}(q_1) = 1$ is equivalent to $(\cdots([xg(a)]g(v))\cdots)g(v) = (\cdots([xg(b)]g(v))\cdots)g(v)$. Thus we must show that

$$(\cdots([xa]v)\cdots)v = (\cdots([xb]v)\cdots)v \rightarrow (\cdots([xg(a)]g(v))\cdots)g(v) = (\cdots([xg(b)]g(v))\cdots)g(v).$$

This implication follows from Lemma 2.6. The argument for $n = 2$ is identical to the argument we used to show that $\hat{g} : Q \rightarrow Q$ is well-defined. Now we turn to higher values of n.

Our goal is to prove that

$$\tilde{\rho}_{a_1}\tilde{\rho}_{b_1}^{-1}\tilde{\rho}_{a_2}\tilde{\rho}_{b_2}^{-1} \cdots \tilde{\rho}_{a_n}\tilde{\rho}_{b_n}^{-1} = 1 \rightarrow \tilde{\rho}_{g(a_1)}\tilde{\rho}_{g(b_1)}^{-1}\tilde{\rho}_{g(a_2)}\tilde{\rho}_{g(b_2)}^{-1} \cdots \tilde{\rho}_{g(a_n)}\tilde{\rho}_{g(b_n)}^{-1} = 1.$$

Our first step will be to rewrite $\tilde{\rho}_{a_1}\tilde{\rho}_{b_1}^{-1}\tilde{\rho}_{a_2}\tilde{\rho}_{b_2}^{-1} \cdots \tilde{\rho}_{a_n}\tilde{\rho}_{b_n}^{-1} = 1$ as an equality that is free of any occurrences of inversion. The idea will be to move all inverted elements from left to right using equalities of the form

$$\tilde{\rho}_v^{-1}\tilde{\rho}_u = \tilde{\rho}_{uv}\tilde{\rho}_{vv}^{-1}.$$

Once we have all inverted elements on the right, we can move them to the right side of the "=" using multiplication on the right, thereby removing all occurrences of inversion.

Assume that we have an expression which is a product of elements of $\tilde{R} = \{\tilde{\rho}_a | a \in A\}$ and inverses of elements of $\tilde{R}$. The **active region** of the expression is defined as follows. If all of the inverted elements

appear to the right of all other elements, then we shall say that the **active region is empty**. Otherwise, the active region will be the subproduct beginning at the left-most inverted element and ending at the right-most non-inverted element. For example, in the expression that concerns us now we have

$$\tilde{\rho}_{a_1} \underbrace{[\tilde{\rho}_{b_1}^{-1}\tilde{\rho}_{a_2}\tilde{\rho}_{b_2}^{-1}\cdots\tilde{\rho}_{a_n}]}_{\text{the active region}} \tilde{\rho}_{b_n}^{-1}.$$

(If $n = 1$ in this expression, then the active region is empty.) Within the active region of this particular expression we may pair up adjacent elements so that we have a sequence of $n-1$ products of the form $\tilde{\rho}_{b_i}^{-1}\tilde{\rho}_{a_{i+1}}$. We parenthesize to indicate these pairs:

$$\tilde{\rho}_{a_1}[(\tilde{\rho}_{b_1}^{-1}\tilde{\rho}_{a_2})(\tilde{\rho}_{b_2}^{-1}\tilde{\rho}_{a_3})\cdots(\tilde{\rho}_{b_{n-1}}^{-1}\tilde{\rho}_{a_n})]\tilde{\rho}_{b_n}^{-1}.$$

To each of these pairs we shall apply the rule $\tilde{\rho}_v^{-1}\tilde{\rho}_u = \tilde{\rho}_{uv}\tilde{\rho}_{vv}^{-1}$. Thinking of the inverted elements as marching from left to right and the non-inverted elements as marching from right to left, we shall view a parenthesized pair $(\tilde{\rho}_{b_i}^{-1}\tilde{\rho}_{a_{i+1}})$ as pair of elements preparing to **salute** each other as they pass on their drill. The salute will involve (i) switching positions and (ii) altering subscripts according to the rule $\tilde{\rho}_v^{-1}\tilde{\rho}_u = \tilde{\rho}_{uv}\tilde{\rho}_{vv}^{-1}$. We mandate that all parenthesized pairs salute each other simultaneously. Let us examine two rounds of the saluting process for the case $n = 4$. From

$$\tilde{\rho}_{a_1} \underbrace{[(\tilde{\rho}_{b_1}^{-1}\tilde{\rho}_{a_2})(\tilde{\rho}_{b_2}^{-1}\tilde{\rho}_{a_3})(\tilde{\rho}_{b_3}^{-1}\tilde{\rho}_{a_4})]}_{\text{the active region}} \tilde{\rho}_{b_4}^{-1}$$

we obtain

$$\tilde{\rho}_{a_1}\tilde{\rho}_{a_2b_1} \underbrace{[(\tilde{\rho}_{b_1b_1}^{-1}\tilde{\rho}_{a_3b_2})(\tilde{\rho}_{b_2b_2}^{-1}\tilde{\rho}_{a_4b_3})]}_{\text{the active region}} \tilde{\rho}_{b_3b_3}^{-1}\tilde{\rho}_{b_4}^{-1}$$

and afterwards

$$\tilde{\rho}_{a_1}\tilde{\rho}_{a_2b_1}\tilde{\rho}_{(a_3b_2)(b_1b_1)} \underbrace{[(\tilde{\rho}_{(b_1b_1)(b_1b_1)}^{-1}\tilde{\rho}_{(a_4b_3)(b_2b_2)})]}_{\text{the active region}} \tilde{\rho}_{(b_2b_2)(b_2b_2)}^{-1}\tilde{\rho}_{b_3b_3}^{-1}\tilde{\rho}_{b_4}^{-1}.$$

After each round of salutes, we are left with a new product of elements, a new active region (possibly empty) and new subscripts on some elements. Let us say a few words about the subscripts that occur. Let $Z = \{a_1, \ldots, a_n, b_1, \ldots, b_n\}$. Define a sequence of members of $\mathrm{Clo}\langle A; xy\rangle$ as follows.

(i) $s_0(x_1) = x_1$, $s_1(x_1, x_2) = x_1x_2$, $s_2(x_1, x_2, x_3, x_4) = (x_1x_2)(x_3x_4)$.

(ii) $s_{n+1}(X, Y) = (s_n(X))(s_n(Y))$, where $X = (x_1, \ldots, x_{2^n})$ and $Y = (x_{2^n+1}, \ldots, x_{2^{n+1}})$.

The following list includes the important points concerning each round of salutes.

(i) The length of the entire product remains the same.

(ii) The length of the active region decreases by 2.

(iii) If two elements $\tilde{\rho}_u^{-1}$ and $\tilde{\rho}_v$ salute in the i^{th} round, then their respective subscripts are of the form $u = s_{i-1}(\bar{p})$ and $v = s_{i-1}(\bar{q})$ for some $\bar{p}, \bar{q} \in Z^{2^{i-1}}$ before the salute and their respective subscripts are $vv = s_i(\bar{q}, \bar{q})$ and $uv = s_i(\bar{p}, \bar{q})$ afterwards.

From (iii) it is clear that the complexity of the subscript after i rounds of saluting depends only on how long the element spends in the active region. If an element spent j of those rounds in the active region, then its subscript will be $s_j(\bar{u})$ for some $\bar{u} \in Z^{2^j}$. Since any element salutes at most $n-1$ other elements we see that after $n-1$ rounds of saluting, the active region is empty. At this point, the complexity of the subscript of the element that is k^{th} from the right equals the complexity of the element k^{th} from the left. The final configuration looks like

$$\tilde{\rho}_{c_1}\tilde{\rho}_{c_2c_3}\tilde{\rho}_{(c_4c_5)(c_6c_7)}\cdots\tilde{\rho}_{s_{n-1}(C)}\tilde{\rho}_{s_{n-1}(D)}^{-1}\cdots\tilde{\rho}_{(d_7d_6)(d_5d_4)}^{-1}\tilde{\rho}_{d_3d_2}^{-1}\tilde{\rho}_{d_1}^{-1}$$

Where each $c_i, d_i \in Z$ and $C, D \in Z^{2^{n-1}}$.

Now we return to our argument that

$$\tilde{\rho}_{a_1}\tilde{\rho}_{b_1}^{-1}\tilde{\rho}_{a_2}\tilde{\rho}_{b_2}^{-1}\cdots\tilde{\rho}_{a_n}\tilde{\rho}_{b_n}^{-1} = 1 \to \tilde{\rho}_{g(a_1)}\tilde{\rho}_{g(b_1)}^{-1}\tilde{\rho}_{g(a_2)}\tilde{\rho}_{g(b_2)}^{-1}\cdots\tilde{\rho}_{g(a_n)}\tilde{\rho}_{g(b_n)}^{-1} = 1.$$

Assume that $\tilde{\rho}_{a_1}\tilde{\rho}_{b_1}^{-1}\cdots\tilde{\rho}_{a_n}\tilde{\rho}_{b_n}^{-1} = 1$. We may write this assumption as

$$\tilde{\rho}_{c_1}\tilde{\rho}_{c_2c_3}\tilde{\rho}_{(c_4c_5)(c_6c_7)}\cdots\tilde{\rho}_{s_{n-1}(C)}\tilde{\rho}_{s_{n-1}(D)}^{-1}\cdots\tilde{\rho}_{(d_7d_6)(d_5d_4)}^{-1}\tilde{\rho}_{d_3d_2}^{-1}\tilde{\rho}_{d_1}^{-1} = 1,$$

as we explained above, and by multiplying on the right to remove inversions we get

$$\tilde{\rho}_{c_1}\tilde{\rho}_{c_2c_3}\tilde{\rho}_{(c_4c_5)(c_6c_7)}\cdots\tilde{\rho}_{s_{n-1}(C)} = \tilde{\rho}_{d_1}\tilde{\rho}_{d_2d_3}\tilde{\rho}_{(d_4d_5)(d_6d_7)}\cdots\tilde{\rho}_{s_{n-1}(D)}.$$

To rewrite this as an equality between elements of $\mathbf{S}$, we use the same trick we have used earlier in this proof. We (i) replace all $\tilde{\rho}$s with ρs and equality with $\equiv$ and then (ii) multiply on the right with some ρ_v^m for large enough m so that equality is obtained. We end up with an equivalent equality

$$\rho_{c_1}\rho_{c_2c_3}\rho_{(c_4c_5)(c_6c_7)}\cdots\rho_{s_{n-1}(C)}\rho_v^m = \rho_{d_1}\rho_{d_2d_3}\rho_{(d_4d_5)(d_6d_7)}\cdots\rho_{s_{n-1}(D)}\rho_v^m.$$

This may be re-expressed as

$$\begin{aligned}(\cdots(\{\ldots[(xc_1)(c_2c_3)][(c_4c_5)(c_6c_7)]\ldots\}\{s_{n-1}(C)\}v)\cdots)\,v = \\ (\cdots(\{\ldots[(xd_1)(d_2d_3)][(d_4d_5)(d_6d_7)]\ldots\}\{s_{n-1}(D)\}v)\cdots)\,v.\end{aligned}$$

We need to prove from this assumption that

$$\tilde{\rho}_{g(a_1)}\tilde{\rho}_{g(b_1)}^{-1}\tilde{\rho}_{g(a_2)}\tilde{\rho}_{g(b_2)}^{-1}\cdots\tilde{\rho}_{g(a_n)}\tilde{\rho}_{g(b_n)}^{-1} = 1.$$

But the simplification procedure we used on the product $\tilde{\rho}_{a_1}\tilde{\rho}_{b_1}^{-1}\cdots\tilde{\rho}_{a_n}\tilde{\rho}_{b_n}^{-1}$ works exactly the same way on the product $\tilde{\rho}_{g(a_1)}\tilde{\rho}_{g(b_1)}^{-1}\cdots\tilde{\rho}_{g(a_n)}\tilde{\rho}_{g(b_n)}^{-1}$ with $g(u)$ in place of u for each $u \in Z \cup \{v\}$. Hence the implication

$$\tilde{\rho}_{a_1}\tilde{\rho}_{b_1}^{-1}\cdots\tilde{\rho}_{a_n}\tilde{\rho}_{b_n}^{-1} = 1 \to \tilde{\rho}_{g(a_1)}\tilde{\rho}_{g(b_1)}^{-1}\cdots\tilde{\rho}_{g(a_n)}\tilde{\rho}_{g(b_n)}^{-1} = 1$$

is equivalent to the implication

$$\begin{aligned}&(\cdots(\{\ldots[(xc_1)(c_2c_3)][(c_4c_5)(c_6c_7)]\ldots\}\{s_{n-1}(C)\}v)\cdots)\,v = \\ &\qquad(\cdots(\{\ldots[(xd_1)(d_2d_3)][(d_4d_5)(d_6d_7)]\ldots\}\{s_{n-1}(D)\}v)\cdots)\,v \to \\ &(\cdots(\{\ldots[(xg(c_1))(g(c_2)g(c_3))][(g(c_4)g(c_5))(g(c_6)g(c_7))]\ldots\}\{s_{n-1}(g(C))\}g(v))\cdots)\,g(v) = \\ &(\cdots(\{\ldots[(xg(d_1))(g(d_2)g(d_3))][(g(d_4)g(d_5))(g(d_6)g(d_7))]\ldots\}\{s_{n-1}(g(D))\}g(v))\cdots)\,g(v).\end{aligned}$$

(Here $g(C)$ and $g(D)$ mean the expressions obtained by applying g to each element from Z that occurs in the sequences C and D.) This latter implication holds as it is a special case of Lemma 2.6. This finishes the proof. □

Lemma 2.7 proves that $P = \{\hat{g} | g \in \mathrm{Pol}_1\mathbf{A}\}$ is a collection of endomorphisms of $\mathbf{H}$. This is the set P mentioned at the beginning of the section.

Now we describe the representing affine algebra for the case where $\mathbf{A}$ is abelian. Its universe will be H. Its fundamental operations will be

(i) The group operations of $\mathbf{H}$,

(ii) the constant operations on H and

(iii) For each $g \in \mathrm{Pol}_1\mathbf{A}$, the unary endomorphism $\hat{g}$.

We fix an element $u \in A$ and define a function $\phi : A \to H$ by the rule $a \mapsto \tilde{\rho}_a\tilde{\rho}_u^{-1}$. For each n-ary fundamental operation f of $\mathbf{A}$ we associate the following unary polynomials. For each $i = 1, \ldots, n$, let $f_i(x) = f^{\mathbf{A}}(u, u, \ldots, u, x, u, \ldots, u)$, x in the i^{th} position only. For each n-ary fundamental operation f of $\mathbf{A}$ we associate the following n-ary operation on H:

$$\hat{f}_1(x_1) + \cdots + \hat{f}_n(x_n) + h$$

where for $U = f^{\mathbf{A}}(u, u, \ldots, u)$ we define $h = \tilde{\rho}_U\tilde{\rho}_u^{-1} \in H$.

LEMMA 2.8 *For all* $\bar{a} \in A^n$

$$\phi(f^{\mathbf{A}}(a_1, \ldots, a_n)) = (\Sigma_{i=1}^n \hat{f}_i(\phi(a_i))) + h.$$

Proof: First, let us unravel the definitions involved. Each $\phi(a_i) = \tilde{\rho}_{a_i}\tilde{\rho}_u^{-1}$ and $\phi(f^{\mathbf{A}}(a_1, \ldots, a_n)) = \tilde{\rho}_{f(a_1,\ldots,a_n)}\tilde{\rho}_u^{-1}$. Further, $\hat{f}_i(\tilde{\rho}_{a_i}\tilde{\rho}_u^{-1}) = \tilde{\rho}_{f_i(a_i)}\tilde{\rho}_{f_i(u)}^{-1}$. Hence, noting that $f_i(u) = f_j(u) = U$, we must prove that

$$\tilde{\rho}_{f(a_1,\ldots,a_n)}\tilde{\rho}_u^{-1} = (\tilde{\rho}_{f_1(a_1)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_2(a_2)}\tilde{\rho}_U^{-1}) \cdots (\tilde{\rho}_{f_n(a_n)}\tilde{\rho}_U^{-1})(\tilde{\rho}_U\tilde{\rho}_u^{-1}).$$

Multiplying both sides on the right by $\tilde{\rho}_u\tilde{\rho}_U^{-1}$, this simplifies to

$$\tilde{\rho}_{f(a_1,\ldots,a_n)}\tilde{\rho}_U^{-1} = (\tilde{\rho}_{f_1(a_1)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_2(a_2)}\tilde{\rho}_U^{-1}) \cdots (\tilde{\rho}_{f_n(a_n)}\tilde{\rho}_U^{-1}).$$

We shall argue by induction in i that

$$\tilde{\rho}_{f(a_1,\ldots,a_i,u,\ldots,u)}\tilde{\rho}_U^{-1} = (\tilde{\rho}_{f_1(a_1)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_2(a_2)}\tilde{\rho}_U^{-1}) \cdots (\tilde{\rho}_{f_i(a_i)}\tilde{\rho}_U^{-1}).$$

The case $i = n$ is the statement of the lemma.

The case $i = 1$ is tautologous. Assume that the statement is true for $i = k$. That is, assume that

$$\tilde{\rho}_{f(a_1,\ldots,a_k,u,\ldots,u)}\tilde{\rho}_U^{-1} = (\tilde{\rho}_{f_1(a_1)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_2(a_2)}\tilde{\rho}_U^{-1}) \cdots (\tilde{\rho}_{f_k(a_k)}\tilde{\rho}_U^{-1})$$

holds for some $k \geq 1$. To show that the statement then holds for $i = k+1$, we must prove that the first two expressions below are equal.

$$\begin{aligned}\tilde{\rho}_{f(a_1,\ldots,a_{k+1},u,\ldots,u)}\tilde{\rho}_U^{-1} &\stackrel{?}{=} (\tilde{\rho}_{f_1(a_1)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_2(a_2)}\tilde{\rho}_U^{-1}) \cdots (\tilde{\rho}_{f_{k+1}(a_{k+1})}\tilde{\rho}_U^{-1}) \\ &= (\tilde{\rho}_{f(a_1,\ldots,a_k,u,\ldots,u)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_{k+1}(a_{k+1})}\tilde{\rho}_U^{-1}).\end{aligned}$$

The second two expressions are equal by hypothesis. It is enough for us to prove that the first expression equals the third:

$$\tilde{\rho}_{f(a_1,\ldots,a_{k+1},u,\ldots,u)}\tilde{\rho}_U^{-1} \stackrel{?}{=} (\tilde{\rho}_{f(a_1,\ldots,a_k,u,\ldots,u)}\tilde{\rho}_U^{-1})(\tilde{\rho}_{f_{k+1}(a_{k+1})}\tilde{\rho}_U^{-1}).$$

(Again we adopt the convention of using $\stackrel{?}{=}$ to indicate an equality that is yet to be established. As in the proof of Lemma 2.5, we shall modify this equality to an equivalent one by a sequence of reversible manuevers and then directly establish the final equality.) Cancelling $\tilde{\rho}_U^{-1}$ on the right of each side of the last displayed equality, we must prove

$$\tilde{\rho}_{f(a_1,\ldots,a_{k+1},u,\ldots,u)} \stackrel{?}{=} \tilde{\rho}_{f(a_1,\ldots,a_k,u,\ldots,u)}\tilde{\rho}_U^{-1}\tilde{\rho}_{f_{k+1}(a_{k+1})}.$$

Note that $\tilde{\rho}_U^{-1}\tilde{\rho}_{f_{k+1}(a_{k+1})} = \tilde{\rho}_{f_{k+1}(a_{k+1})U}\tilde{\rho}_{UU}^{-1}$, so this may be rewritten as

$$\tilde{\rho}_{f(a_1,\ldots,a_{k+1},u,\ldots,u)} \stackrel{?}{=} \tilde{\rho}_{f(a_1,\ldots,a_k,u,\ldots,u)}\tilde{\rho}_{f_{k+1}(a_{k+1})U}\tilde{\rho}_{UU}^{-1}$$

or just

$$\tilde{\rho}_{f(a_1,\ldots,a_{k+1},u,\ldots,u)}\tilde{\rho}_{UU} \stackrel{?}{=} \tilde{\rho}_{f(a_1,\ldots,a_k,u,\ldots,u)}\tilde{\rho}_{f_{k+1}(a_{k+1})U}.$$

In terms of elements of $\mathbf{S}$, this is just

$$\rho_{f(a_1,\ldots,a_{k+1},u,\ldots,u)}\rho_{f(u,\ldots,u)f(u,\ldots,u)} \stackrel{?}{=} \rho_{f(a_1,\ldots,a_k,u,\ldots,u)}\rho_{f(u,\ldots,u,a_{k+1},u,\ldots,u)f(u,\ldots,u)},$$

since $U = f(u, \ldots, u)$ and $f_{k+1}(a_{k+1}) = f(u, \ldots, u, a_{k+1}, u, \ldots, u)$ with a_{k+1} in the $(k+1)^{\text{rst}}$ position. But is even true that in $\mathbf{S}$ we have

$$\rho_{f(a_1,\ldots,a_{k+1},u,\ldots,u)}\rho_{f(u,\ldots,u)f(u,\ldots,u)} = \rho_{f(a_1,\ldots,a_k,u,\ldots,u)}\rho_{f(u,\ldots,u,a_{k+1},u,\ldots,u)f(u,\ldots,u)},$$

since this is just the statement

$$[(xf(a_1,\dots,a_{k+1},u,\dots,u))(f(u,\dots,u)f(u,\dots,u))] = \\ [(xf(a_1,\dots,a_k,u,u,\dots,u))(f(u,\dots,u,a_{k+1},u\dots,u)f(u,\dots,u))]$$

which we proved in Lemma 2.1. This finishes the argument for the inductive step and so completes the proof. □

We define $\widehat{\mathbf{H}}$ to be the algebra whose universe is H and whose fundamental operations are the collection of all $(\Sigma_{i=1}^{n}\hat{f}_i(x_i)) + h$; one for each fundamental operation f of $\mathbf{A}$. Clearly $\widehat{\mathbf{H}}$ is a reduct of the unnamed affine algebra defined in the paragraph preceding Lemma 2.8. The result of Lemma 2.8 is that $\phi : \mathbf{A} \to \widehat{\mathbf{H}}$ is a homomorphism. From this we get our main theorem.

THEOREM 1.1 *Let* $\mathbf{A}$ *be an abelian algebra with a central term* $t(x,y) = xy$. $\mathbf{A}$ *has a quasi-affine representation* Φ *such that* $\Phi(a) = \Phi(b)$ *holds if and only if for some* n *it is the case that*

$$(\cdots((aa)a)\cdots)a = (\cdots((ab)a)\cdots)a \quad \text{and} \quad a(\cdots(a(aa))\cdots) = a(\cdots(a(ba))\cdots)$$

where there are $n+1$ *occurrences of* a *on the left side of each equation and* n *occurrences of* a *on the right.*

Proof: In this section we have described a representation ϕ for any abelian algebra with a binary central term. In our representation we have that for some ℓ, any $v \in A$ and any $x \in A$ it is the case that

$$\begin{aligned} \phi(a) = \phi(b) \quad &\leftrightarrow \tilde{\rho}_a\tilde{\rho}_u^{-1} = \tilde{\rho}_b\tilde{\rho}_u^{-1} \\ &\leftrightarrow \tilde{\rho}_a = \tilde{\rho}_b \\ &\leftrightarrow \rho_a \equiv \rho_b \\ &\leftrightarrow \rho_a\rho_v^{\ell} = \rho_b\rho_v^{\ell} \\ &\leftrightarrow (\cdots((xa)v)\cdots)v = (\cdots((xb)v)\cdots)v. \end{aligned}$$

In particular, using that $\mathbf{A}$ is abelian, we set $x = v = a$ and obtain that for some ℓ

$$\phi(a) = \phi(b) \leftrightarrow (\cdots((aa)a)\cdots)a = (\cdots((ab)a)\cdots)a$$

where there are $\ell + 1$ occurrences of a on the left side of each equation and ℓ occurrences of a on the right.

Similarly, we may construct a representation ϕ' based on the dual term $t'(x,y) = t(y,x)$. We get that $\phi'(a) = \phi'(b)$ if and only if for some m it is the case that $a(\cdots(a(aa))\cdots) = a(\cdots(a(ba))\cdots)$ where there are $m+1$ occurrences of a on the left side of each equation and m occurrences of a on the right. If we choose $n = \max\{\ell, m\}$ and set $\Phi = \phi \times \phi' : \mathbf{A} \to \widehat{\mathbf{H}} \times \widehat{\mathbf{H}}'$, then Φ has the properties stated in the theorem. □

Next we show that most of the arguments we have used work equally well in some cases where $\mathbf{A}$ is non-abelian.

THEOREM 2.9 *Let* $\mathbf{A} = \langle A; xy\rangle$ *be an idempotent, entropic binar.* $\mathbf{A}$ *has a quasi-affine representation* Φ *where* $\Phi(a) = \Phi(b)$ *if and only if there exist* $c_1,\dots,c_m,d_1,\dots,d_n \in A$ *such that*

$$(\forall x)(\cdots(xa)c_1\cdots)c_m = (\cdots(xb)c_1\cdots)c_m \quad \text{and} \quad (\forall x)\ d_n(\cdots d_1(ax)\cdots) = d_n(\cdots d_1(bx)\cdots).$$

If xy *is right cancellative, then* $\Phi(a) = \Phi(b)$ *if and only if* $(\forall x)\ xa = xb$. *If* xy *is also left cancellative, then* Φ *is faithful.*

Proof: As in the proof of Theorem 1.1, let $R = \{\rho_a | a \in A\}$ and let $\mathbf{S}$ be the semigroup of self maps of $\mathbf{A}$ generated by R. The fundamental operation of $\mathbf{A}$ is central and idempotent, by assumption, so Lemmas 2.2, 2.4 and 2.5 hold. For a fixed $u \in A$ we have a function $\phi : A \to H$ defined by $a \mapsto \tilde{\rho}_a\tilde{\rho}_u^{-1}$. Our goal is to prove that ϕ extends to a representation.

Consider the automorphism

$$\alpha : \mathbf{G} \to \mathbf{G} : g \mapsto \tilde{\rho}_u^{-1} g \tilde{\rho}_u.$$

We claim that $\alpha(H) \subseteq H$. To show this, it will suffice to prove that $\alpha(Q) \subseteq Q$ where $Q = \{\tilde{\rho}_a \tilde{\rho}_b^{-1} | a, b \in A\}$ since Q generates $\mathbf{H}$. We shall use the fact that for any $a, v \in A$ we have $\rho_a \rho_{vv} = \rho_v \rho_{av}$ which we established in Lemma 2.2. Since xy is idempotent, $\rho_{vv} = \rho_v$ and so we have $\rho_a \rho_v = \rho_v \rho_{av}$. A fortiori, we have $\tilde{\rho}_a \tilde{\rho}_v = \tilde{\rho}_v \tilde{\rho}_{av}$ which we choose to write as $\tilde{\rho}_{av} = \tilde{\rho}_v^{-1} \tilde{\rho}_a \tilde{\rho}_v$. Using this equality in the case where $v = u$ we calculate that

$$\begin{aligned} \alpha(\tilde{\rho}_a \tilde{\rho}_b^{-1}) &= \tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_b^{-1} \tilde{\rho}_u \\ &= (\tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_u)(\tilde{\rho}_u^{-1} \tilde{\rho}_b^{-1} \tilde{\rho}_u) \\ &= (\tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_u)(\tilde{\rho}_u^{-1} \tilde{\rho}_b \tilde{\rho}_u)^{-1} \\ &= \tilde{\rho}_{au} \tilde{\rho}_{bu}^{-1} \in Q. \end{aligned}$$

This shows that $\alpha(Q) \subseteq Q$ and so $\alpha(H) \subseteq H$.

Claim. For all $a, b \in A$, $\phi(ab) = \alpha(\phi(a)) + (1 - \alpha)(\phi(b))$.

We begin by computing in $\mathbf{H}$, but after the first line we compute in $\mathbf{G}$ (so we change from additive to multiplicative notation).

$$\begin{aligned} \alpha(\phi(a)) + (1 - \alpha)(\phi(b)) &= \alpha(\phi(a)) - \alpha(\phi(b)) + \phi(b) \\ &= [\tilde{\rho}_u^{-1}(\tilde{\rho}_a \tilde{\rho}_u^{-1})\tilde{\rho}_u][\tilde{\rho}_u^{-1}(\tilde{\rho}_b \tilde{\rho}_u^{-1})\tilde{\rho}_u]^{-1}[\tilde{\rho}_b \tilde{\rho}_u^{-1}] \\ &= [\tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_u^{-1} \tilde{\rho}_u][\tilde{\rho}_u^{-1} \tilde{\rho}_u \tilde{\rho}_b^{-1} \tilde{\rho}_u][\tilde{\rho}_b \tilde{\rho}_u^{-1}] \\ &= [\tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_b^{-1} \tilde{\rho}_u][\tilde{\rho}_b \tilde{\rho}_u^{-1}] \end{aligned}$$

We need to interrupt the argument in order to make the following observations: If $c, d \in A$, $h \in H$, then (i) $\tilde{\rho}_c \tilde{\rho}_d^{-1} \in H$ (since $\tilde{\rho}_c \tilde{\rho}_d^{-1} \in Q$) and (ii) $\tilde{\rho}_c^{-1} h \tilde{\rho}_c = \tilde{\rho}_d^{-1} h \tilde{\rho}_d$ (since this is equivalent to saying that $\tilde{\rho}_c \tilde{\rho}_d^{-1}$ and h commute. Both elements belong to $\mathbf{H}$ and $\mathbf{H}$ is commutative, so this is clear.) From observation (i) it follows that $\tilde{\rho}_a \tilde{\rho}_b^{-1} \in H$ in our previous derivation. Applying observation (ii) to this, we get that $\tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_b^{-1} \tilde{\rho}_u = \tilde{\rho}_b^{-1} \tilde{\rho}_a \tilde{\rho}_b^{-1} \tilde{\rho}_b$. Now we resume our derivation at the last line.

$$\begin{aligned} [\tilde{\rho}_u^{-1} \tilde{\rho}_a \tilde{\rho}_b^{-1} \tilde{\rho}_u][\tilde{\rho}_b \tilde{\rho}_u^{-1}] &= [\tilde{\rho}_b^{-1} \tilde{\rho}_a \tilde{\rho}_b^{-1} \tilde{\rho}_b][\tilde{\rho}_b \tilde{\rho}_u^{-1}] \\ &= (\tilde{\rho}_b^{-1} \tilde{\rho}_a \tilde{\rho}_b) \tilde{\rho}_u^{-1} \\ &= \tilde{\rho}_{ab} \tilde{\rho}_u^{-1} \\ &= \phi(ab). \end{aligned}$$

This completes the proof of the claim

Let $\widehat{\mathbf{H}} = \langle H; \alpha(x) + (1 - \alpha)(y) \rangle$. The claim proves that $\phi : \mathbf{A} \to \widehat{\mathbf{H}}$ is a homomorphism. Clearly $\widehat{\mathbf{H}}$ is a reduct of the affine algebra obtained by adjoining the unary endomorphism α to the abelian group $\mathbf{H}$. Hence ϕ is a quasi-affine representation. For any $a, b \in A$ we have $\phi(a) = \phi(b)$ if and only if

$$\tilde{\rho}_a \tilde{\rho}_u^{-1} = \phi(a) = \phi(b) = \tilde{\rho}_b \tilde{\rho}_u^{-1}$$

which occurs exactly when $\tilde{\rho}_a = \tilde{\rho}_b$. This happens if and only if $\rho_a \equiv \rho_b$ which, by Lemma 2.4, is equivalent to $\rho_a w = \rho_b w$ for some $w \in S$. If we can express w as $\rho_{c_1} \cdots \rho_{c_m}$ for $c_1, \ldots, c_m \in A$, then $\rho_a w = \rho_b w$ says that the polynomials

$$(\cdots (xa) c_1 \cdots) c_m \quad \text{and} \quad (\cdots (xb) c_1 \cdots) c_m$$

are equal. Thus, $\phi(a) = \phi(b)$ if and only if

$$(\forall x)(\cdots (xa) c_1 \cdots) c_m = (\cdots (xb) c_1 \cdots) c_m$$

holds.

We can finish the proof of the first claim of the theorem by using the same trick we used in the proof of Theorem 1.1. We let $\phi' : \mathbf{A} \to \widehat{\mathbf{H}}$ by a representation based on yx in place of xy and set $\Phi = \phi \times \phi'$. For this Φ the first claim of the theorem holds.

For the second claim of the theorem, assume that xy is right cancellative. Then by Lemma 2.4 we have that $\equiv$ is the equality relation. Thus $\phi(a) = \phi(b)$ is equivalent to $\rho_a = \rho_b$. This is just the

statement that $(\forall x)xa = xb$. If $\phi(a) = \phi(b)$, then for any $c \in A$ we have $ca = cb$, so for all $x \in A$ we have $(ca)(cx) = (cb)(cx)$. With the entropic and idempotent laws we may write this as

$$(\forall x)c(ax) = c(bx)$$

which proves that $\phi'(a) = \phi'(b)$. It follows that $\ker \phi \subseteq \ker \phi'$ and so

$$\begin{aligned} \Phi(a) = \Phi(b) &\leftrightarrow \phi(a) = \phi(b) \ \& \ \phi'(a) = \phi'(b) \\ &\leftrightarrow \phi(a) = \phi(b) \\ &\leftrightarrow (\forall x)xa = xb. \end{aligned}$$

The third claim of the theorem easily follows from the second. □

The argument in Theorem 2.9 can be extended to the case where $\mathbf{A}$ has several fundamental operations which need not be binary. If it is the case that $\mathbf{A}$ is idempotent and entropic, then our proof can be modified to construct a representation in this more general setting. For a given fundamental operation f of $\mathbf{A}$, one defines the endomorphism which is the i^{th} coefficient of the corresponding operation of $\widehat{\mathbf{H}}$ in the way that we defined α to be the first coefficient of the operation corresponding to xy. In particular, it can be shown that if $\mathbf{A}$ is an idempotent entropic algebra with multiple cancellative binary operations, then $\mathbf{A}$ is quasi-affine.

Now we show that, in some circumstances, quasi-affine representations have strongly abelian kernels.

Definition 2.10 A congruence θ on $\mathbf{A}$ is said to be **strongly abelian** if for any $(n+1)$-ary term $s(x, \bar{y})$ and all $a \ \theta \ b$, $u_i \ \theta \ v_i \ \theta \ w_i$ it is the case that

$$s^{\mathbf{A}}(a, \bar{u}) = s^{\mathbf{A}}(b, \bar{v}) \rightarrow s^{\mathbf{A}}(a, \bar{w}) = s^{\mathbf{A}}(b, \bar{w}).$$

THEOREM 2.11 *If $\mathbf{A}$ is a right cancellative binar which satisfies*

(i) $xx = x$ *(Idempotent Law),*

(ii) $(xy)(zu) = (xz)(yu)$ *(Entropic Law),*

then $\mathbf{A}$ has a strongly abelian congruence θ such that $\mathbf{A}/\theta$ is quasi-affine and each θ-class is the universe of a left-zero subsemigroup of $\mathbf{A}$.

If $\mathbf{A}$ is an idempotent abelian binar, then $\mathbf{A}$ has a strongly abelian congruence θ such that $\mathbf{A}/\theta$ is quasi-affine.

Proof: First assume that $\mathbf{A}$ is a right cancellative, idempotent, entropic binar. Let $\theta = \ker \Phi$ where Φ is the representation defined in Theorem 2.9. As proven in Theorem 2.9, $(a, b) \in \theta$ if and only if the polynomials ρ_a and ρ_b are equal. Now pick any $(n+1)$-ary term s and $a \ \theta \ b$, $u_i \ \theta \ v_i \ \theta \ w_i$. We shall argue by induction on the complexity of s that the implication in Definition 2.10 holds. Clearly when s is projection onto a variable, the implication holds. Assume that $s^{\mathbf{A}}(a, \bar{u}) = s^{\mathbf{A}}(b, \bar{v})$. If $s(x, \bar{y}) = (s_1(x, \bar{y})s_2(x, \bar{y}))$, then

$$s_2(a, \bar{w}) \ \theta \ s_2(a, \bar{u}) \ \theta \ s_2(b, \bar{v}) \ \theta \ s_2(b, \bar{w}),$$

so

$$\rho_{s_2(a,\bar{w})} = \rho_{s_2(a,\bar{u})} = \rho_{s_2(b,\bar{v})} = \rho_{s_2(b,\bar{w})}.$$

Hence we get

$$\begin{aligned} (s_1^{\mathbf{A}}(a, \bar{u})s_2^{\mathbf{A}}(a, \bar{u})) &= s^{\mathbf{A}}(a, \bar{u}) \\ &= s^{\mathbf{A}}(b, \bar{v}) \\ &= (s_1^{\mathbf{A}}(b, \bar{v})s_2^{\mathbf{A}}(b, \bar{v})) \\ &= [s_1^{\mathbf{A}}(b, \bar{v})]\rho_{s_2(b,\bar{v})} \\ &= [s_1^{\mathbf{A}}(b, \bar{v})]\rho_{s_2(a,\bar{u})} \\ &= (s_1^{\mathbf{A}}(b, \bar{v})s_2^{\mathbf{A}}(a, \bar{u})). \end{aligned}$$

Cancelling $s_2^{\mathbf{A}}(a, \bar{u}))$ from the the right in $s_1^{\mathbf{A}}(a, \bar{u})s_2^{\mathbf{A}}(a, \bar{u}) = s_1^{\mathbf{A}}(b, \bar{v})s_2^{\mathbf{A}}(a, \bar{u})$ we get $s_1^{\mathbf{A}}(a, \bar{u}) = s_1^{\mathbf{A}}(b, \bar{v})$. Since s_1 has smaller complexity than s, we may use the implication of Definition 2.10 to deduce that

$$s_1^{\mathbf{A}}(a, \bar{w}) = s_1^{\mathbf{A}}(b, \bar{w}).$$

Together with $\rho_{s_2(a,\bar{w})} = \rho_{s_2(b,\bar{w})}$ we get

$$\begin{aligned} s^{\mathbf{A}}(a,\bar{w}) &= (s_1^{\mathbf{A}}(a,\bar{w})s_2^{\mathbf{A}}(a,\bar{w})) \\ &= s_1^{\mathbf{A}}(a,\bar{w})\rho_{s_2(a,\bar{w})} \\ &= s_1^{\mathbf{A}}(b,\bar{w})\rho_{s_2(b,\bar{w})} \\ &= (s_1^{\mathbf{A}}(b,\bar{w})s_2^{\mathbf{A}}(b,\bar{w})) \\ &= s^{\mathbf{A}}(b,\bar{w}). \end{aligned}$$

This concludes the induction. Hence θ is strongly abelian.

Let U be a θ-class. If $a, b \in U$, then the entropic law yields

$$(\underline{a}b)(\underline{aa}) = (\underline{a}a)(\underline{b}a).$$

Using the strong term condition to change all underlined parameters to a, we get $(ab)(aa) = (aa)(aa)$. We may cancel (aa) on the right to obtain $ab = aa$. Since $\mathbf{A}$ is idempotent, $ab = a$. Thus xy is a left-zero semigroup operation on U. This establishes the first claim of the theorem.

Now assume that $\mathbf{A}$ is abelian. As proved in the introduction, any idempotent abelian binar satisfies the entropic law. Hence the fundamental operation commutes with itself and with every constant operation. It follows that for any $a \in A$, the polynomials $\rho_a = xa$ and $\lambda_a = ax$ are endomorphisms of $\mathbf{A}$. By the term condition we get that

$$\underline{a}x = \underline{a}y \leftrightarrow \underline{b}x = \underline{b}y.$$

That is, $\ker \lambda_a = \ker \lambda_b$ for any $a, b \in A$. Furthermore,

$$\underline{a}(\underline{b}x) = \underline{a}(\underline{b}y) \leftrightarrow \underline{c}(\underline{d}x) = \underline{c}(\underline{d}y)$$

for any $a, b, c, d \in A$. Hence $\ker \lambda_a \circ \lambda_b = \ker \lambda_c \circ \lambda_d$ for any $a, b, c, d \in A$. In fact, any composition of n functions of the form λ_{a_i}, $a_i \in A$, has the same kernel as any other composition of n such functions and that kernel is just $\ker \lambda_a^n$ for any fixed $a \in A$. This statement is true with ρ in place of λ throughout. Let us fix one choice of $a \in A$ for the rest of the proof. Let β be the congruence $\bigcup_{n<\omega} \ker\lambda_a^n$. Let γ be the congruence $\bigcup_{n<\omega} \ker\rho_a^n$. We claim that

(*i*) $\beta \wedge \gamma = 0_{\mathbf{A}}$.

(*ii*) $\mathbf{B} = \mathbf{A}/\beta$ is left cancellative and $\mathbf{C} = \mathbf{A}/\gamma$ is right cancellative.

For (*i*), choose $(b, c) \in \beta \wedge \gamma - 0_{\mathbf{A}}$. This implies that $\lambda_a^m(b) = \lambda_a^m(c)$ for some m and $\rho_a^n(b) = \rho_a^n(c)$ for some n. We can replace (b, c) with $(b', c') = (\lambda_a^k(b), \lambda_a^k(c))$ where k is chosen maximally so that $\lambda_a^k(b) \neq \lambda_a^k(c)$. Since $\lambda_a \in \mathrm{Pol}_1\mathbf{A}$, we still have $(b', c') \in \gamma$ but now $(b', c') \in \ker\lambda_a$. With the same type of argument on the right, we can replace (b', c') with some $(b'', c'') \in \ker\lambda_a \wedge \ker\rho_a - 0_{\mathbf{A}}$. Changing notation, we assume that our original pair (b, c) was chosen from $\ker\lambda_a \wedge \ker\rho_a - 0_{\mathbf{A}}$. But now we have a problem. We have

$$b = \lambda_b(b) = \lambda_b(c) = \rho_c(b) = \rho_c(c) = c$$

since $\ker \lambda_a = \ker \lambda_b$ and $\ker \rho_a = \ker \rho_c$. This contradicts our assumption that $b \neq c$, so it must be that $\beta \wedge \gamma = 0_{\mathbf{A}}$.

To prove (*ii*) we must show, for example, that $bx\ \beta\ by \rightarrow x\ \beta\ y$. If $(bx, by) \in \beta$, then $\lambda_a^n\lambda_b(x) = \lambda_a^n\lambda_b(y)$ for some n. Hence

$$(x, y) \in \ker \lambda_a^n\lambda_b = \lambda_a^{n+1} \leq \beta.$$

The same argument with ρ_a in place of λ_a works for γ.

Now $\mathbf{A}/\gamma$ is a right cancellative, idempotent, entropic binar. By the first part of the theorem, there is a congruence $\sigma \in \mathrm{Con}\,\mathbf{A}$ such that σ/γ is a strongly abelian congruence on $\mathbf{A}/\gamma$ and $(\mathbf{A}/\gamma)/(\sigma/\gamma)$ is quasi-affine. Similarly, there is a congruence $\varsigma \in \mathrm{Con}\,\mathbf{A}$ such that ς/β is a strongly abelian congruence and $(\mathbf{A}/\beta)/(\varsigma/\beta)$ is quasi-affine. It follows that $\varsigma \times \sigma$ is a strongly abelian congruence on $(\mathbf{A}/\beta) \times (\mathbf{A}/\gamma)$ and the corresponding factor algebra is quasi-affine. If we let $\theta = \eta^{-1}(\varsigma \times \sigma)$ where $\eta : \mathbf{A} \rightarrow (\mathbf{A}/\beta) \times (\mathbf{A}/\gamma)$ is the natural homomorphism, then θ is strongly abelian (since (*a*) the strongly abelian implication is

universally quantified and (b) the fact that $\beta \wedge \gamma = 0$ implies that η is 1-1). Furthermore, the induced homomorphism

$$\bar{\eta} : \mathbf{A}/\theta \to (\mathbf{A}/\varsigma) \times (\mathbf{A}/\sigma)$$

is an embedding of $\mathbf{A}/\theta$ into a product of quasi-affine algebras. Hence $\mathbf{A}/\theta$ is quasi-affine. This proves the second statement of the theorem. □

Using tame congruence theory (see [4]) it can be proved that any finite abelian algebra has a strongly solvable congruence θ such that $\mathbf{A}/\theta$ is quasi-affine. ($\theta \in \mathrm{Con}\,\mathbf{A}$ is **strongly solvable** if there is a chain of congruences $0 = \theta_0 \leq \cdots \leq \theta_n = \theta$ such that θ_{i+1}/θ_i is strongly abelian in $\mathbf{A}/\theta_i$.) I do not know if every abelian algebra has a quasi-affine representation with a strongly *abelian* kernel, but think it is likely for finite algebras.

3 Further Remarks

In this section we discuss possible extensions of Theorem 1.1. The reason for including these remarks is that some of the ideas underlying our arguments are unlikely to be recognized after simply reading through the calculations that lead up to the proof of Theorem 1.1.

Assume that $\phi : \mathbf{A} \to \mathbf{H}$ is a quasi-affine representation where $\mathbf{H}$ is a reduct of some affine algebra, $\mathbf{M}$. By replacing $\mathbf{H}$ and $\mathbf{M}$ if necessary, we may assume that $\mathbf{A}$, $\mathbf{H}$ and $\mathbf{M}$ have a close relationship, to be described shortly. We shall assume that this relationship holds throughout this paragraph and the next. First, we may assume that the clone of $\mathbf{M}$ is generated by $x - y + z$ and the operations of $\mathbf{H}$. This is possible since the clone so generated will be that of an affine algebra, $\mathbf{M}'$, and $\mathbf{H}$ is also a reduct of $\mathbf{M}'$. We may further assume that $\mathbf{M}$ is generated by the set $\phi(A)$. Finally, replacing $\mathbf{M}$ with a homomorphic image if necessary, we may assume that every non-zero congruence on $\mathbf{M}$ has a non-zero restriction to $\mathbf{A}$. With these assumptions in place we say that $\mathbf{A}$ is **minimally represented on** $\mathbf{M}$. When $\mathbf{A}$ is minimally represented on $\mathbf{M}$, it is not hard to see that any two term or polynomial operations of $\mathbf{H}$ which agree on $\phi(A)$ also agree on $H = M$. Such a representation of $\mathbf{A}$ on $\mathbf{H}$ where $\mathbf{H}$ is a reduct of $\mathbf{M}$ induces well-defined clone homomorphisms

$$\epsilon : \mathrm{Clo}\,\mathbf{A} \to \mathrm{Clo}\,\mathbf{H}, \qquad \mu : \mathrm{Clo}\,\mathbf{H} \to \mathrm{Clo}\,\mathbf{M}$$

(and similar homomorphisms between polynomial clones). Here ϵ is an epimorphism (in fact a surjection) and μ is a monomorphism. Clo $\mathbf{M}$ is generated by $\mu(\mathrm{Clo}\,\mathbf{H})$ together with $x-y+z$, so $\mu(\mathrm{Clo}\,\mathbf{H})$ might be considered to be "large" in Clo $\mathbf{M}$, but μ does not have to be an epimorphism of clones. (To eliminate $\mathbf{H}$ from our discussion, note that μ is an epimorphism if and only if $\mu\epsilon$ is an epimorphism.) For an example where μ is not an epimorphism, take $\mathbf{A}$ to be a 10-element set with no operations. Then Clo $\mathbf{M}$ is generated by a ternary abelian group operation $x - y + z$, $\mathbf{M}$ has at least 10 distinct elements and $\mathbf{M}$ is generated by those elements. Let σ be any permutation of M which does not commute with $x - y + z$ and let C be the clone on M generated by the operations $x - y + z$ and $\sigma^{-1}(\sigma(x) - \sigma(y) + \sigma(z))$. There are two obvious maps from Clo $\mathbf{M}$ to C: f which sends $x - y + z$ to $x - y + z$ and g which sends $x - y + z$ to $\sigma^{-1}(\sigma(x) - \sigma(y) + \sigma(z))$. One has that $f \neq g$, but $f\mu\epsilon = g\mu\epsilon =$ the unique clone homomorphism from Clo $\mathbf{A}$ to C. This shows that, no matter how $\mathbf{M}$ is chosen, $\mu\epsilon$ will not be an epimorphism when $\mathbf{A}$ is a 10-element set with no operations.

However, for some quasi-affine representations it *may* occur that $\mu\epsilon$ is an epimorphism. The best known example where this occurs is where $\mathbf{A} = \mathbf{N}$ is the semigroup of natural numbers and $\mathbf{M} = \mathbf{Z}$ is the group of integers. The classical construction of $\mathbf{Z}$ from $\mathbf{N}$ is perhaps the first "affinization" argument. The construction of $\mathbf{Z}$ from $\mathbf{N}$ might have been the blueprint for some later affinization arguments, e.g., those found in [3], [5] and [6]. In each of these papers where a quasi-affine representation is obtained for an algebra $\mathbf{A}$ it is possible to show that the clone homomorphism from Clo $\mathbf{A}$ to the clone of the representing algebra $\mathbf{M}$ is an epimorphic embedding of clones. The main difference one finds in the present paper is that we construct quasi-affine representations where the induced clone homomorphism does not need to be epimorphic, but as a consequence the representations are not always faithful.

Since every affine algebra is a reduct of a module expanded by constants, the difficulty in constructing a representation of $\mathbf{A}$ is the problem of associating a module to $\mathbf{A}$. Let us look at a classical method for constructing a module, and then discuss a new view of essentially the same construction.

We begin with the following data: a pair $\langle \mathbf{G}; \mathbf{H} \rangle$ where $\mathbf{G}$ is a group and $\mathbf{H}$ is an abelian subgroup of $\mathbf{G}$. The group $\mathbf{H}$ will be the underlying abelian group of our module. For each element $g \in G$ we have a corresponding inner automorphism which, restricted to $\mathbf{H}$, gives an endomorphism:

$$\alpha_g : \mathbf{H} \to \mathbf{H} : h \mapsto ghg^{-1}.$$

In this way, $\mathbf{H}$ may be viewed as an $\mathbf{R}$-module where $\mathbf{R}$ may be taken to be the group ring of $\mathbf{G}$ with integer coefficients (or the subring of End($\mathbf{H}$) generated by the endomorphisms α_g, $g \in G$). So, as a general philosophy, to come up with a module which we can associate to an abelian algebra $\mathbf{A}$, we might begin by looking for some pair $\langle \mathbf{G}; \mathbf{H} \rangle$ associated to $\mathbf{A}$ where $\mathbf{G}$ is a group and $\mathbf{H}$ is an abelian subgroup of $\mathbf{G}$.

Our other method for constructing a module begins with the following data: $\langle \mathbf{G}; \tau \rangle$ where $\mathbf{G}$ is a group and τ is an abelian tolerance on $\mathbf{G}$. Group tolerances are identical with group congruences and each congruence corresponds uniquely to a normal subgroup; abelian tolerances/congruences correspond to abelian normal subgroups. Hence, the data required for this construction is the same as the data for the previous construction. The tolerance τ may be viewed as a subgroup of $\mathbf{G} \times \mathbf{G}$ which contains the diagonal. If $\mathbf{D}$ is the diagonal subgroup, then we will construct our module on $\tau/\mathbf{D}$: the set of left cosets of $\mathbf{D}$ in the group τ.

Let $\mathbf{H}$ be the abelian normal subgroup of $\mathbf{G}$ that corresponds to τ. Each left coset of $\mathbf{D}$ contains a unique element of the form $\langle h, 1 \rangle$ where $h \in H$. E.g., the unique element in $\langle a, b \rangle D$ of this form is $\langle ab^{-1}, 1 \rangle$. Hence we may put abelian group operations on $\tau/\mathbf{D}$ by mimicking the operations of $\mathbf{H}$: to add $\langle a, b \rangle D$ to $\langle c, d \rangle D$ we first rewrite these as $\langle h, 1 \rangle D$ and $\langle k, 1 \rangle D$ where $h = ab^{-1}$ and $k = cd^{-1}$. Then we define

$$\langle a, b \rangle D + \langle c, d \rangle D = \langle h, 1 \rangle D + \langle k, 1 \rangle D = \langle hk, 1 \rangle D = \langle ab^{-1}cd^{-1}, 1 \rangle D.$$

Negation is defined by

$$-\langle a, b \rangle D = -\langle h, 1 \rangle D = \langle h^{-1}, 1 \rangle D = \langle ba^{-1}, 1 \rangle D = \langle b, a \rangle D$$

and 0 is defined to be $\langle 1, 1 \rangle D = D$. The abelian group laws for $\tau/\mathbf{D}$ follow from those of $\mathbf{H}$.

Clearly, τ acts by multiplication (on the left) on the set of left cosets of $\mathbf{D}$. This action is affine with respect to the described abelian group operations, so if we restrict the action to the subgroup $\mathbf{D} < \tau$, then we get that $\mathbf{D}$ acts on these cosets by left multiplication in a way that preserves 0. Hence, $\mathbf{D}$ acts as a group of endomorphisms of the abelian group structure on $\tau/\mathbf{D}$. This gives us a module structure on $\tau/\mathbf{D}$. This module is essentially the same as the one defined earlier; since the diagonal subgroup $\mathbf{D} < \tau < \mathbf{G}^2$ is isomorphic to $\mathbf{G}$, the abelian group structure on $\tau/\mathbf{D}$ is isomorphic to $\mathbf{H}$, and $\mathbf{D}$ acts on $\tau/\mathbf{D}$ "essentially" by conjugation. By this, we mean that if $\langle g, g \rangle \in D$ acts on the coset $\langle h, 1 \rangle D$ by left multiplication, we get

$$\langle g, g \rangle (\langle h, 1 \rangle D) = \langle gh, g \rangle D = \langle ghg^{-1}, 1 \rangle D.$$

So in fact, the module we have described is essentially the same as the one described previously.

We introduced the second construction for the following reason. It is nearly true that to each abelian algebra we can associate a pair $\langle \mathbf{G}; \tau \rangle$ where $\mathbf{G}$ is a group and τ is an abelian tolerance on $\mathbf{G}$. What is actually true is that it is always possible to associate a pair $\langle \mathbf{S}; \tau \rangle$ where $\mathbf{S}$ is a monoid and τ is an abelian tolerance on $\mathbf{S}$. (Here an **abelian tolerance** is just what it sounds like: τ is an abelian tolerance on the monoid $\mathbf{S}$ if $s(x, \bar{y})$ is an $(n+1)$-ary monoid term, $\langle f, g \rangle \in \tau$ and $\langle u_i, v_i \rangle \in \tau$, then the implication

$$s(f, \bar{u}) = s(f, \bar{v}) \to s(g, \bar{u}) = s(g, \bar{v})$$

holds.) The monoid we associate to the abelian algebra $\mathbf{A}$ is just $\mathbf{S} = \mathrm{Pol}_1 \mathbf{A}$. The tolerance on this monoid is just the **twin relation**: $\langle f(x), g(x) \rangle \in \tau$ if there is a term $t(x, \bar{y})$ such that $f(x) = t^{\mathbf{A}}(x, \bar{a})$ and $g(x) = t^{\mathbf{A}}(x, \bar{b})$ for some $\bar{a}, \bar{b} \in A^n$. The fact that τ is an abelian tolerance follows easily from the fact that $\mathbf{A}$ is abelian.

Now, we have pointed out that to any abelian algebra one may associate a pair $\langle \mathbf{S}; \tau \rangle$ where $\mathbf{S}$ is a monoid and τ is an abelian tolerance on $\mathbf{S}$. If $\mathbf{S}$ happened to be a group, then we already know how to construct a module from this data; but, of course, $\mathbf{S}$ is a group only when $\mathbf{A}$ is trivial. Nevertheless, this often doesn't matter. To see this, let's examine what happens when $\mathbf{A}$ is an affine algebra. For the rest of this paragraph, $\mathbf{A}$ is affine, $\mathbf{S}$ equals $\mathrm{Pol}_1\mathbf{A}$, and τ is the twin relation. Let T equal the set of unary polynomials of $\mathbf{A}$ of the form $x + a$, $a \in A$. I.e., T is the set of additive translations. Composition of translations is an abelian group operation on T, and under composition T is isomorphic to the underlying additive semigroup of $\mathbf{A}$. It turns out that τ is the submonoid of $\mathbf{S}^2 = (\mathrm{Pol}_1\mathbf{A})^2$ consisting of pairs of unary polynomials of the form $\langle \alpha(x) + m, \alpha(x) + n \rangle$, $m, n \in A$. Each pair in τ may be thought of as representing a translation, since there is exactly one $t \in T$ such that $t(\alpha(x) + m) = \alpha(x) + n$: it is $t(x) = x + (n - m)$. However, different pairs in τ may represent the same member of T. For example, $\langle \alpha(x) + m, \alpha(x) + n \rangle$ and $\langle \beta(x) + r, \beta(x) + s \rangle$ represent the same element of T if $m - n = r - s$. We factor the monoid τ by the relation $\sim$ that identifies two pairs if they represent the same translation. If D is the diagonal of $(\mathrm{Pol}_1\mathbf{A})^2$, then it turns out that $\sim$ is precisely the left congruence on the monoid τ which is generated by $D \times D$. Furthermore, D is a $\sim$-class, since a pair $\langle f, g \rangle$ is in D iff $f = g$ iff $\langle f, g \rangle \in \tau$ represents the identity translation. We can define abelian group operations on $\tau/\sim$ in the expected way. Let $0 = D/\sim$. Let $-\langle f, g \rangle/\sim = \langle g, f \rangle/\sim$. If f or g is the identity function and h or k is the identity function, then define

$$(\langle f, g \rangle/\sim) + (\langle h, k \rangle/\sim) := \langle fh, gk \rangle/\sim .$$

It turns out that $+$, $-$ and 0 are well-defined total abelian group operations on $\tau/\sim$. Since $\sim$ is a left congruence, the monoid τ acts by left multiplication on $\tau/\sim$. The diagonal submonoid $\mathbf{D} < \tau$ acts as a monoid of abelian group endomorphisms. We get that

$$\langle \tau/\sim; +, -, 0, d(d \in D) \rangle$$

is a $\mathbf{D}$-module. This module is polynomially equivalent to the underlying module of $\mathbf{A}$. The first main point to observe is that *everything works just as before even though* $\mathbf{S}$ *is not a group.* The second main point is that *this procedure always produces the correct representing module if* $\mathbf{A}$ *is affine.* This suggests how we might proceed when $\mathbf{A}$ is not (known to be) affine. The problem is to decide how to define τ, $\sim$, $+$, -, 0 and $\mathbf{D}$ in this generality.

We now relax our assumption that $\mathbf{A}$ is affine to the assumption that $\mathbf{A}$ is abelian. We still define $\mathbf{S} = \mathrm{Pol}_1\mathbf{A}$ and $\tau =$ the twin relation. Next, it seems like a good idea to choose $\sim$ to be the left congruence on τ generated by $D \times D$. In the affine case, $\sim$ is the unique left congruence on τ which has D as a congruence class, but there may be many such left congruences when $\mathbf{A}$ is merely abelian. If we choose $\sim$ to be any left congruence other than the one generated by $D \times D$, our choice for $\sim$ might be incompatible with some or all of the possible representations of $\mathbf{A}$. By choosing the least left congruence on τ which has D as a class, we do not exclude any representations. This choice gives us a set $\tau/\sim$ on which to build a module structure. There is no difficulty in defining 0, negation or a left $\mathbf{D}$-action on this set just as we did in the last paragraph. But our definition for addition does not always give a well-defined total operation. (In some cases it will. For example, if $\mathbf{A}$ is a commutative cancellative monoid, then $\langle \tau/\sim; +, -, 0, d(d \in D) \rangle$ is a total algebra equal to the universal group of $\mathbf{A}$.) At best, we can only construct what might be called a "partial module."

We have just described a failed attempt to construct a representing module for an arbitrary abelian algebra. However, it indicates a good outline for a successful representation if one adopts certain additional hypotheses chosen so that the partial module obtained is total. Rather than work with all of $\mathrm{Pol}_1\mathbf{A}$, one might think to restrict the above discussion to a sub*group* $\mathbf{G}$ of $\mathrm{Pol}_1\mathbf{A}$. If one restricts the twin relation τ to $\mathbf{G}$, one obtains an abelian tolerance on $\mathbf{G}$. As remarked earlier, group tolerances are congruences and abelian tolerances correspond to abelian normal subgroups. Hence, if one restricts to a subgroup $\mathbf{G}$ of $\mathrm{Pol}_1\mathbf{A}$, then there *will be* an abelian normal subgroup $\mathbf{H} \lhd \mathbf{G}$ such that any family of mutually twin polynomials in $\mathbf{G}$ lie in the same coset of $\mathbf{H}$. This leads to an expansion of $\mathbf{H}$ to a module and it suggests a mapping of $\mathbf{A}$ into a coset of $\mathbf{H}$: take a binary term xy and map $a \mapsto \rho_a = xa$ (provided that each $\rho_a \in G$). This is a function from A into G whose image lies in a single coset of $\mathbf{H}$, since any pair $\langle \rho_a, \rho_b \rangle$ is a pair of twins. The abelian group operations on $\mathbf{H}$ may be translated to the coset of $\mathbf{H}$

containing all $\{\rho_a \mid a \in A\}$ or the image of φ may be translated to the coset H where we may directly use the abelian group operations of $\mathbf{H}$. We have the beginnings of a representation. Hence, we are led to consider the case of an abelian algebra which has a binary term xy where the right translations ρ_a generate a subgroup of $\mathrm{Pol}_1\mathbf{A}$. We may take $\mathbf{G}$ to be the group so generated and the twin relation, τ, restricted to $\mathbf{G}$ will be an abelian congruence. The function $\varphi(a) = \rho_a$ is a mapping of A into a single congruence class in $\mathbf{G}$. We can translate the image of φ into H by fixing $u \in A$ and defining $\phi(x) = \varphi(x)\varphi(u)^{-1}$. That is, $\phi : A \to H : a \mapsto \rho_a\rho_u^{-1}$ for a fixed $u \in A$. The reader will recognize that this paragraph describes the essential features of first half of the argument for Theorem 1.1.

But, assuming that the right translations $\{\rho_a \mid a \in A\}$ generate a subgroup of $\mathrm{Pol}_1\mathbf{A}$ is too restrictive. Here is why we say this. If $\mathrm{Pol}_1\mathbf{A}$ has a subgroup $\mathbf{G}$ which contains all $\rho_a = xa \in \mathrm{Pol}_1\mathbf{A}$, then for each $a \in A$ the function $x \mapsto xa$ has a polynomial inverse. It is hard to imagine that this would ever happen naturally except when the term xy is invertible in its first variable. Here we say that a polynomial $f(x, \bar{y})$ is **invertible in its first variable** if there is a polynomial $g(x, \bar{y})$ such that

$$\mathbf{A} \models f(g(x, \bar{y}), \bar{y}) = x = g(f(x, \bar{y}), \bar{y}).$$

Furthermore, if we hope that our representation is faithful, then we will want distinct $a, b \in A$ to yield distinct operations $\rho_a = xa$ and $\rho_b = xb$. Using the term condition, it is easy to see that this is equivalent to cancellativity of xy in its second variable. But now, with xy invertible in its first variable and cancellative in its second variable, we can only be talking about an algebra which looks dangerously similar to a division quasigroup. In particular, any algebra which has a polynomial which is invertible in two distinct variables has division quasigroup polynomials, as is easy to see. What makes this situation "dangerous" is that any abelian algebra with division quasigroup polynomials is affine; hence, constructing a quasi-affine representation for such an algebra is certainly a step backwards. Thus, it is our desire to avoid trivial cases that makes us reluctant to restrict our attention to subgoups of $\mathrm{Pol}_1\mathbf{A}$. So, instead of dealing with a subgroup of $\mathrm{Pol}_1\mathbf{A}$, one might choose to deal with a submonoid embeddable in a group or, as we have done in this paper, with a subsemigroup of $\mathrm{Pol}_1\mathbf{A}$ that has a "good" homomorphism into a group. These generalizations introduce a new interesting problem. When one begins with a subgroup $\mathbf{G} \subseteq \mathrm{Pol}_1\mathbf{A}$, then the restriction of τ to $\mathbf{G}$ is a congruence which has a congruence class equal to an abelian subgroup $\mathbf{H} \triangleleft \mathbf{G}$. One may proceed to define a representation on an expansion of $\mathbf{H}$. But if one begins only with a subsemigroup $\mathbf{S} \subseteq \mathrm{Pol}_1\mathbf{A}$ which has an abelian tolerance τ and one embeds $\mathbf{S}$ into a group $\mathbf{G}$, will τ extend to an abelian congruence on $\mathbf{G}$? The answer is 'not necessarily.' It is this difficulty which motivated our decision to let the binary term $t(x, y) = xy$ be a *central* binary term in Theorem 1.1. The assumption that xy is a central term came in quite handy in the proofs of every lemma that led up to the proof of Theorem 1.1, but primarily it was chosen make Lemma 2.5 work. If one begins with an abelian algebra $\mathbf{A}$ and a binary term xy, one can construct the semigroup $\mathbf{S}$ generated by the right translations. The twin relation τ on $\mathbf{S}$ is an abelian tolerance. If one wants to proceed along the lines we have described *without* the assumption that xy is central, then we need a homomorphism of $\mathbf{S}$ into a group which has an abelian congruence whose restriction to $\mathbf{S}$ contains τ. This leads us to the following question.

Question. If $\mathbf{S}$ is a semigroup which is embeddable into a group $\mathbf{G}$ and τ is an abelian tolerance on $\mathbf{S}$, then is it possible to embed $\mathbf{S}$ into a group $\mathbf{G}'$ which has an abelian congruence whose restriction to $\mathbf{S}$ contains τ?

In general, one cannot always take $\mathbf{G} = \mathbf{G}'$. As an example, let $\mathbf{S}$ be the free semigroup on the set X, $|X| > 1$. The relation τ which relates any two words of $\mathbf{S}$ which have the same length is an abelian congruence on $\mathbf{S}$. Now $\mathbf{S}$ is embeddable into the free group $\mathbf{G}$ generated by the set X by a semigroup homomorphism which is the identity on X. But $\mathbf{G}$ has no non-trivial abelian congruence, hence $\mathbf{G}$ has no abelian congruence whose restriction to $\mathbf{S}$ contains τ. If $\mathbf{S}$ is embeddable into a group $\mathbf{G}'$ which has an abelian congruence α whose restriction to $\mathbf{S}$ contains τ, then $\mathbf{G}'$ may be chosen to be a group of fractions for $\mathbf{S}$. Now, since $\tau \subseteq \alpha|_{\mathbf{S}}$, there is an induced homomorphism $h : \mathbf{S}/\tau \to \mathbf{G}'/\alpha$ where $h(S/\tau)$ generates $\mathbf{G}'/\alpha$. But $\mathbf{S}/\tau$ is 1-generated, so $\mathbf{G}'/\alpha$ is cyclic. It follows that $\mathbf{G}'$ must have an abelian congruence α such that $\mathbf{G}'/\alpha$ is cyclic. A positive answer to the question we listed would imply, therefore, that any free semigroup is embeddable into a group $\mathbf{G}'$ which has an abelian normal subgroup

H such that $\mathbf{G}'/\mathbf{H}$ is cyclic. This statement, which turns out to be true, is not at all obvious. It hints that a positive answer to our question will not be easily proven.

References

[1] P. Dubreil, *Sur les problèmes d'immersion et la théorie des modules*, C. R. Acad. Sci. (Paris) **216** (1943) 625–627.

[2] R. Freese and R. McKenzie, *Commutator Theory for Congruence Modular Varieties*, LMS Lecture Notes, No. 125, 1987.

[3] C. Herrmann, *Affine algebras in congruence modular varieties*, Acta Sci. Math. **41** (1979), 119–125.

[4] D. Hobby and R. McKenzie, *The Structure of Finite Algebras*, Contemporary Mathematics, American Mathematics Society, Providence, Rhode Island, 1988.

[5] K. A. Kearnes, *Idempotent simple algebras*, manuscript.

[6] K. A. Kearnes, *The structure theorem for abelian algebras*, in preparation.

[7] R. McKenzie, G. McNulty and W. Taylor, *Algebras, Lattices and Varieties, vol. 1*, Wadsworth & Brooks/Cole, 1987.

[8] R. Quackenbush, *Quasi-affine algebras*, Algebra Universalis **20** (1985), 318–327.

1980 *Mathematics subject classification* (1985): Primary 08A05, Secondary 08A40.