A correspondence between balanced varieties and inverse monoids

Mark V. Lawson
Department of Mathematics
Heriot-Watt University
Riccarton
Edinburgh EH14 4AS
Scotland
e-mail: M.V.Lawson@ma.hw.ac.uk

July 19, 2005

Abstract

There is a well-known correspondence between varieties of algebras and fully invariant congruences on the appropriate term algebra. A special class of varieties are those which are balanced, meaning they can be described by equations in which the same variables appear on each side. In this paper, we prove that the above correspondence, restricted to balanced varieties, leads to a correspondence between balanced varieties and inverse monoids. In the case of unary algebras, we recover the theorem of Meakin and Sapir that establishes a bijection between congruences on the free monoid with n generators and wide, positively self-conjugate inverse submonoids of the polycyclic monoid on n generators.

In the case of varieties generated by linear equations, meaning those equations where each variable occurs exactly once on each side, we can replace the clause monoid above by the linear clause monoid. In the case of algebras with a single operation of arity n, we prove that the linear clause monoid is isomorphic to the inverse monoid of right ideal isomorphisms between the finitely generated essential right ideals of the free monoid on n letters, a monoid previously studied by Birget in the course of work on the Thompson group V and its analogues.

We show that Dehornoy's geometry monoid of a balanced variety is a special kind of inverse submonoid of ours.

Finally, we construct groups from the inverse monoids associated with a balanced variety and examine some conditions under which they still reflect the structure of the underlying variety. Both free groups and Thompson's groups $V_{n,1}$ arise in this way.

2000 AMS Subject Classification: 20M18, 20L05, 08B99.

1 Introduction

In [7], Patrick Dehornoy showed that an inverse monoid could be associated with each variety that can be described by means of a set of of balanced equations, where a balanced equation is one in which the variables occurring on each side of the equation are the same. This monoid, which Dehornoy termed the 'structural monoid', and later the 'geometry monoid', of the variety, acts on the set of terms. The equivalence relation it determines is the fully invariant congruence associated with the original variety. The geometry monoid is therefore intimately connected with the structure of the variety. In addition, the universal groups of the geometry monoids are interesting. For example, the universal group of the geometry monoid associated with the variety of semigroups is none other than Thompson's group F [8, 9, 10].

Dehornoy's idea [7] relies on the notion of a 'minimal presentation' of a variety. He associates an inverse monoid with each minimal presentation, and then proves that the inverse monoids associated with different minimal presentations are isomorphic. It is this common inverse monoid that is defined to be the geometry monoid of the variety. Dehornoy states that it can be regarded as being 'essentially' the family of identities that hold in the variety.

The main goal of this paper is to show that the set of *all* identities that hold in a balanced variety V can be used to construct an inverse monoid, which I shall denote by $\mathcal{I}(\mathsf{V})$, and call the 'inverse monoid of the variety'. Dehornoy's geometry monoid will turn out to be a special inverse submonoid of $\mathcal{I}(\mathsf{V})$.

2 Statement of the main theorem

We begin with some definitions from universal algebra; for standard results, I refer the reader to [2], [6] or [13].

Let $X = \{x_i \colon i \in \mathbb{N}\}$, our set of *variables*. If $A \subseteq X$ then 1_A denotes the identity function defined on A. Let Ω be the *operator domain*: that is, a list of function symbols and their arities. We shall refer to Ω-*algebras* or *algebras of type* Ω. We denote by $T(X) = T_\Omega(X)$ the *term algebra* of type Ω over X. Operations of arity zero are called *constants*. Terms having no variables are called *ground terms*. If $A \subseteq X$ then $T(A) = T_\Omega(A)$. The term algebra $T(A)$ is itself an algebra of type Ω. A *variety of* Ω-*algebras* is a collection of Ω-algebras closed under arbitrary direct products, homomorphic images, and subalgebras. If s is a term, then $\mathbf{v}(s)$ denotes the set of variables occurring in s. If $s, t \in T(X)$, then I shall write $s \approx t$ to be the *equation* they determine. I shall also feel free to regard an equation $s \approx t$ as an ordered pair (s, t). If I write $s = t$ then I mean the terms s and t are identically equal. A *fully invariant congruence* $\mathcal{G}$ on $T_\Omega(X)$ is a relation $\approx$ satisfying the following five conditions:

(C1) $s \approx s \in \mathcal{G}$ for each term $s \in T_\Omega(X)$.

(C2) If $s \approx t \in \mathcal{G}$ then $t \approx s \in \mathcal{G}$.

(C3) If $s \approx t, t \approx u \in \mathcal{G}$ then $s \approx u \in \mathcal{G}$.

(C4) If $s(x_1, \ldots, x_n) \approx t(x_1, \ldots, x_n)$ and $u_1, \ldots, u_n \in T_\Omega(X)$ then

$$s(u_1, \ldots, u_n) \approx t(u_1, \ldots, u_n).$$

(C5) If $s_i \approx t_i \in \mathcal{G}$ for $1 \leq i \leq n$ and ρ is an n-ary operation symbol in Ω then $\rho(s_1, \ldots, s_n) \approx \rho(t_1, \ldots, t_n)$.

The set of all equations that hold in every algebra of a variety forms a fully invariant congruence, and the set of all algebras that satisfy all the equations in a fully invariant congruence is a variety. In fact, the following result is true. See Theorems IV.1.2 and IV.3.1 of [6] for a proof.

Result 2.1 (Birkhoff's variety theorem) *There is a bijection between varieties of Ω-algebras and fully invariant congruences on $T_\Omega(X)$.* ■

Let V be a variety and $\mathcal{G}$ its corresponding fully invariant congruence. A *presentation* of V is a set of equations $\mathcal{E}$ such that the smallest fully invariant congruence on $T(X)$ containing $\mathcal{E}$ is $\mathcal{G}$. The proof of the following may be found in [2].

Result 2.2 (Birkhoff's completion theorem for equational logic) *Let $\mathcal{E}$ be a set of equations, and let $\mathcal{G}$ be the fully invariant congruence they generate. Then the equation $s \approx t$ belongs to $\mathcal{G}$ iff $s \approx t$ can be derived from $\mathcal{E}$ using the following operations:*

(EL1) *$\mathcal{E} \vdash s \approx t$ for each $s \approx t \in \mathcal{E}$.*

(EL2) *$\mathcal{E} \vdash s \approx s$ for each term s.*

(EL3) *If $\mathcal{E} \vdash s \approx t$ then $\mathcal{E} \vdash t \approx s$.*

(EL4) *If $\mathcal{E} \vdash s \approx t$ and $\mathcal{E} \vdash t \approx u$ then $\mathcal{E} \vdash s \approx u$.*

(EL5) *If $\mathcal{E} \vdash s(x_1, \ldots, x_n) \approx t(x_1, \ldots, x_n)$ and $u_1, \ldots, u_n \in T_\Omega(X)$ then*

$$\mathcal{E} \vdash s(u_1, \ldots, u_n) \approx t(u_1, \ldots, u_n).$$

(EL6) *If $\mathcal{E} \vdash s_i \approx t_i$ for $1 \leq i \leq n$ and ρ is an n-ary operation then*

$$\mathcal{E} \vdash \rho(s_1, \ldots, s_n) \approx \rho(t_1, \ldots, t_n).$$

■

An equation $s \approx t$ is said to be *balanced* if the set of variables occurring in s is the same as the set of variables occurring in t. A presentation is *balanced* if every equation in it is balanced. Using Result 2.2, it is easy to show that if the fully invariant congruence $\mathcal{G}$ is generated by a balanced presentation then every equation occurring in $\mathcal{G}$ is balanced; in this case, I shall say that the

fully invariant congruence $\mathcal{G}$ is *balanced*. Thus balanced varieties correspond to balanced fully invariant congruences.

In this paper, an inverse monoid will be constructed from each balanced variety by refining Result 2.1. We now review the basic definitions of inverse semigroup theory, and describe the extra structure that the inverse monoids associated with a balanced variety will have.

A *semigroup* is a set equipped with an associate binary operation, and a *monoid* is a semigroup with an identity. A *zero* 0 in a semigroup S is an element such that $0s = 0 = s0$ for all elements s. *Monoid homomorphisms* preserve identities, and *homomorphisms of semigroups with zero* preserve zeros. A semigroup S is said to be *inverse* if for each $s \in S$ there exists a unique element, denoted s^{-1} and called the *inverse* of s, such that $s = ss^{-1}s$ and $s^{-1} = s^{-1}ss^{-1}$. An element e of S is called an *idempotent* if $e^2 = e$. The set of all idempotents of S, denoted $E(S)$, forms a commutative subsemigroup of S. If $s, t \in S$ we define the relation $s \leq t$ if $s = te$ for some idempotent e. It can be proved that $\leq$ is a partial order, called the *narural partial order*. With respect to this order the set of idempotents becomes a meet semilattice.

If S is an inverse semigroup, then we can define a groupoid, in the sense of category theory, denoted $(S, \cdot)$, where $s \cdot t$ is defined and equals st iff $s^{-1}s = tt^{-1}$. The groupoid of S and its natural partial order together determine the semigroup multiplication, since

$$st = (se) \cdot (et)$$

where $e = s^{-1}stt^{-1}$ and $se \leq s$ and $et \leq t$. If $\theta\colon S \to T$ is a function between two inverse semigroups, then it is a homomorphism iff it is a functor for the groupoid structure, order preserving, and is a semilattice map between the respective semilattices of idempotents.

The *symmetric inverse monoid* $I(X)$ is the set of all partial bijections of the set X under composition of partial functions. It is an inverse monoid. An *action* of an inverse monoid S on a set X is a monoid homomorphism from S to $I(X)$. Such an action induces an equivalence relation on the set X.

A subset A of an inverse semigroup S is said to be an *order ideal* if $s \leq a \in A$ implies $a \in A$. If S is an inverse monoid and T an inverse submonoid of S, we say that T is a *wide* inverse submonoid if the idempotents in S and T are the same. If T is a wide inverse submonoid of S then T is also an order ideal. For proofs of all the above assertions see [15].

An *inverse Ω-algebra* is an inverse monoid S which is also an Ω-algebra such that the Ω-operations are semigroup homomorphisms. For example, suppose that Ω consists of one binary operation symbol $\times$. Then the inverse monoid S is an inverse Ω-algebra if in addition to its semigroup operation, denoted by concatenation, it is also equipped with a binary operation, I shall denote by $\otimes$, such that $S \times S \to S$ given by $(s, t) \mapsto s \otimes t$ is a semigroup homomorphism. This means that if $(s, t), (s', t') \in S \times S$ then

$$(s \otimes t)(s' \otimes t') = ss' \otimes tt'.$$

Remark Observe that $(s \otimes t)^{-1} = s^{-1} \otimes t^{-1}$, since

$$(s \otimes t)(s^{-1} \otimes t^{-1})(s \otimes t) = ss^{-1}s \otimes tt^{-1}t = s \otimes t,$$

and

$$(s^{-1} \otimes t^{-1})(s \otimes t)(s^{-1} \otimes t^{-1}) = s^{-1}ss^{-1} \otimes t^{-1}tt^{-1} = s^{-1} \otimes t^{-1}.$$

In addition,

$$s \otimes t = (s \otimes 1)(1 \otimes t) = (1 \otimes t)(s \otimes 1).$$

These observations generalise to operations of arbitrary arity. Thus if Ω consists of one operator symbol of arity n, then an inverse Ω-algebra is equivalent to an inverse monoid equipped with n semigroup endomorphisms θ_i such that $\theta_i(s)\theta_j(t) = \theta_j(t)\theta_i(s)$ for all $s, t \in S$, and $\theta_i(1) = \theta_j(1)$ for all i and j.

An *inverse Ω-subalgebra* T of S is an inverse submonoid T which is also an Ω-algebra with respect to the induced operations. The intersection of any set of inverse Ω-subalgebras is again an inverse Ω-subalgebra so we can talk meaningfully of the inverse Ω-algebra generated by a subset.

We can now state the main theorem of this paper.

Theorem 2.3 (Correspondence Theorem) *With each operator domain Ω we can associate an inverse Ω-monoid $\mathcal{CM}_\Omega$, called the* clause monoid over Ω. *There is a bijection between the balanced Ω-varieties V and the wide inverse Ω-subalgebras $\mathcal{I}(\mathsf{V})$ of $\mathcal{CM}_\Omega$.* ■

3 Proof of the main theorem

Our main goal is to construct from each balanced Ω-variety V an inverse monoid denoted $\mathcal{I}(\mathsf{V})$. The construction of this inverse monoid is in fact an application of a general technique for constructing inverse monoids from categories acting suitably on sets (or groupoids), described in [16, 19]. However, I shall proceed without reference to this more general theory except to call upon it to substantiate results whose proofs would otherwise be tedious.

We begin by studying the properties of substitutions, since they play a big role in our construction. If A and B are subsets of X, the set of variables, we may consider homomorphisms $f\colon T(A) \to T(B)$ in the usual way. Every homomorphism $f\colon T(A) \to T(B)$ is determined by the values of f restricted to A, $f|A$, called a *substitution*. I denote the set A by $\mathrm{dom}(f)$.

Remark It is important to observe that in this paper a substitution will be a partial function: it is the restriction of f to A. This is different from the way that substitutions are usually handled, which is as functions defined on all variables but the identity on all but a finite number of them. However, the translation between the two approaches is straightforward.

A homomorphism $f\colon T(A) \to T(B)$ is said to be *full* if

$$B = \bigcup_{s \in T(A)} \mathbf{v}(f(s)).$$

Denote by $\mathcal{C}$ the set of algebras $T(A)$, where A is a finite subset of X, together with the full homomorphisms between them. A full homomorphism $f\colon T(A) \to T(B)$ is determined by a *full substitution*: the restriction $f|A$ has the property that each element of B occurs in $f(a)$ for some $a \in A$. Let $f\colon T(A) \to T(B)$ and $f'\colon T(A') \to T(B')$ be two full homomorphisms. Suppose that A and A' are disjoint. Then $(f|A) \cup (f'|A')$ is a function from $A \cup A'$ to $T(B \cup B')$ and so extends to a homomorphism from $T(A \cup A')$ to $T(B \cup B')$. It is clearly full. I shall denote this homomorphism by $f + g$, their disjoint union.

Lemma 3.1 *The set $\mathcal{C}$ is a right cancellative category in which the isomorphisms are those full homomorphisms $f\colon T(A) \to T(B)$ such that f induces a bijection from A to B.*

Proof We begin by characterising the full homomorphisms. Let $f\colon T(A) \to T(B)$ be full, and suppose that $x_b \in \mathbf{v}(f(s))$ for some $s \in T(A)$, then $x_b \in \mathbf{v}(f(x_a))$ for some $x_a \in A$. It follows that $f\colon T(A) \to T(B)$ is full iff for each $b \in B$ there exists $a \in A$ such that $b \in \mathbf{v}(f(a))$.

To show that $\mathcal{C}$ is a category, it is enough to note that the identity homomorphism defined on $T(A)$ is full, and that the composition of full homomorphisms is full follows from the characterisation of full homomorphisms above.

We now prove that it is right cancellative. Let

$$T(C) \stackrel{f}{\leftarrow} T(B) \stackrel{g}{\leftarrow} T(A)$$

and

$$T(C) \stackrel{f'}{\leftarrow} T(B) \stackrel{g}{\leftarrow} T(A)$$

be full homomorphisms such that $fg = f'g$. I shall prove that $f = f'$. Let $x_b \in B$. Since g is full there exists $x_a \in A$ such that the variable x_b occurs in the term $g(x_a)$. By assumption $f'(g(x_a)) = f(g(x_a))$. It follows that $f'(x_b) = f(x_b)$. Since x_b was an arbitrary variable in B we deduce that $f' = f$, as required.

The function $f\colon T(A) \to T(B)$ is an isomorphism if there is a function $g\colon T(B) \to T(A)$ such that gf is the identity on A, and fg is the identity on B. It follows that f induces a bijection between A and B. ■

The isomorphisms in $\mathcal{C}$ are called *renaming isomorphisms*.

If $f \in \mathcal{C}$ and $s \in T(X)$ then we define $f \cdot s = f(s)$ iff $\mathrm{dom}(f) = \mathbf{v}(s)$. The proofs of the following are straightforward.

Lemma 3.2

(i) *If $A = \mathbf{v}(s)$ then $1_A \cdot s = s$.*

(ii) *If $f \cdot s$ is defined and f: $T(A) \to T(B)$ then $\mathbf{v}(f \cdot s) = B$.*

(iii) *Let f: $T(A) \to T(B)$ and g: $T(B) \to T(C)$. Then $(gf) \cdot s$ is defined iff $g \cdot (f \cdot s)$ is defined, and they are equal.*

(iv) *If $f \cdot s = g \cdot s$ then $f = g$.* ■

Remark Parts (i)–(iii) of the above lemma tell us that the category $\mathcal{C}$ *acts* on the set of terms $T(X)$.

We say that a pair of terms s and t are *disjoint* if $\mathbf{v}(s) \cap \mathbf{v}(t) = \emptyset$. Such a pair of disjoint terms is said to be *unifiable* iff there exist substitutions f and g such that $f \cdot s = g \cdot t$.

Lemma 3.3 *Let $s = \rho(u_1, \ldots, u_m)$ and $t = \sigma(v_1, \ldots, v_n)$ be a pair of disjoint terms where ρ and σ are in the operator domain, and where the u_i are disjoint, and the v_j are disjoint. Then s and t are unifiable iff $\rho = \sigma$, so that $m = n$, and u_i and v_i are unifiable for $1 \leq i \leq m$.*

Proof Suppose that

$$f \cdot s = g \cdot t.$$

We may write $f = f_1 + \ldots + f_m$ and $g = g_1 + \ldots + g_n$, both disjoint unions, such that

$$f \cdot s = \rho(f_1 \cdot u_1, \ldots, f_m \cdot u_m)$$

and

$$g \cdot t = \sigma(g_1 \cdot v_1, \ldots, g_n \cdot v_n).$$

Thus $f \cdot s = g \cdot t$ implies $\rho = \sigma$, $m = n$ and $f_i \cdot u_i = g_i \cdot v_i$. Hence, in particular, f_i and g_i are unifiable. The converse is proved similarly. ■

The next result is a special case of the *unification algorithm* [2], adapted to our approach to substitutions.

Theorem 3.4 *Let s and t be a pair of disjoint unifiable terms. Then there are substitutions f and g such that $f \cdot s = g \cdot t$, and if f' and g' are substitutions such that $f' \cdot s = g' \cdot t$ then there is a substitution h such that $f' = hf$ and $g' = hg$.* ■

Notation We use the notation of the previous theorem. We put

$$f = \mathrm{mgu}_1(s, t) \,\mathrm{and}\, g = \mathrm{mgu}_2(s, t)$$

and

$$f \cdot s = s \wedge t = g \cdot t.$$

It is easy to deduce, using the fact that $\mathcal{C}$ is right cancellative, that f and g are unique upto a renaming isomorphism, and so $s \wedge t$ is unique upto a renaming

substitution.

Let s be a term. Define

$$\mathcal{C} \cdot s = \{f \cdot s \colon f \in \mathcal{C}\}.$$

This is therefore the set of all terms that can be obtained from the term s by using substitutions. The lemma below is crucial to our main construction.

Lemma 3.5 *Let s and t be terms. If $\mathcal{C} \cdot s \cap \mathcal{C} \cdot t \neq \emptyset$ then there is a term $s \wedge t$ such that*

$$\mathcal{C} \cdot s \cap \mathcal{C} \cdot t = \mathcal{C} \cdot (s \wedge t).$$

Proof By Lemma 3.2(iv), $\mathcal{C} \cdot s = \mathcal{C} \cdot s'$ iff s and s' differ by a renaming isomorphism. Thus without loss of generality, we may assume that $\mathbf{v}(s) \cap \mathbf{v}(t) = \emptyset$. Since $\mathcal{C} \cdot s \cap \mathcal{C} \cdot t \neq \emptyset$, the terms s and t are unifiable. By Theorem 3.4, there are substitutions f and g such that $f \cdot s = g \cdot t$, and if f' and g' are substitutions such that $f' \cdot s = g' \cdot t$ then there is a substitution h such that $f' = hf$ and $g' = hg$. Put

$$s \wedge t = f \cdot s = g \cdot t.$$

Then

$$\mathcal{C} \cdot (s \wedge t) \subseteq \mathcal{C} \cdot s \cap \mathcal{C} \cdot t.$$

Let $u \in \mathcal{C} \cdot s \cap \mathcal{C} \cdot t$. Then $u = f' \cdot s = g' \cdot t$ for some substitutions f' and g'. It follows from the properties given above that $u \in \mathcal{C} \cdot (s \wedge t)$. ■

Remark Let s and t be a pair of disjoint terms such that

$$\mathcal{C} \cdot s \cap \mathcal{C} \cdot t = \mathcal{C} \cdot u.$$

Then there exist substitutions f and t such that

$$u = f \cdot s = g \cdot t.$$

Suppose that $f' \cdot s = g' \cdot t$. Then $f' \cdot s = g' \cdot t = h \cdot u$ for some h. It follows that $(hf) \cdot s = f' \cdot s$ and $(hg) \cdot t = g' \cdot t$. By Lemma 3.2(iv), we have that $hf = f'$ and $hg = g'$. We deduce that for disjoint terms s and t, we have that $\mathcal{C} \cdot s \cap \mathcal{C} \cdot t \neq \emptyset$ iff s and t are unifiable, in which case $u = s \wedge t$, upto a renaming isomorphism.

We can now define the inverse monoid $\mathcal{I}(\mathsf{V})$ associated with the balanced variety V.

Proposition 3.6 *With each variety V defined by a balanced presentation, we can associate an inverse monoid $\mathcal{I}(\mathsf{V})$.*

Proof Let $\mathcal{G}$ be the fully invariant congruence corresponding to V. I shall denote elements of $\mathcal{G}$ by ordered pairs (s, t). Define a (partial) action of $\mathcal{C}$ on

$\mathcal{G}$ as follows: if $(s,t) \in \mathcal{G}$ and f: $T(A) \to T(B)$ is a full homomorphism then define

$$f \cdot (s,t) = (f \cdot s, f \cdot t) = (f(s), f(t))$$

if $A = \mathbf{v}(s)(= \mathbf{v}(t))$. The pair $(f(s), f(t))$ belongs to $\mathcal{G}$ because the congruence $\mathcal{G}$ is fully invariant. Define a relation $\preceq$ on $\mathcal{G}$ by

$$(s',t') \preceq (s,t)$$

iff there exists a substitution f such that $(s',t') = f \cdot (s,t)$. The relation $\preceq$ is a preorder: it is reflexive because each term s is fixed by the identity substitution defined on the variables occurring in s; it is transitive by composing full homomorphisms. It follows that the preorder induces an equivalence relation, which I shall denote by $\equiv$, on the set $\mathcal{G}$. The $\equiv$-equivalence class containing (s,t) will be denoted $[s,t]$. In addition, the set of $\equiv$-equivalence classes is ordered when we define $[s',t'] \leq [s,t]$ iff $(s',t') \preceq (s,t)$. Observe that by Lemma 3.2(iv), $[s,t] = [s',t']$ iff there is a there is a renaming isomorphism f such that

$$f \cdot s' = s \text{ and } f \cdot t' = t.$$

We define the set $\mathcal{I}(\mathsf{V})$ to consist of all the equivalence classes $[s,t]$ together with an adjoined zero 0. We extend the partial order we have defined to the whole of $\mathcal{I}(\mathsf{V})$ by making 0 the smallest element.

We now prove that we can define a multiplication on the set $\mathcal{I}(\mathsf{V})$ in such a way that it becomes an inverse monoid. First, the zero 0 behaves like a zero for the multiplication. Let $[s,t], [u,v] \in \mathcal{I}(\mathsf{V})$, both non-zero. Without loss of generality, we may assume that t and u have no variables in common. There are two possibilities:

(1) $\mathcal{C} \cdot s \cap \mathcal{C} \cdot u = \emptyset$. In this case, we define $[s,t][u,v] = 0$.

(2) $\mathcal{C} \cdot s \cap \mathcal{C} \cdot u \neq \emptyset$. In this case, by Lemma 3.5, put $f = \mathrm{mgu}_1(t,u)$ and $g = \mathrm{mgu}_2(t,u)$, and define $[s,t][u,v] = [f \cdot s, g \cdot v]$. Observe that, by assumption, the variables in s and t are the same as the variables in u and v respectively, so that both $f \cdot s$ and $g \cdot v$ are defined. In addition, from $s \approx t$ and $u \approx v$ we get $f \cdot s \approx f \cdot t$ and $g \cdot u \approx g \cdot v$ by (C4), and so $f \cdot s \approx g \cdot v \in \mathcal{G}$ by (C3).

We have to check that this multiplication is well-defined and associative: this can be done directly, but it follows from the general theory described in [16, 19]. This definition gives us the semigroup (with zero) structure of $\mathcal{I}(\mathsf{V})$.

It is straightforward to check that the nonzero idempotents of $\mathcal{I}(\mathsf{V})$ are those elements of the form $[s,s]$, where s is any term; this is a reflection of axiom (C1), and that the idempotents commute. If $[s,t] \in \mathcal{I}(\mathsf{V})$ then $[t,s] \in \mathcal{I}(\mathsf{V})$ by (C2). An easy calculation shows that

$$[s,t] = [s,t][t,s][s,t] \text{ and } [t,s] = [t,s][s,t][t,s].$$

Thus $\mathcal{I}(\mathsf{V})$ is an inverse semigroup in which $[s,t]^{-1} = [t,s]$. The natural partial order in this inverse semigroups agrees with the order we defined above. Thus

the form of the natural partial order is related to axiom (C4). Let $x \in X$ be any variable. Then $[s, s] \leq [x, x]$. It follows that $[x, x]$ is the maximum idempotent and so is the identity of the semigroup. Thus $\mathcal{I}(\mathsf{V})$ is an inverse monoid (with zero). ■

The monoid constructed above is called the *inverse monoid of the (balanced) variety* V.

Remark It is interesting to observe that we have used all the axioms defining a fully invariant congruence except (C5). We shall see below how this axiom can be interpreted in terms of the structure of $\mathcal{I}(\mathsf{V})$.

The set of *all* possible equations $s \approx t$ where $\mathbf{v}(s) = \mathbf{v}(t)$ is clearly a fully invariant congruence. The corresponding inverse monoid is denoted $\mathcal{CM}_\Omega$ and is called the *clause monoid (over* Ω*)*. The monoid $\mathcal{I}(\mathsf{V})$ is an inverse submonoid of $\mathcal{CM}_\Omega$ for any balanced variety of algebras of type Ω. The clause monoid was defined in [16], and was originally motivated by work of Girard in linear logic [12].

We shall now show how (C5) comes into play.

Proposition 3.7 *Let* V *be a balanced variety of* Ω*-algebras. Then* $\mathcal{I}(\mathsf{V})$ *is an* Ω*-algebra. In particular, for each* n*-ary operation* ρ *in* Ω*, we may define an* n*-ary operation* $\hat{\rho}$ *on* $\mathcal{I}(\mathsf{V})$ *in such a way that the following two properties hold:*

(i) $\hat{\rho}(\mathbf{f_1}, \ldots, \mathbf{f_n})\hat{\rho}(\mathbf{g_1}, \ldots, \mathbf{g_n}) = \hat{\rho}(\mathbf{f_1 g_1}, \ldots, \mathbf{f_n g_n})$.

(ii) $\hat{\rho}(\mathbf{f_1}, \ldots, \mathbf{f_n})^{-1} = \hat{\rho}(\mathbf{f_1}^{-1}, \ldots, \mathbf{f_n}^{-1})$.

Where the $\mathbf{f}_i$ *and the* $\mathbf{g}_i$ *are elements of* $\mathcal{I}(\mathsf{V})$*. In addition, if any of the* $\mathbf{f}_i = 0$ *then* $\hat{\rho}(\mathbf{f_1}, \ldots, \mathbf{f_n}) = 0$*. Finally, the function* $\hat{\rho}$ *is injective.*

Proof It is enough to prove (i), since the proof of (ii) follows from inverse semigroup theory and (i). We shall begin by showing how we may define the stated operations.

Let ρ be an n-ary operation symbol in Ω. We shall define an n-ary operation $\hat{\rho}$ on $\mathcal{I}(\mathsf{V})$. Let $[s_i, t_i] \in \mathcal{I}(\mathsf{V})$ where $1 \leq i \leq n$. I shall assume that $\mathbf{v}(s_i) \cap \mathbf{v}(s_j) = \emptyset$ for $i \neq j$. Since $\mathbf{v}(s_i) = \mathbf{v}(t_i)$ for each $1 \leq i \leq n$, it follows that $\mathbf{v}(t_i) \cap \mathbf{v}(t_j) = \emptyset$ for $i \neq j$ as well. Define the n-ary operation

$$\hat{\rho}([s_1, t_1], \ldots, [s_n, t_n]) = [\rho(s_1, \ldots, s_n), \rho(t_1, \ldots, t_n)].$$

Observe that the right-hand side is an element of $\mathcal{I}(\mathsf{V})$ because from $s_i \approx t_i$ for $1 \leq i \leq n$ we have that $\rho(s_1, \ldots, s_n) \approx \rho(t_1, \ldots, t_n)$ by (C5).

To show $\hat{\rho}$ is well-defined, we need only show that if $[s'_i, t'_i] = [s_i, t_i]$ are such that $\mathbf{v}(s'_i) \cap \mathbf{v}(s'_j) = \emptyset$ for $i \neq j$ then

$$[\rho(s_1, \ldots, s_n), \rho(t_1, \ldots, t_n)] = [\rho(s'_1, \ldots, s'_n), \rho(t'_1, \ldots, t'_n)].$$

Let f_i be the relabelling isomorphism that maps s_i to s'_i and t_i to t'_i. Then the disjoint union of the f_i is a relabelling isomorphism that maps

$$(\rho(s_1, \ldots, s_n), \rho(t_1, \ldots, t_n))$$

to

$$(\rho(s'_1, \ldots, s'_n), \rho(t'_1, \ldots, t'_n)).$$

It follows that $\hat{\rho}$ is a well-defined n-ary operation on $\mathcal{I}(\mathsf{V})$.

It remains to show that (i) holds. We shall prove that

$$\hat{\rho}([s_1, t_1], \ldots, [s_n, t_n])\hat{\rho}([u_1, v_1], \ldots, [u_n, v_n]) = \hat{\rho}([s_1, t_1][u_1, v_1], \ldots, [s_n, t_n][u_n, v_n]).$$

By definition

$$\hat{\rho}([s_1, t_1], \ldots, [s_n, t_n]) = [\rho(s_1, \ldots, s_n), \rho(t_1, \ldots, t_n)]$$

and

$$\hat{\rho}([u_1, v_1], \ldots, [u_n, v_n]) = [\rho(u_1, \ldots, u_n), \rho(v_1, \ldots, v_n)].$$

We now consider the product

$$[\rho(s_1, \ldots, s_n), \rho(t_1, \ldots, t_n)][\rho(u_1, \ldots, u_n), \rho(v_1, \ldots, v_n)].$$

Without loss of generality, we can assume that $\mathbf{v}(\rho(t_1, \ldots, t_n))$ and $\mathbf{v}(\rho(v_1, \ldots, v_n))$ are disjoint. By Lemma 3.3, these two terms are unifiable iff t_i and u_i are unifiable for $1 \leq i \leq n$. If $f_i \cdot t_i = g_i \cdot u_i$ then we may define $f = f_1 + \ldots + f_n$ and $g = g_1 + \ldots + g_n$, and it follows that

$$f \cdot \rho(t_1, \ldots, t_n) = g \cdot \rho(v_1, \ldots, v_n).$$

We deduce that the product

$$[\rho(s_1, \ldots, s_n), \rho(t_1, \ldots, t_n)][\rho(u_1, \ldots, u_n), \rho(v_1, \ldots, v_n)]$$

is non-zero precisely when all the products

$$[s_i, t_i][u_i, v_i]$$

are non-zero. This non-zero product will be

$$[f \cdot \rho(s_1, \ldots, s_n), g \cdot \rho(v_1, \ldots, v_n)],$$

which is equal to

$$[\rho(f_1 \cdot s_1, \ldots, f_n \cdot s_n), \rho(g_1 \cdot v_1, \ldots, g_n \cdot v_n)].$$

By definition this is just

$$\hat{\rho}([f_1 \cdot s_1, g_1 \cdot v_1], \ldots, [f_n \cdot s_n, g_n \cdot v_n])$$

which is equal to

$$\hat{\rho}([s_1, t_1][u_1, v_1], \ldots, [s_n, t_n][u_n, v_n]),$$

as required. This proves (i).

It remains to prove that $\hat{\rho}$ is injective. Suppose that

$$\hat{\rho}([s_1, t_1], \ldots, [s_n, t_n]) = \hat{\rho}([s'_1, t'_1], \ldots, [s'_n, t'_n]).$$

Then by definition

$$[\rho(s_1, \ldots, s_n), \rho(t_1, \ldots t_n)] = [\rho(s'_1, \ldots, s'_n), \rho(t'_1, \ldots t'_n)].$$

Thus

$$\rho(s_1, \ldots, s_n) = f \cdot \rho(s'_1, \ldots, s'_n) \text{ and } \rho(t_1, \ldots t_n) = f \cdot \rho(t'_1, \ldots t'_n)$$

for some renaming isomorphism f. From then proof of Lemma 3.3, we deduce that $[s_i, t_i] = [s'_i, t'_i]$ for $1 \leq i \leq n$, as required. ■

We can now prove the Correspondence Theorem.

Proof of Theorem 2.3 It is immediate from the definitions that for each balanced variety V, the inverse monoid $\mathcal{I}(\mathsf{V})$ is a wide inverse Ω-subalgebra of $\mathcal{CM}_\Omega$. Let V and V' be two balanced varieties such that $\mathcal{I}(\mathsf{V}) = \mathcal{I}(\mathsf{V}')$. Let (s, t) be a balanced equation holding in V. Then $[s, t] \in \mathcal{I}(\mathsf{V})$ and so $[s, t] \in \mathcal{I}(\mathsf{V}')$. Thus (s', t') is a balanced equation holding in V' where $(s', t') = f \cdot (s, t)$ and f is a renaming isomorphism. Thus $f^{-1} \cdot (s', t') = (s, t)$. By (C4), the equation (s, t) holds in V'. By symmetry, the varieties V and V' satisfy the same equations, and so $\mathsf{V} = \mathsf{V}'$ by Result 2.1. We have proved that $\mathcal{I}$ is injective. We now prove that it is surjective. Let S be an inverse submonoid of $\mathcal{CM}_\Omega$ satisfying the given conditions. Define

$$\mathcal{G} = \{s \approx t \colon [s, t] \in S\}.$$

We prove that $\mathcal{G}$ is a fully invariant congruence on $T_\Omega(X)$:

(C1) holds because S is a *wide* inverse submonoid.

(C2) holds because S is an *inverse* submonoid.

(C3) holds because S is a *submonoid*.

(C4) holds because S is a wide submonoid and so an order ideal; recall that $[s', t'] \leq [s, t]$ iff $s' = f \cdot s$ and $t' = f \cdot t$ for some element f of $\mathcal{C}$.

(C5) holds because S is an Ω-subalgebra.

Thus $\mathcal{G}$ is a fully invariant congruence. Clearly, $S = \mathcal{I}(\mathsf{V})$ where V is the Ω-variety corresponding to $\mathcal{G}$. ■

Inverse monoids are the abstract counterparts of inverse monoids of partial bijections. The clause monoid $\mathcal{CM}_\Omega$ has been defined in terms of equivalence classes of equations. To conclude this section, I shall show that it can be isomorphically represented by partial bijections on the term algebra.

Let s and t be terms. A function

$$\alpha\colon \mathcal{C} \cdot s \to \mathcal{C} \cdot t$$

is called a *$\mathcal{C}$-isomorphism* if it satisfies the following three conditions:

(IM1) It is a bijection.

(IM2) $\mathbf{v}(\alpha(u)) = \mathbf{v}(u)$, for each $u \in \mathcal{C} \cdot s$.

(IM3) $\alpha(f \cdot u) = f \cdot \alpha(u)$ for all $u \in \mathcal{C} \cdot s$ and $f \in \mathcal{C}$.

Denote the set of $\mathcal{C}$-isomorphisms of $T_\Omega(X)$ by $I_\Omega = I_\mathcal{C}(T_\Omega(X))$. For each n-ary operation symbol ρ in Ω, define the n-operation $\check{\rho}$ on I_Ω as follows. Let $f_1 \ldots, f_n$ be n elements of I. Then $\check{\rho}(f_1, \ldots, f_n)$ has as domain all terms of the form $\rho(u_1, \ldots, u_n)$ such that all $f_i \cdot u_i$ are defined. When this occurs define

$$\check{\rho}(f_1, \ldots, f_n)(\rho(u_1, \ldots, u_n)) = \rho(f_1 \cdot u_1, \ldots, f_n \cdot u_n).$$

Lemma 3.8 *I_Ω is an inverse Ω-algebra.*

Proof The proof that we have an inverse monoid is a special case of the results to be found in [16]. It remains to be proved that this inverse monoid is actually an inverse Ω-algebra. We prove first that $\check{\rho}(f_1, \ldots, f_n)$ is a $\mathcal{C}$-isomorphism. Let $f_i\colon \mathcal{C} \cdot s_i \to \mathcal{C} \cdot t_i$ be the $\mathcal{C}$-isomorphisms. Without loss of generality, we may assume that the $\mathbf{v}(s_i)$ are disjoint. It follows that

$$\check{\rho}\colon \mathcal{C} \cdot \rho(s_1, \ldots, s_n) \to \mathcal{C} \cdot \rho(t_1, \ldots, t_n).$$

To prove that we have an inverse Ω-algebra, we need to check that

$$\check{\rho}(f_1, \ldots, f_n)\check{\rho}(g_1, \ldots, g_n) = \check{\rho}(f_1 g_1, \ldots, f_n g_n),$$

which is straightforward. ∎

The semigroup-isomorphism part of the result below is just an application of Theorem 2.7 of [16].

Proposition 3.9 *There is an isomorphism*

$$\phi\colon \mathcal{CM}_\Omega \to I_\Omega$$

of inverse monoids and Ω-algebras. In addition, if V is a balanced Ω-variety, and if $[u, v] \in \mathcal{I}(\mathsf{V})$ then $s = \phi([u, v])(t)$ iff $s \approx t$ is an equation holding in V.

Proof Given $[s, t] \in \mathcal{CM}_\Omega$, we define $\phi([s, t])$ to be the function with domain $\mathcal{C} \cdot t$ and codomain $\mathcal{C} \cdot s$, and mapping $f \cdot t$ to $f \cdot s$. This function is well-defined, because if $u \in \mathcal{C} \cdot t$ then $u = f \cdot t$ and f is uniquely determined by u using Lemma 3.2(iv). If $[s', t'] = [s, t]$ then $s' = g \cdot s$ and $t' = g \cdot t$ for a renaming isomorphism g. Hence $\mathcal{C} \cdot s' = \mathcal{C} \cdot s$ and $\mathcal{C} \cdot t' = \mathcal{C} \cdot t$. If $u \in \mathcal{C} \cdot t'$ then $u = f' \cdot t'$.

It follows that $u = f' \cdot (g \cdot t)$ and so $u = (f'g) \cdot t$. Thus by Lemma 3.2(iv), we have that $f = f'g$. Now $u \mapsto f' \cdot s'$. But $f' \cdot s' = f'(g \cdot s) = f \cdot s$. We have shown that the function $\phi([s,t])$ is well-defined and independent of the choice of representative from the equivalence class $[s,t]$. Observe that $\phi([s,t])(t) = s$.

Suppose that $\phi([s,t]) = \phi([u,v])$. Then $u = \phi([s,t])(v)$. Thus $v \in \mathcal{C} \cdot t$ and so $v = f \cdot t$ and $u = f \cdot s$ for some f. Also $s = \phi([u,v])(t)$. Thus $t \in \mathcal{C} \cdot v$ and so $t = g \cdot v$ and $s = g \cdot u$. Hence $v = (fg) \cdot v$ and $t = (gf) \cdot t$. By Lemma 3.2(iv), f is a renaming isomorphism and so $[s,t] = [u,v]$. We have proved that ϕ is injective.

Let α: $\mathcal{C} \cdot s \to \mathcal{C} \cdot t$ be an arbitrary element of I_Ω. Let $\alpha(s) = t'$, where $t' \in \mathcal{C} \cdot t$. Thus $t' = f \cdot t$ for some f. Since α is a bijection, there is an element $g \cdot s$ such that $\alpha(g \cdot s) = t$. It follows that $(gf) \cdot t = t$. Thus by Lemma 3.2(iv), gf is the identity function on the variables of t. Now observe that

$$\alpha((fg) \cdot s) = \alpha(f \cdot (g \cdot s)) = f \cdot t = t' = \alpha(s).$$

By injectivity, $(fg) \cdot s = s$. Thus fg is the identity function on the set of variables of s. Hence f is a renaming isomorphism. By (IM2), we have that $\mathbf{v}(s) = \mathbf{v}(t')$. Thus $[t', s] \in \mathcal{CM}_\Omega$, and it is now immediate that $\phi([t', s]) = \alpha$.

We have proved that ϕ is a well-defined bijection. It only remains to check that ϕ is a homomorphism, but this can either be proved directly or follows from [16].

To prove that ϕ is isomorphism of Ω-algebras, we need to prove that

$$\phi\hat{\rho} = \check{\rho}\phi$$

where $\hat{\rho}$ is the n-ary operation on $\mathcal{CM}_\Omega$ and $\check{\rho}$ is the n-ary operation on I_Ω. But this follows from the definitions and the proof of Lemma 3.8.

Finally, we prove the last claim. Suppose $s \approx t$ holds in V. Then $[s,t] \in \mathcal{I}(\mathsf{V})$ and $s = \phi([s,t])(t)$. Now suppose that $s = \phi([u,v])(t)$ where $u \approx v$ holds in V. Then $t = f \cdot v$ and $s = f \cdot u$. By assumption, $u \approx v$ holds in V and so $f \cdot u \approx f \cdot v$ holds in V by (C4). Thus $s \approx t$ holds in V, as required. ■

4 An example: unary algebras

In this section, I shall interpret the Correspondence Theorem in the simplest interesting case: that of an operator domain Ω consisting only of unary operation symbols. Let $\Omega = \{l_1, \ldots, l_n\}$ where $n \geq 1$. We begin by describing the term algebra $T_\Omega(X)$. A typical term is

$$l_1(l_2(l_3(x)))$$

where $x \in X$. The first point to note is that the brackets here are not necessary. So this term can be written just as well as

$$l_1 l_2 l_3(x).$$

Let $L = \{l_1, \ldots, l_n\}$. The set L^* is the free monoid on L, whose identity we denote by ε. Each term in $T_\Omega(X)$ can be written unambiguously as $u(x)$ where $u \in L^*$ and $x \in X$; if $u = \varepsilon$ then $\varepsilon(x)$ is simply x. An equation over $T_\Omega(X)$ is balanced precisely when the variables ocurring on each side of the equation are the same. Elements of $\mathcal{CM}_\Omega$ have the form

$$[u_1(x), u_2(x)].$$

But x can be replaced by any variable in X. It follows that there is a bijection between the non-zero elements of $\mathcal{CM}_\Omega$ and pairs of strings (u_1, u_2). Consequently, define

$$P_n = \{(u_1, u_2)\colon\ u_1, u_2 \in L^*\} \cup \{0\}$$

We shall now show that P_n can be endowed with a multiplication that makes it isomorphic to $\mathcal{CM}_\Omega$. The key, of course, is to describe the unification algorithm for terms in $T_\Omega(X)$ for our choice of Ω.

Let $u(x)$ and $v(y)$ be a pair of terms where $x \neq y$. Then they are unifiable iff there are terms $u'(z)$ and $v'(z)$ such that $uu'(z) = vv'(z)$; this occurs iff u is a prefix of v, or v is a prefix of u. It follows that we should define the following product on P_n:

$$(u_1, u_2)(v_1, v_2) = \begin{cases} (u_1 w, v_2) & \text{if } v_1 = u_2 w \\ (u_1, v_2 w) & \text{if } u_2 = v_1 w \\ 0 & \text{else} \end{cases}$$

The set P_n equipped with this product is called the *polycyclic monoid on n generators.*[1] The monoid $\mathcal{CM}_\Omega$ is also an Ω-algebra. By Proposition 3.7, we may define $\hat{l_i}$ on P_n by

$$\hat{l_i}(u_1, u_2) = (l_i u_1, l_i u_2);$$

it maps zero to zero. Observe that in this case

$$\hat{l_i}(u_1, u_2) = (l_i, \varepsilon)(u_1, u_2)(l_i, \varepsilon)^{-1}.$$

We have therefore proved the following theorem.

Proposition 4.1 *Let Ω be an operator domain with n unary operation symbols. Then the clause monoid $\mathcal{CM}_\Omega$ is isomorphic to the polycyclic monoid on n generators.* ■

We now have to determine what the fully invariant balanced congruences on $T_\Omega(X)$.

Proposition 4.2 *There is a bijection between the fully invariant balanced congruences on $T_\Omega(X)$ and the semigroup congruences on L^*.*

[1] You can find out more about this monoid in my book [15]. The multiplication there is defined in terms of suffixes rather than prefixes, but this yields an isomorphic monoid.

Proof Let ρ be a fully invariant balanced congruence on $T_\Omega(X)$. Define

$$\rho^c = \{(u, v) \in L^* \times L^* \colon u(x) \approx v(x) \in \rho \text{ for some variable } x\}.$$

Since ρ is an equivalence relation, it follows that ρ^c is an equivalence relation; since ρ is fully invarient, ρ^c is a right congruence; since ρ is an Ω-congruence, ρ^c is a left congruence. Thus ρ^c is a congruence on L^*.

Let α be a congruence on L^*. Define

$$\alpha^a = \{u(x) \approx v(x) \colon x \in X \text{ and } u \,\alpha\, v\}.$$

It is easy to check that α^a is a fully invariant congruence on $T_\Omega(X)$, and that $(\alpha^a)^c = \alpha$ and $(\rho^c)^a = \rho$. ■

We shall now characterise the wide, inverse Ω-subalgebras of $\mathcal{CM}_\Omega$ where Ω consists only of unary operation symbols. An inverse submonoid S of P_n is said to be *positively self-conjugate* if

$$(u, \varepsilon)(v, w)(u, \varepsilon)^{-1} \in S$$

for each $(v, w) \in S$ and $(u, \varepsilon) \in P_n$.

Proposition 4.3 *The wide inverse submonoids of P_n that are also Ω-subalgebras are precisely the wide, positively self-conjugate inverse submonoids.*

Proof This follows from our description of the unary operations $\hat{l}_i$ defined on P_n prior to Proposition 4.1. ■

If we combine Propositions 4.2 and 4.3 with Theorem 2.3, we obtain the following version of the Correspondence Theorem for the case where the operator domain consists solely of unary operation symbols.

Theorem 4.4 *Let L^* be the free monoid on n generators. Then there is a bijection between the congruences on L^* and the wide, positively self-conjugate inverse submonoids of P_n.* ■

Remark Theorem 4.4 was first proved by Meakin and Sapir [20]. It follows that our Theorem 2.3 can be viewed as a generalisation of their result to all balanced varieties. Explicitly, if ρ is a congruence on L^* then the corresponding inverse submonoid of P_n consists of those pairs (u, v) such that $u \,\rho\, v$. ■

5 The linear case

The theory developed in Section 3 dealt with arbitrary balanced varieties. In this section, I shall concentrate on a class of balanced varieties called the 'linear'

varieties. These have nicer properties than arbitrary balanced varieties and so it will make sense to develop a version of the general theory tailored to this case.

A term s is said to be *linear* if each variable that occurs occurs exactly once. I shall denote by $T^l(X) = T^l_\Omega(X)$ the set of all *linear terms*. An equation $s \approx t$ is said to be *linear* if in addition to $\mathbf{v}(s) = \mathbf{v}(t)$ we have that each variable that occurs in s (respectively t) occurs exactly once. Thus linear equations are balanced equations in which both terms are linear. A variety is said to be *linear* if it has a presentation by means of linear equations.

Remark The fully invariant congruence associated with a linear variety need not consist entirely of linear equations. For example, if Ω consists of a single binary operation $*$, then the variety determined by the equation $x * y \approx y * x$ is linear. However $x * x \approx x * x$ belongs to the corresponding fully invariant congruence, and is clearly not linear.

In this section, a refinement of the Correspondence Theorem will be proved which holds for linear varieties. The Remark above shows that it is not an immediate corollary of Theorem 2.3.

Lemma 5.1 *Let s and t be a pair of disjoint linear terms. If they are unifiable with $f = \mathrm{mgu}_1(s,t)$ and $g = \mathrm{mgu}_2(s,t)$ then $f \cdot s$ and $g \cdot t$ are linear terms. In particular, the only obstruction to such terms being unifiable is that there is a mismatch between operators at some point.*

Proof The key to the proof of this lemma is Lemma 3.3 and the unification algorithm described, say, in [2]. I shall write a given pair of disjoint linear terms thus: $s =? t$. This indicates that we are trying to find substitutions that unify them. We now define an operation decompose:

- If $s = \rho(u_1, \ldots, u_n)$ and $t = \rho(v_1, \ldots, v_n)$ then we can carry out the operation decompose to obtain the set $u_1 =? v_1, \ldots, u_n =? v_n$. Observe that each pair of terms we get is linear and disjoint, and that all the u_i are disjoint from each other, as are the v_i.
- If $s = \rho(u_1, \ldots, u_m)$ and $t = \sigma(v_1, \ldots, v_n)$ where $\rho \neq \sigma$, then I shall say that the operation decompose *fails*

Let s and t be a pair of disjoint linear terms. Carry out the operation decompose on $s =? t$ and then iteratively on their results. This process will terminate in one of two ways: either decompose fails along the way — in which case it is clear that s and t are not unifiable — or there comes a point when the procedure decompose can no longer be applied. Let us suppose that the latter occurs and we end up with the following set of pairs of disjoint terms: $u_1 =? v_1, \ldots, u_n =? v_n$. Each equation must have one of the following forms: $x =? y$ where x and y are constants; $x =? y$ where x and y are variables; $x =? v$ where x is a variable and v is a term; $u =? y$ where u is a term and y is a variable. In the first case, if the constants are the same there is nothing to do, otherwise they fail to be

unifiable (a special case of a mismatch of operators); in all other cases, reorient the pairs so that a variable is on the left-hand side. It is immediate that the variables on the left-hand side are all distinct, and none of them occurs on the right-hand side. It follows from Section 4.6 of [2], that s and t are unifiable. Indeed, the substitutions f and g doing the job can be constructed from the set $u_1 =? v_1, \ldots, u_n =? v_n$ and so $f \cdot s$ (and $g \cdot t$) will be linear. ■

Remark It follows from the above lemma that when the operator domain consists of a single operation every pair of disjoint linear terms is unifiable.

The following is immediate from Lemma 5.1.

Lemma 5.2 *Let $[s,t], [s',t'] \in \mathcal{CM}_\Omega$ be such that $s \approx t$ and $s' \approx t'$ are both linear. Then if their product is nonzero it equals some $[u,v]$ where $u \approx v$ is linear.* ■

Lemma 5.2 implies that the set of all elements of $\mathcal{CM}_\Omega$ of the form $[s,t]$, where $s \approx t$ is linear, forms an inverse submonoid; we call it the *linear clause monoid (over Ω)*, denoted by $\mathcal{LCM}_\Omega$. The linear clause monoid is used by Abramsky [1] in formulating a theory of reversible computation.

Proposition 5.3 *The linear clause monoid over Ω is an inverse Ω-subalgebra of the clause monoid over Ω.*

Proof It is enough to prove that the Ω-algebra operations on $\mathcal{CM}_\Omega$ restrict to $\mathcal{LCM}_\Omega$. But this follows from the way these operations are defined in the proof of Proposition 3.7. ■

Remark In the case where Ω consists only of unary operation symbols (see Section 4), the clause monoid and the linear clause monoid are the same since a term cannot have a repeated variable. In all other cases, they are different.

For the purposes of this paper, I shall define an equation $s \approx t$ to be *strongly balanced* if it is balanced and the number of times a variable appears in s is equal to the number of times the same variable appears in t. In other words, we count occurrences according to their multiplicities.

Lemma 5.4 *Let $\mathcal{E}$ be a set of strongly balanced equations, and let $\mathcal{G}$ be the fully invariant congruence generated by $\mathcal{E}$. The every equation in $\mathcal{G}$ is strongly balanced.*

Proof We use Result 2.2. New equations are introduced by (EL1) and (EL2). In both cases, the equations introduced will be strongly balanced.

If $s \approx t$ is strongly balanced, then clearly $t \approx s$ is strongly balanced.

If $s \approx t$ and $t \approx u$ are both strongly balanced then $s \approx u$ is strongly balanced.

If $s \approx t$ is strongly balanced then $f(s) \approx f(t)$ is strongly balanced for every substitution f.

If $s_i \approx t_i$ is strongly balanced for $1 \leq i \leq n$ then $\rho(s_1, \ldots, s_n) \approx \rho(t_1, \ldots, t_n)$ is strongly balanced.

The result now follows by induction on the length of a derivation of an equation from $\mathcal{E}$. ■

We shall now modify some of the derivations introduced in Result 2.2:

(LEL2) $\mathcal{E} \vdash s \approx s$ for each *linear* term s.

(LEL5) If $\mathcal{E} \vdash s \approx t$ then $\mathcal{E} \vdash f(s) \approx f(t)$ where $s \approx t$ is a linear equation and for every substitution f where $f(a)$ is linear for each a in the domain of f and $\mathbf{v}(f(a)) \cap \mathbf{v}(f(b)) = \emptyset$ for $a \neq b$.

(LEL6) If $\mathcal{E} \vdash s_i \approx t_i$ for $1 \leq i \leq n$ where $s_i \approx t_i$ is a linear equation for each i and $\mathbf{v}(s_i) \cap \mathbf{v}(s_j) = \emptyset$ for $i \neq j$ and ρ is an n-ary operation then $\mathcal{E} \vdash \rho(s_1, \ldots, s_n) \approx \rho(t_1, \ldots, t_n)$.

Let $\mathcal{E}$ be a set of linear equations. If $\mathcal{E} \vdash s \approx t$ where only the operations (EL1), (LEL2), (EL3), (EL4), (LEL5), (LEL6) are used then I shall write

$$\mathcal{E} \vdash' s \approx t.$$

It is evident that

$$\mathcal{E} \vdash' s \approx t \,\text{implies}\, \mathcal{E} \vdash s \approx t$$

and

$$\mathcal{E} \vdash' s \approx t \,\text{implies}\, s \approx t \,\text{is linear.}$$

Lemma 5.5 *Let $\mathcal{E}$ be a set of linear equations, containing all equations of the form $s \approx s$ where s is linear, and let $s \approx t$ be a linear equation such that $\mathcal{E} \vdash s \approx t$. Then $\mathcal{E} \vdash' s \approx t$.*

Proof We shall prove that every equation appearing in the proof of $s \approx t$ is linear and that $\mathcal{E} \vdash' s \approx t$. We shall prove the result by induction on the length of the derivation of the equation. A linear equation derived in one step must be proved using either (EL1) or (EL2). In the case of (EL2), we are deriving a linear equation in one step and so we need only apply (LEL2).

Our induction hypothesis is that if $s \approx t$ is a linear equation such that $\mathcal{E} \vdash s \approx t$ in n or fewer steps then $\mathcal{E} \vdash' s \approx t$ and all equations that occur in this derivation are linear.

Let $s \approx t$ be a linear equation such that $\mathcal{E} \vdash s \approx t$ in $n+1$ steps. There are three possibilities to be considered.

Suppose that $s \approx t$ is derived by (EL3) from equations $s \approx u$ and $u \approx t$ which occur earlier in the proof. Observe that linear equations are strongly balanced. Thus by Lemma 5.4, every equation appearing in the derivation of $s \approx t$ from $\mathcal{E}$ will be strongly balanced. Hence both equations $s \approx u$ and $u \approx t$ are strongly balanced. By assumption, $s \approx t$ is linear. It follows that $s \approx u$ and $u \approx t$ are both linear. By the induction hypothesis, $\mathcal{E} \vdash' s \approx u$ and $\mathcal{E} \vdash' u \approx t$. Hence $\mathcal{E} \vdash' s \approx t$.

Suppose that $s \approx t$ is derived by (EL5) from an equation $s' \approx t'$ occurring earlier. It follows that $s = f(s')$ and $t = f(t')$ for some substitution f. We know that s' and t' contain the same variables counting multiplicities, and that s (and t) are linear. Hence s' and t' must be linear. By the induction hypothesis $\mathcal{E} \vdash' s' \approx t'$. It follows that we must in fact be applying (LEL5).

Suppose that $s \approx t$ is derived by (EL6) from earlier equations. Then $s = \rho(s_1, \ldots, s_n)$ and $t = \rho(t_1, \ldots, t_n)$ where $\mathcal{E} \vdash s_i \approx t_i$. Since $s \approx t$ is linear then all the $s_i \approx t_i$ must be linear and, in addition, the variables in the s_i must be disjoint from each other. By the induction hypothesis, $\mathcal{E} \vdash' s_i \approx t_i$. In addition, $s \approx t$ is derived by applying (LEL6). ■

If V is a linear variety we put

$$\mathcal{LI}(\mathsf{V}) = \mathcal{LCM}_\Omega \cap \mathcal{I}(\mathsf{V}).$$

It is a wide inverse Ω-subalgebra of the linear clause monoid.

Lemma 5.6 *Let S be a wide inverse Ω-subalgebra of the linear clause monoid. Put $\mathcal{E} = \{s \approx t\colon [s,t] \in S\}$. If $\mathcal{E} \vdash' u \approx v$ then $[u,v] \in S$.*

Proof This follows by Lemma 5.5. We prove the result by induction on the length of a derivation $\mathcal{E} \vdash' s \approx t$.

(EL1) and (EL2) come for free, the latter because our monoid is wide in the linear clause monoid.

(EL3) is closure under inverses.

(EL4) is a special case of closure under products.

(LEL5) follows from the fact that S is an order ideal of the linear clause monoid.

(LEL6) follows from the fact that S is an Ω-subalgebra. ■

We have the following refinement of Theorem 2.3 in the case of linear varieties.

Theorem 5.7 (Linear Correspondence Theorem) *With each operator domain Ω we can associate the* linear clause monoid $\mathcal{LCM}_\Omega$. *There is a bijection between the linear Ω-varieties V and the wide inverse Ω-subalgebras of $\mathcal{LCM}_\Omega$.*

Proof Observe first that $[s,t] \in \mathcal{LI}(\mathsf{V})$ iff $s \approx t$ is a linear equation holding in V. Let $\mathcal{E}$ be the set of all linear equations holding in V. By assumption they generate the fully invariant congruence that determines V. Let $\mathcal{E}'$ be the set of all linear equations holding in V' and generating the fully invariant congruence that determines V'. Suppose that $\mathcal{LI}(\mathsf{V}) = \mathcal{LI}(\mathsf{V}')$. Let $s \approx t \in \mathcal{E}$. Then $[s,t] \in \mathcal{LI}(\mathsf{V}) = \mathcal{LI}(\mathsf{V}')$. Thus $s \approx t \in \mathcal{E}'$. By a symmetrical argument it follows that $\mathcal{E} = \mathcal{E}'$. Thus $\mathsf{V} = \mathsf{V}'$.

We now prove that $\mathcal{LI}$ is surjective. Let S be an inverse submonoid of $\mathcal{LCM}_\Omega$ that satisfies the conditions of the theorem. Let

$$\mathcal{E} = \{s \approx t\colon [s,t] \in S\},$$

and let V be the variety generated by the linear equations $\mathcal{E}$. Clearly, $S \subseteq \mathcal{LI}(\mathsf{V})$. So it is enough ot prove that the reverse inclusion holds. To do this, we have to show that for every linear equation $s \approx t$ that holds in V, we must have that $[s,t] \in S$. But this follows from Lemmas 5.5 and 5.6. ■

We shall now modify Proposition 3.9 to obtain an isomorphic representation of the linear clause monoid by means of partial bijections.

Let $\mathcal{C}^l$ denote those full substitutions

$$f\colon A \to T^l(B)$$

(note codomain) such that the variables in $f(a)$ and $f(b)$ are disjoint when $a \neq b$.

Lemma 5.8 *The set $\mathcal{C}^l$ is a category such that the following properties hold.*

(i) *If $s \in T^l(A)$ and $f\colon A \to T^l(B)$ then $f \cdot s \in T^l(B)$.*

(ii) *Given $f\colon A \to T^l(B)$ and $s,t \in T^l(A)$. Then $f \cdot s = f \cdot t$ implies $s = t$.*

(iii) *Let s and t be disjoint linear terms such that $\mathcal{C}^l \cdot s \cap \mathcal{C}^l \cdot t \neq \emptyset$. Then there is a linear term $s \wedge t$ such that*

$$\mathcal{C}^l \cdot s \cap \mathcal{C}^l \cdot t = \mathcal{C}^l \cdot (s \wedge t).$$

It follows that $\mathcal{C}^l$ is cancellative.

Proof It is straightforward to check that $\mathcal{C}^l$ is a right cancellative category.

(i) Clear.

(ii) We have to prove the following. Let s,t be a pair of terms such that $\mathbf{v}(s) = \mathbf{v}(t)$ and such that no variable is repeated in s or in t, and suppose that f is such that $f \cdot s = f \cdot t = v$ where v has no repeated variable. Then $s = t$. We shall prove the result by induction. If s is a ground term then t is a ground term and the result holds. If s is a variable then t is the same variable and the result holds. Suppose now that $s = \rho(s_1, \ldots, s_n)$. Then because s and t are unifiable we must have that $t = \rho(t_1, \ldots, t_n)$. Let the variables occurring in s (and so in t) be $x_1, \ldots, x_m$. Let $f = f_1 + \ldots + f_m$ where $f_i(x_i) = u_i$. By assumption, $\mathbf{v}(u_i) \cap \mathbf{v}(u_j) = \emptyset$ if $i \neq j$. Write $f = f'_1 + \ldots + f'_n$ so that $f \cdot s = \rho(f'_1 \cdot s_1, \ldots, f'_n \cdot s_n)$, and $f = f''_1 + \ldots + f''_n$ so that $f \cdot t = \rho(f''_1 \cdot t_1, \ldots, f''_n \cdot t_n)$. It follows that $f'_i \cdot s_i = f''_i \cdot t_i$ for each i. Now f'_i is a disjoint union of some of the f_j, as is f''_i. From the definition of the f_j it follows that f'_i and f''_i must each be disjoint unions of the same f_j. Hence $f'_i = f''_i$. But then it follows that s_i and t_i must contain the same variables. By induction $s_i = t_i$, and so $s = t$, as required.

(iii) This is just a restatement of Lemma 5.1 combined with Lemma 3.5.

The fact that $\mathcal{C}^l$ is left cancellative, and so cancellative, follows from (ii). ■

The action of $\mathcal{C}$ on $T_\Omega(X)$ which follows as a result of Lemma 3.2 restricts to an action of $\mathcal{C}^l$ on $T^l_\Omega(X)$ by Lemma 5.8. We define a $\mathcal{C}^l$-*isomorphism* in

an analogous way to a $\mathcal{C}$-isomorphism. Denote by $I_\Omega^l = I_{\mathcal{C}^l}(T_\Omega^l(X))$ the set of all $\mathcal{C}^l$-isomorphisms. This forms an inverse monoid ([16]) and an Ω-algebra (Lemma 3.8). The proof of the following follows from Lemma 5.8, Theorem 2.7 of [16], and Proposition 3.9.

Proposition 5.9 *There is an isomorphism of inverse Ω-algebras*

$$\phi\colon \mathcal{LCM}_\Omega \to I_\Omega^l.$$

■

6 An example: linear clause monoids with one operation

In this section, Ω will be an operator domain with a single function symbol of arity n. The linear clause monoid over Ω will be denoted $\mathcal{LCM}_n$, and its isomorphic copy via Proposition 5.9 by I_n^l. Our goal is to show that these linear clause monoids are isomorphic to some monoids introduced by Scott [21] and developed by Birget [4] in the course of his work on the Thompson groups. We begin by defining precisely what these monoids are.

Let M be a monoid. A subset R of M is a *right ideal* if $RM \subseteq R$. A function $\theta\colon R \to R'$ is called a *right ideal isomorphism* if R and R' are right ideals of M and $\theta(rm) = \theta(r)m$ for all $r \in R$ and $m \in M$. The collection of right ideal isomorphisms of M is an inverse monoid. A right ideal R of M is said to be *essential* if $R \cap R' \neq \emptyset$ for all right ideals R'. The collection of right ideal isomorphisms between the essential right ideals of M is an inverse monoid. The proofs of the following can be found in Appendix A and Lemma 3.3 of [4].

Result 6.1 *Let M be a free monoid.*

(i) *Each right ideal R is of the form $R = ZM$ where Z is a uniquely determined prefix code.*

(ii) *Each essential right ideal R is of the form $R = ZM$ where Z is a maximal prefix code.*

(iii) *Each essential finitely generated right ideal R is of the form $R = ZM$ where Z is a maximal finite prefix code.* ■

Birget's paper is a good source of information on (maximal) prefix codes.

The collection of right ideal isomorphisms between the essential finitely generated right ideals of the free monoid on n generators is an inverse monoid, which I shall denote by T_n. The goal of this section can now be stated: to prove that $\mathcal{LCM}_n$ is isomorphic to T_n. By Result 6.1, the structure of T_n will be bound up with the properties of maximal prefix codes in the free monoid on n generators. We begin, therefore, by obtaining some properties of such codes.

Let $A_n = \{a_1, \ldots, a_n\}$ be a finite alphabet with n letters. The free monoid on A_n is A_n^*. Its elements are called *strings*. The empty string is denoted by ε. If $x = uv$ where x, u, and v are strings then u is called a *prefix* of x; it is *proper* if v is not empty. A pair of strings is said to be *prefix comparable* if one of the strings is a prefix of the other. Let C and D be (maximal) prefix codes in A_n^*. We define $C \leq D$ iff each element of C is a prefix of an element of D, and each element in D has an element of C as a prefix.

Lemma 6.2 *The relation defined above is a partial order on the set of (maximal) prefix codes.*

Proof It is clear that the relation is reflexive and transitive. Suppose that $C \leq D$ and $D \leq C$. Let $x \in C$. Then x is a prefix of some $y \in D$. In its turn, y is a prefix of some z in C. Thus x is a prefix of z and $x, z \in C$. Since C is a prefix code, it follows that $x = z$ and so, in particular, $x = y$. We have therefore shown that $C \subseteq D$. The reverse inclusion follows by symmetry. Hence $C = D$, as required. ■

Lemma 6.3 *Let C and D be maximal prefix codes. Then there is a maximal prefix code $C \wedge D$ which is a least upper bound of C and D with respect to the ordering $\leq$.*

Proof If x and y are strings define $x \wedge y$ to be x if y is a prefix of x, and y if x is a prefix of y, otherwise it is not defined. Observe that both x and y are prefixes of $x \wedge y$. Put

$$E = \{x \wedge y \colon x \in C, y \in D\}.$$

By construction and our observation above, each element of E has an element of C (resp. an element of D) as a prefix. In addition, each element of C (resp. D) is a prefix of an element of E. Let $x \in C$. Since D is a maximal prefix code, there is a string $y \in D$ such that x and y are prefix comparable. Thus $x \wedge y$ is defined and belongs to E. But x is a prefix of $x \wedge y$. By symmetry, every element of D is a prefix of an element of E.

Our proof that E is a maximal prefix code is in two steps. We begin by showing that E is a prefix code. Let $u, v \in E$ such that u is a prefix of v. By definition, $u = x \wedge y$ and $v = w \wedge z$ where $x, w \in C$ and $y, z \in D$. There are four cases to be considered:

- x is a prefix of y, and w is a prefix of z. Thus $u = y$ and $v = z$. But $y, z \in D$ and so $y = z$, since D is a prefix code, giving $u = v$.
- x is a prefix of y, and z is a prefix of w. Thus $u = y$ and $v = w$. Let $y = xa$ and $w = zb$. Since y is a prefix of w, we have that y and z are prefix comparable. Thus $y = z$ since $y, z \in D$, a prefix code. Hence $w = xab$ and so $w = x$ since $x, w \in C$. Thus $w = z$ and so $y = w$, as required.

- y is a prefix of x, and w is a prefix of z. Thus $u = x$ and $v = z$. Let $x = ya$ and $z = wb$. Now x is a prefix of wb. So x and w are prefix comparable. But $x, w \in C$, a prefix code. Thus $x = w$. It follows that $z = xb = yab$. But $y, z \in D$ and so $z = y$. Thus $z = x$, as required.
- y is a prefix of x and z is a prefix of w. Thus $u = x$ and $v = w$. Since $x, w \in C$ we have that $x = w$, and so $u = v$, as required.

Thus E is a prefix code. To show that E is a maximal prefix code, let z be any string. We show that z is prefix comparable to a string in E. Since C is a maximal prefix code, there is an $x \in C$ such that x and z are prefix comparable. Suppose that z is a prefix of x. Then we know that x is a prefix of an element of E and so z is a prefix of an element of E. Now suppose that x is a prefix of z. Since D is a maximal prefix code, z is prefix comparable with an element $y \in D$. It follows that x and y are prefix comparable and so $x \wedge y$ is defined and belongs to E. Clearly z and $x \wedge y$ are prefix comparable.

We have therefore proved that $C, D \leq E$.

To finish off, let E' be any maximal prefix code such that $C, D \leq E'$. We shall prove that $E \leq E'$. First, let $z \in E$. Then $z = x \wedge y$ for some $x \in C$ and $y \in D$. If $x \wedge y = x$ then x is a prefix of some element of E'; if $x \wedge y = y$ then y is a prefix of some element of E'. Thus in both cases, $z = x \wedge y$ is a prefix of an element of E'. Second, let $z \in E'$. Then there exists $x \in C$ such that x is a prefix of z, and there exists $y \in D$ such that y is a prefix of z. Thus x and y are prefix comparable and so $x \wedge y$ is defined and belongs to E. But $x \wedge y$ is a prefix of z. ■

Lemma 6.4 *Let Z and Z' be maximal prefix codes in A_n^*. Then $Z' \leq Z$ iff $ZA_n^* \subseteq Z'A_n^*$.*

Proof Let $Z' \leq Z$ and $x \in ZA_n^*$. Then $x = zy$ where $z \in Z$ and y is a string. By asumption, there exists $z' \in Z'$ such that $z = z'u$ for some string u. Thus $x = zy = z'uy$ and so $x \in Z'A_n^*$.

Let $ZA_n^* \subseteq Z'A_n^*$ and $z \in Z$. Then $z = z'u$ for some $z' \in Z'$ and string u. So z has as a prefix an element of Z'. Now let $z' \in Z'$. Because Z is a maximal prefix code, there exists $z \in Z$ such that z and z' are prefix comparable. If z' is a prefix of z then we are done. We therefore suppose that z is a prefix of z'. So let $z' = zu$ for some string u. By our subset inclusion, we have that $z = z''v$ for some $z'' \in Z'$ and string v. Thus $z' = z''vu$. But Z' is a prefix code and so $z' = z''$ giving $u = v\varepsilon$. Thus $z' = z$ and so z' is a prefix of an element of Z in this case as well. ■

The following is now immediate by Lemmas 6.3 and 6.4.

Proposition 6.5 *Let Z and Z' be finite maximal prefix codes. Then*

$$ZA_n^* \cap Z'A_n^* = (Z \wedge Z')A_n^*.$$

■

Under our assumption on the operator domain, a linear term s can be regarded as a rooted tree in which every interior vertex has outdegree n. Thus we may associate with s a maximal prefix code Z_s over the alphabet A_n. Each element of Z_s describes the position of a leaf and so of a variable.

Example The linear term $(x \otimes y) \otimes z$ is mapped to the maximal prefix code $\{a_1a_1, a_1a_2, a_2\}$.

Lemma 6.6 *Let s and t be linear terms. Then $Z_s \leq Z_t$ iff $t = f \cdot s$ for some substitution f.*

Proof Let s be a linear term and f a substitution such that $f \cdot s$ is linear. Then it is easy to check that $Z_s \leq Z_{f \cdot s}$.

Conversely, let t be a linear term such that $Z_s \leq Z_t$. We shall show that we can define a substitution f such that $t = f \cdot s$. Let $x \in Z_s$. Define Z to be that subset of Z_t consisting of all strings y such that x is a prefix of y. We assume that Z is non-empty. Let $x^{-1}Z$ be the elements of Z with the common prefix x removed. We prove that $x^{-1}Z$ is a maximal prefix code. From the fact that Z_t is a prefix code, it is easy to see that $x^{-1}Z$ is a prefix code. To show that it is a maximal prefix code, let w be any string. Consider the string xw. Since Z_t is a maximal prefix code there is a string $v \in Z_t$ such that xw and v are prefix comparable. Suppose that xw is a prefix of v. Then x is a prefix of v. Thus $v \in Z$ and so $x^{-1}v \in x^{-1}Z$ and w is a prefix of $x^{-1}v$. Suppose that v is a prefix of xw. If x is a prefix of v then we are back to the previous case. Thus v is a prefix of x; we show that this case cannot occur. Since $Z_s \leq Z_t$, there exists $x' \in Z_s$ such that x' is a prefix of v. But then x' is a prefix of x and so $x' = x$ giving $v = x$. This implies $x^{-1}Z$ is empty, contradicting our assumption.

For each x in Z_s define Z_x to consist of all those strings in Z_t that have x as a proper prefix. For those sets Z_x which are non-empty, we know by the result above that $x^{-1}Z_x$ is a maximal prefix code. For each $x \in Z_s$ for which $x^{-1}Z_x$ is non-empty define $f'(x)$ to be a linear term whose underlying tree is $x^{-1}Z_x$; in addition, if $x \neq x'$ ensure that the variables in $f'(x)$ are all different from the variables in $f'(x')$. Now t and $f'(s)$ will differ only by a renaming isomorphism g. Define $f = gf'$. This gives $t = f \cdot s$ as required. ■

Proposition 6.7 *Let the operator domain Ω consist of a single function symbol of arity n. The semilattice whose elements are the sets $\mathcal{C}^l \cdot s$, where s is a linear term, under the order of subset inclusion, is isomorphic to the semilattice ZA_n^* of finitely generated essential right ideals of A_n^* under subset inclusion.*

Proof Define

$$\Theta(\mathcal{C}^l \cdot s) = Z_s A_n^*.$$

This is well-defined since two linear terms which differ by a renaming isomorphism give rise to the same underlying maximal prefix code. It is evident that

Θ is a bijection. Now observe that $\mathcal{C}^l \cdot t \subseteq \mathcal{C}^l \cdot s$ iff $t = f \cdot s$ for some substitution f which, by Lemma 6.6, holds iff $Z_s \leq Z_t$. But by Lemma 6.4, $Z_s \leq Z_t$ iff $Z_t A_n^* \subseteq Z_s A_n^*$. It follows that the partially ordered sets are isomorphic and so the semilattices are isomorphic. ■

The proof of the following lemma is easy.

Lemma 6.8 *Let Z_i be n maximal prefix codes in A_n^*. Then*

$$\bigcup_{i=1}^{n} a_i Z_i$$

is a maximal prefix code. ■

Let

$$\beta_i \colon Z_i A_n^* \to Z_i' A_n^*$$

be n right ideal isomorphisms between finitely generated essential right ideals. By Lemma A2 of [4], β_i determines and is determined by the bijection it induces between Z_i and Z_i'. Define an n-ary operation $\tilde{\rho}$ on T_n as follows. The partial bijection $\tilde{\rho}(\beta_1, \ldots, \beta_n)$ has domain $(\bigcup a_i Z_i)A_n^*$, codomain $(\bigcup a_i Z_i')A_n^*$ and rule $\tilde{\rho}(a_i z_i) = a_i \beta_i(z_i)$. By Lemma 6.8 and Lemma A2 of [4], it is a well-defined element of T_n.

Proposition 6.9 *With the above definitions, T_n is an inverse Ω-algebra.*

Proof Let $\beta_i \colon Y_i A_n^* \to Y_i' A_n^*$ and $\gamma_i \colon Z_i A_n^* \to Z_i' A_n^*$ be two sets of n elements of I_n. We prove that

$$\tilde{\rho}(\beta_1, \ldots, \beta_n)\tilde{\rho}(\gamma_1, \ldots, \gamma_n) = \tilde{\rho}(\beta_1\gamma_1, \ldots, \beta_n\gamma_n).$$

By Lemma 6.3, we can deduce that

$$(\bigcup a_i Y_i)A_n^* \wedge (\bigcup a_i Z_i')A_n^* = \bigcup a_i(Y_i \wedge Z_i').$$

Thus the domain of

$$\tilde{\rho}(\beta_1, \ldots, \beta_n)\tilde{\rho}(\gamma_1, \ldots, \gamma_n)$$

is

$$(\bigcup_{i=1}^{n} a_i(\gamma^{-1}(Y_i \wedge Z_i')))A_n^*.$$

It is easy to check that this is also the domain of $\tilde{\rho}(\beta_1\gamma_1, \ldots, \beta_n\gamma_n)$. The equality of the two functions is now readily checked. ■

The main result of this section is the following.

Theorem 6.10 *The linear clause monoid $\mathcal{LCM}_n$ is isomorphic to T_n both as an inverse semigroup and as an Ω-algebra.*

Proof The proof is via the isomorphism of Proposition 5.9.

Observe that if $\alpha\colon \mathcal{C}^l \cdot s \to \mathcal{C}^l \cdot t$ is a $\mathcal{C}^l$-isomorphism then we may assume without loss of generality that $\alpha(s) = t$. To see why, suppose that $\alpha(s) = t' = f \cdot t$ for some f. Since α is surjective there exists $g \cdot s$ such that $\alpha(g \cdot s) = t$. Hence $t = (gf) \cdot t$. But then f is a renaming isomorphism by Lemma 3.2(iv). Thus $\mathcal{C}^l \cdot t = \mathcal{C}^l \cdot t'$ and $\alpha(s) = t'$.

Let

$$\alpha\colon \mathcal{C}^l \cdot s \to \mathcal{C}^l \cdot t$$

be a $\mathcal{C}^l$-isomorphism such that $\alpha(s) = t$. By assumption, s and t are linear terms with the same variables. Relabelling if necessary, I shall assume that the variables are $x_1, \ldots, x_m$. Denote by $w(s)_i$ the string in Z_s that describes the position of the variable x_i. Denote by $w(t)_i$ the string in Z_t that describes the position of the variable x_i. I shall define a right ideal isomorphism

$$\bar{\alpha}\colon Z_s A_n^* \to Z_t A_n^*.$$

To do this, it is enough to define a bijection from Z_s to Z_t. Define

$$\bar{\alpha}(w(s)_i) = w(t)_i.$$

We shall prove that $\alpha \mapsto \bar{\alpha}$ is an isomorphism of monoids. It is clearly a well-defined function. Observe that this map induces the isomorphism of semilattices established in Proposition 6.7.

Prove of injectivity: suppose that $\alpha\colon \mathcal{C}^l \cdot s \to \mathcal{C}^l \cdot t$ and $\beta\colon \mathcal{C}^l \cdot u \to \mathcal{C}^l \cdot v$ are such $\alpha(s) = t$ and $\beta(u) = v$, and that $\bar{\alpha} = \bar{\beta}$. Clearly $Z_s = Z_u$ and $Z_t = Z_v$ so the terms s and u (resp. t and v) have the same underlying trees. Let the variables occurring in s be $x_1, \ldots, x_m$. By definition, if $w(s)_i$ is the string labelling the position of x_i in s, then $\bar{\alpha}(w(s)_i)$ is the string labelling the position of x_i in t. Let the variables occurring in u be $y_1, \ldots, y_m$. By definition, if $w(u)_j$ is the string labelling the position of y_j in u, then $\bar{\beta}(w(u)_j)$ is the string labelling the position of y_j in v. If $w(s)_i$ is the string labelling the position of x_i in s then it labels the variable $j_{f(i)}$ in u. By abuse of notation, f induces a bijection from the x_i to the y_j. Thus $f \cdot s = u$. Now $\bar{\alpha}(w(s)_i) = w(t)_i$, $w(s)_i = w(u)_{f(i)}$, and $\bar{\alpha} = \bar{\beta}$. Hence $\bar{\alpha}(w(s)_i) = \bar{\beta}(w(u)_{f(i)}) = w(v)_{f(i)}$. We therefore have that $f \cdot t = v$. It is now clear that $\alpha = \beta$.

Proof of surjectivity: let $\alpha'\colon Z_1 A_n^* \to Z_2 A_n^*$ be an element of I_n. Observe that α' induces a bijection from Z_1 to Z_2. Let s be any linear term such that $Z_s = Z_1$ over the variables $x_1 \ldots, x_m$. Let t be a linear term such that $Z_t = Z_2$ and if $\alpha'(w) = w'$ where $w \in Z_1$ and w is labelled with the variable x_i in s then t is labelled by x_i at w'. Define $\alpha\colon \mathcal{C}^l \cdot s \to \mathcal{C}^l \cdot t$ by $\alpha(s) = t$. Then $\bar{\alpha} = \alpha'$.

The function $\alpha \mapsto \bar{\alpha}$ is order preserving: let $\alpha\colon \mathcal{C}^l \cdot s \to \mathcal{C}^l \cdot t$ and $\beta\colon \mathcal{C}^l \cdot u \to \mathcal{C}^l \cdot v$ be $\mathcal{C}^l$-isomorphisms such that β is a restriction of α. Then there is a substitution f such that $u = f \cdot s$ and $v = f \cdot t$. Thus by Lemma 6.6, we have that $Z_u A_n^* \subseteq Z_s A_n^*$ and $Z_v A_n^* \subseteq Z_t A_n^*$. Let $z \in Z_u$. Then $z = z'p$ for some $z' \in Z_s$ and string p. Let x_i be the variable in s that the string z' indicates. The string p picks out a variable y_j in $f(x_i)$. Thus $w(s)_i = z'$ and $w(f(x_i))_j = p$. By

definition $\bar{\beta}(w(u)_j) = w(v)_j$. $\bar{\alpha}(w(u)_j) = \bar{\alpha}(w(s)_i p) = \bar{\alpha}(w(s)_i)p = w(t)_i p$. However, from β a restriction of α we get that $v = f \cdot t$. Now x_j occurs uniquely in $f(x_i)$. Hence $w(v)_j = w(t)_i p$, as required.

The function $\alpha \to \bar{\alpha}$ preserves the groupoid product: let α: $\mathcal{C}^l \cdot s \to \mathcal{C}^l \cdot t$ and β: $\mathcal{C}^l \cdot t \to \mathcal{C}^l \cdot u$ be $\mathcal{C}^l$-isomorphisms. The proof that $\overline{\beta\alpha} = \bar{\beta}\bar{\alpha}$ is immediate from the definition.

To finish off, we have to show that the Ω-algebra structures are isomorphic. To that end, we need to prove that

$$\overline{\check{\rho}(f_1, \ldots, f_n)} = \rho(\bar{f}_1, \ldots, \bar{f}_n).$$

This is straightforward and hinges on the fact that

$$Z_{\rho(s_1,\ldots,s_n)} = \bigcup a_i Z_{s_i} = \tilde{\rho}(Z_{s_1}, \ldots, Z_{s_n}).$$

■

7 Dehornoy's geometry monoid

In Section 1, Dehornoy's geometry monoid constructed from a balanced variety V was briefly introduced. In this section, the relationship between the geometry monoid and the inverse monoid $\mathcal{I}(\mathsf{V})$ will be described.

Our starting point is Proposition 3.9, where an isomorphism

$$\phi\colon \mathcal{CM}_\Omega \to I_\Omega$$

was established. In addition, for each balanced variety V, and each $[u, v] \in \mathcal{I}(\mathsf{V})$, we have that $s = \phi([u, v])(t)$ iff $s \approx t$ is an equation holding in V. We deduce that the action of $\mathcal{CM}_\Omega$ on the set of terms induced by ϕ leads to an action of $\mathcal{I}(\mathsf{V})$ on the set of terms whose corresponding equivalance relation is $\approx$ — the set of equations that hold in V.

Notation I shall use ϕ to denote the restriction map to $\mathcal{I}(\mathsf{V})$, and write $\phi_{[s,t]}$ for $\phi([s, t])$ in what follows.

In view of our observation above, we shall now characterise those inverse submonoids of $\mathcal{I}(\mathsf{V})$ that induce the equivalence relation $\approx$ on $T_\Omega(X)$ via the homomorphism ϕ.

Proposition 7.1 *Let S be an inverse submonoid of $\mathcal{I}(\mathsf{V})$ where V is a balanced variety whose associated fully invariant congruence is $\mathcal{G}$. Then S induces the congruence $\approx$ via the homomorphism ϕ: $\mathcal{I}(\mathsf{V}) \to I_\Omega$ if and only if for each element $\mathbf{f} \in \mathcal{I}(\mathsf{V})$ there exists $\mathbf{s} \in S$ such that $\mathbf{f} \leq \mathbf{s}$.*

Proof Suppose that S induces the congruence $\approx$. Let $[s, t] \in \mathcal{I}(\mathsf{V})$. Thus $s \approx t \in \mathcal{G}$. By assumption, there exists $[u, v] \in S$ such that $\phi_{[u,v]}(t) = s$. By

the definition of ϕ, it follows that there is an element $f \in \mathcal{C}$ such that $s = f \cdot v$ and $t = f \cdot v$. Thus $[s,t] \leq [u,v]$, as required.

The converse is almost immediate. ■

Proposition 7.1 suggests the following definition. Let S be an inverse submonoid of an inverse monoid T. We say that S is a *dominating* inverse submonoid of T if for each $t \in T$ there exists $s \in S$ such that $t \leq s$. Using this terminology, we may paraphrase Proposition 7.1 in the following terms: the dominating inverse submonoids of $\mathcal{I}(\mathsf{V})$ are precisely the ones that induce the equivalence relation $\approx$ under the given action. We shall express this by saying that the dominating inverse submonoids of $\mathcal{I}(\mathsf{V})$ are the ones that are still 'closely linked' to the structure of the variety V.

We now reconsider the inverse monoid $\mathcal{I}(\mathsf{V})$. It contains a substantial amount of information about the balanced variety V and so could be used as an algebraic tool for studying the variety. It has the practical drawback, however, that it is defined in terms of *all* equations holding in V. We would like to find a smaller inverse monoid that is still 'closely linked' in to the structure of the variety. We shall eventually show that this leads naturally to Dehornoy's geometry monoid.

Let S be an inverse Ω-algebra. It follows that for each subset X of S we can define $\langle X \rangle$ to be the inverse Ω-subalgebra generated by X. We shall now give an explicit construction of $\langle X \rangle$. Let S be an inverse Ω-algebra, and let X be a subset of S. Define

$$\langle X \rangle_\Omega$$

to be the Ω-subalgebra generated by $X \cup \{1\}$, where 1 is the identity of S. For each subset Y of S define

$$\mathrm{Inv}(Y)$$

to be the inverse submonoid generated by Y.

Lemma 7.2 *Let S be an inverse Ω-algebra, and X a subset of S. Then*

$$\langle X \rangle = \mathrm{Inv}(\langle X \rangle_\Omega).$$

Proof We may paraphrase the result by saying that the inverse Ω-algebra generated by X is obtained by taking the inverse submonoid generated by the Ω-subalgebra generated by X.

Clearly

$$X \subseteq \mathrm{Inv}(\langle X \rangle_\Omega) \subseteq \langle X \rangle.$$

Since $\mathrm{Inv}(\langle X \rangle_\Omega)$ is an inverse submonoid of S, the theorem will be proved if we can show that it is an Ω-subalgebra of S. Let $\hat{\rho}$ be an n-ary operation on S, and let $s_1, \ldots, s_n \in \mathrm{Inv}(\langle X \rangle_\Omega)$. We shall prove that

$$\hat{\rho}(s_1, \ldots, s_n) \in \mathrm{Inv}(\langle X \rangle_\Omega).$$

By assumption,

$$s_{ij} = s_{i1} \ldots s_{im}$$

is a product where $s_{ij} \in \langle X \rangle_\Omega$. The fact that each s_i is written as a product of m elements is no loss in generality since we can pad products out to the requisite same length by using the identity of S. Now

$$\hat{\rho}(s_1, \ldots, s_n) = \hat{\rho}(s_{11} \ldots s_{1m}, \ldots, s_{n1} \ldots s_{nm}).$$

But this can be written as the product

$$\hat{\rho}(s_{11}, \ldots, s_{m1}) \ldots \hat{\rho}(s_{1m}, \ldots, s_{nm})$$

since $\hat{\rho}$ is a semigroup homomorphism from S^n to S. But

$$\hat{\rho}(s_{1j}, \ldots, s_{nj}) = \hat{\rho}(s_{1j}, 1, \ldots, 1) \ldots \hat{\rho}(1, \ldots, 1, s_{nj}).$$

Thus we have shown that $\hat{\rho}(s_1, \ldots, s_n)$ can be written as a product of elements of the form $\hat{\rho}(1, \ldots, 1, s, 1, \ldots, 1)$ where $s \in \langle X \rangle_\Omega$. But then

$$\hat{\rho}(1, \ldots, 1, s, 1, \ldots, 1) \in \langle X \rangle_\Omega.$$

It follows that

$$\hat{\rho}(s_1, \ldots, s_n) \in \mathrm{Inv}(\langle X \rangle_\Omega),$$

as required. ■

Let $\mathcal{E}$ be a non-empty subset of $\mathcal{G}$. We write $[\mathcal{E}]$ to denote the set of elements $[s, t]$ where $s \approx t \in \mathcal{E}$. Define $\mathcal{I}(\mathsf{V}, \mathcal{E})$ to be the inverse Ω-subalgebra of $\mathcal{I}(\mathsf{V})$ generated by $[\mathcal{E}]$. The action of $\mathcal{I}(\mathsf{V})$ on $T_\Omega(X)$ induces an action of $\mathcal{I}(\mathsf{V}, \mathcal{E})$ on $T_\Omega(X)$.

Proposition 7.3 *Let V be a balanced variety with fully invariant congruence $\mathcal{G}$, and let $\mathcal{E}$ be a non-empty subset of $\mathcal{G}$. Then $\mathcal{E}$ is a presentation of $\mathcal{G}$ if and only if the equivalence relation that $\mathcal{I}(\mathsf{V}, \mathcal{E})$ induces on $T_\Omega(X)$ is precisely $\approx$.*

Proof Suppose first that $\mathcal{E}$ is a presentation of $\mathcal{G}$. It is enough to prove that for each $u \approx v$ in $\mathcal{G}$ there exists $[s, t] \in \mathcal{I}(\mathsf{V}, \mathcal{E})$ such that $\phi_{[s,t]}(u) = v$. The equations of $\mathcal{G}$ are obtained from those of $\mathcal{E}$ by using the rules of equational logic described in Result 2.2. To prove the result we shall use the following observations. The identity of $\mathcal{I}(\mathsf{V}, \mathcal{E})$ is denoted by 1.

- For any term u we have that $1(u) = u$, and 1 belongs to $\mathcal{I}(\mathsf{V}, \mathcal{E})$.
- We have that $\phi_{[s,t]}(t) = s$ for each $[s, t] \in \mathcal{I}(\mathsf{V}, \mathcal{E})$.
- If $\phi_{[s,t]}(u) = v$ then $\phi^{-1}_{[s,t]}(v) = u$.
- If $\phi_{[s,t]}(u) = v$ and $\phi_{[s',t']}(v) = w$ then $\phi_{[s',t'][s,t]}(u) = w$.
- Suppose that $\phi_{[s,t]}(u) = v$ and that $f \in \mathcal{C}$ such that $f \cdot u$ and $f \cdot v$ are defined. Then $\phi_{[s,t]}(f \cdot u) = f \cdot v$.

- Suppose that $\phi_{[s_i,t_i]}(u_i) = v_i$ for $1 \leq i \leq n$ and $\hat{\rho}$ is an n-ary operation. Put $[s,t] = \hat{\rho}([s_1,t_1],\ldots,[s_n,t_n])$. Then

$$\phi_{[s,t]}(\rho(u_1,\ldots,u_n)) = \rho(v_1,\ldots,v_n).$$

Let $u \approx v \in \mathcal{G}$. Then $\mathcal{E} \vdash u \approx v$. We have to prove that there is $[s,t] \in \mathcal{I}(\mathsf{V},\mathcal{E})$ such that $\phi_{[s,t]}(u) = v$. We prove this result by induction on the length of a derivation. Assume result holds for all equations that can be derived in n or fewer steps. Then the above items show that we can prove the result for an equation that is derived in $n+1$ steps.

Conversely, suppose that the equivalence relation that $\mathcal{I}(\mathsf{V},\mathcal{E})$ induces on $T_\Omega(X)$ is precisely $\approx$. We prove that $\mathcal{E}$ is a presentation of $\mathcal{G}$. Observe first that if $[u,v] \in \mathcal{I}(\mathsf{V},\mathcal{E})$ then $\mathcal{E} \vdash u \approx v$. This follows by Lemma 7.2. Next, let $s \approx t$ be an element of $\mathcal{G}$. Then $[s,t] \in \mathcal{I}(\mathsf{V})$. By assumption, and Proposition 7.1, we have that $[s,t] \leq [u,v]$ for some element $[u,v] \in \mathcal{I}(\mathcal{E},\mathsf{V})$. It follows that $\mathcal{E} \vdash s \approx t$. Thus $\mathcal{E}$ generates $\mathcal{G}$, and so is a presentation. ■

By Propositions 7.1 and 7.3, we have the following.

Corollary 7.4 *Let V be a balanced variety with fully invariant congruence $\mathcal{G}$, and let $\mathcal{E}$ be a non-empty subset of $\mathcal{G}$. Then $\mathcal{E}$ is a presentation of $\mathcal{G}$ if and only if $\mathcal{I}(\mathsf{V},\mathcal{E})$ is a dominating inverse submonoid of $\mathcal{I}(\mathsf{V})$.* ■

Let $\mathcal{E}$ be a presentation of $\mathcal{G}$ with the property that for each $s \approx t$ belonging to $\mathcal{E}$ the element $[s,t]$ is a maximal element of $\mathcal{I}(\mathsf{V})$. Following Dehornoy, I shall say that such a presentation is *minimal* — the conflict in terminology comes about simply from the way that the order is defined in the inverse semigroup $\mathcal{I}(\mathsf{V})$.

Proposition 7.5 *Every balanced variety has a minimal presentation.*

Proof Let V be a balanced variety, and $\mathcal{G}$ its associated fully invariant congruence. We claim that every element of $\mathcal{I}(\mathsf{V})$ lies beneath a maximal element. To see why this is sufficient to prove the proposition, let $\mathcal{E}$ be all those equations $s \approx t$ such that $[s,t]$ is maximal in $\mathcal{I}(\mathsf{V})$. Then by our claim, $\mathcal{I}(\mathsf{V},\mathcal{E})$ is a dominating inverse submonoid of $\mathcal{I}(\mathsf{V})$ and so by Corollary 7.4, $\mathcal{E}$ is a presentation of V. Hence $\mathcal{E}$ is a minimal presentation.

To prove the claim we need some definitions. Let $s \approx t$ be a balanced equation. Its *operator complexity* is the sum of the total number of operators occurring in s and t. Its *variable complexity* is the number of distinct variables occurring in s. A full substitution $f\colon A \to T(B)$ is said to be a *variable substitution* if $f\colon A \to B$. In which case, f is a surjection. Observe that if $(s',t') = f \cdot (s,t)$ then the operator complexity of (s',t') is greater than or equal to the operator complexity of (s,t) since operators cannot be erased by a substitution.

Let $(s,t) \in \mathcal{G}$, and consider the set U of all $(u,v) \in \mathcal{G}$ such that $(s,t) = f \cdot (u,v)$ and where the operator complexity of (u,v) is a minimum. Amongst

these (u, v) pick one where the variable complexity in u is as large as possible. I claim that such a (u, v) is a maximal element of $\mathcal{I}(\mathcal{V})$. To see why, suppose that $[u, v] \leq [u', v']$. Then $(u, v) = f \cdot (u', v')$. Now there is a substitution f' such that $f' \cdot (u', v') = (s, t)$. Thus the operator complexity of (u', v') is less than or equal to that of (u, v). But from the choice of (u, v), the pairs (u, v) and (u', v') have the same operator complexity. Since operators cannot be erased by a substitution and none are inserted in this case, f must be a variable substitution. Now the variable complexity of (u, v) is as large as possible. Thus the variable complexities of (u, v) and (u', v') are the same. It follows that the domain and codomain of f have the same number of elements. But f is a surjective function between two finite sets with the same cardinality. Hence f is bijective, and so a renaming isomorphism. Thus $[u, v] = [u', v']$, as required. ■

Lemma 7.6 *Let $\mathcal{E}$ be a minimal presentation of $\mathcal{G}$. Then $\mathcal{I}(\mathsf{V}, \mathcal{E})$ is contained in every dominating inverse Ω-subalgebra of $\mathcal{I}(\mathcal{V})$.*

Proof Let S be a dominating inverse Ω-subalgebra of $\mathcal{I}(\mathcal{V})$ and let $s \approx t$ be an element of $\mathcal{E}$. By assumption, there exists $[u, v] \in S$ such that $[s, t] \leq [u, v]$. But $[s, t]$ is maximal and so $[s, t] = [u, v]$. Thus each generator of $\mathcal{I}(\mathsf{V}, \mathcal{E})$ belongs to S. It follows that $\mathcal{I}(\mathsf{V}, \mathcal{E})$ is a subset of S. ■

By Lemma 7.6 and Corollary 7.4, we deduce the following.

Corollary 7.7 *If $\mathcal{E}$ and $\mathcal{E}'$ are two minimal presentations of V. Then*

$$\mathcal{I}(\mathsf{V}, \mathcal{E}) = \mathcal{I}(\mathsf{V}, \mathcal{E}').$$

■

Let $\mathcal{E}$ be a minimal presentation of a balanced variety V. Define

$$\mathcal{D}(\mathsf{V}) = \mathcal{I}(\mathsf{V}, \mathcal{E}),$$

the geometry monoid associated with the balanced variety V. Corollary 7.7 shows that this monoid is independent of the minimal generating set used. In addition, we have shown that $\mathcal{D}(\mathsf{V})$ is a dominating inverse Ω-subalgebra of $\mathcal{I}(\mathsf{V})$ contained in ever other dominating inverse Ω-subalgebra of $\mathcal{I}(\mathsf{V})$. Thus it is the minimum dominating inverse Ω-subalgebra of $\mathcal{I}(\mathsf{V})$.

We may summarise by saying that the monoid $\mathcal{I}(\mathsf{V})$ is the most natural one attached to V in the sense that it uses all the identities, whereas the monoid $\mathcal{D}(\mathsf{V})$ is likely to be more useful for calculational purposes.

Proposition 7.8 *Let $\mathcal{E}$ be a minimal presentation of a variety V consisting of linear equations. Then the geometric monoid $\mathcal{D}(\mathsf{V})$ is an inverse submonoid of the linear clause monoid, and a dominating inverse submonoid of $\mathcal{LI}(\mathsf{V})$.*

Proof We prove that each element $[s,t] \in \mathcal{D}(\mathsf{V})$ is such that $s \approx t$ is linear. By Lemma 7.2, the proof of this result comes in two parts. First, $\langle [\mathcal{E}] \rangle_\Omega$ consists of linear elements: this follows by the proof of Proposition 3.7. Second, $\mathrm{Inv}(\langle [\mathcal{E}] \rangle_\Omega)$ consists of linear elements: the inverse of a linear element is linear, and the product of linear elements is linear by Lemma 5.2. By the above, we clearly have that $\mathcal{D}(\mathsf{V})$ is an inverse submonoid of $\mathcal{LI}(\mathsf{V})$ and so the result follows by Corollary 7.4. ■

8 From inverse monoids to groups

With each balanced variety V, we have associated inverse monoids $\mathcal{I}(\mathsf{V})$ and $\mathcal{D}(\mathsf{V})$ whose actions on the set of terms yield the fully invariant congruence associated with V in both cases. The goal of this section is to find some sufficient conditions for these inverse monoids to be replaceable by a group which is still tied to the structure of the original variety in the sense that the group can be made to act on the set of terms and induce by this action the associated fully invariant congruence. To do this, we need to outline first how groups (and their actions) may be constructed from inverse semigroups (and their actions).

We begin by summarising how to obtain groups from inverse semigroups; see [18] for details. If S is a semigroup, we write $S^* = S \setminus \{0\}$.[2] Let S be inverse (with zero). A function $\alpha\colon S^* \to H$ to a group such that $\alpha(st) = \alpha(s)\alpha(t)$ whenever $st \neq 0$ is called a *prehomomorphism*. It can be proved that $s \leq s'$ implies that $\alpha(s) = \alpha(s')$. With each inverse semigroup S, we can associate a group $G(S)$ and a prehomomorphism $\tau\colon S^* \to G(S)$, which is characterised by the following property: if $\alpha\colon S^* \to H$ is any prehomomorphism to a group then there exists a unique group homomorphism $\beta\, G(S) \to H$ such that $\beta\tau = \alpha$. We call $G(S)$ the *universal group* of S. Define the relation σ on S by $s\,\sigma\,t$ iff $a \leq s, t$ for some $a \in S$. Then σ is a congruence on S called the *minimum group congruence*. If S does not have a zero, then $G(S) = S/\sigma$.

An inverse semigroup S is said to be E^*-unitary if $0 \neq e \leq s$, where e is an idempotent, implies that s is an idempotent. It is said to be *strongly E^*-unitary* iff $\tau\colon S^* \to G(S)$ has the property that $\tau(s) = 1$ implies that s is an idempotent. Every strongly E^*-unitary inverse semigroup is E^*-unitary. An inverse semigroup without zero that is strongly E^*-unitary is said to be *E-unitary*.

We shall also need a slight weakening of the notion of a group action. Let G be a group and X a set. We say that G *acts partially* on the set X [14] if there is a partial function from $G \times X$ to X, denoted by $(g,x) \mapsto g \cdot x$, satisfying the following three conditions:

(PA1) $\exists 1 \cdot x$ for all $x \in X$ and $1 \cdot x = x$

(PA2) $\exists g \cdot (h \cdot x)$ implies that $\exists (gh) \cdot x$ and $g \cdot (h \cdot x) = (gh) \cdot x$.

[2] The reader is not going to be confused over this usage and the one in Section 6 where it was used to denote free monoids.

(PA3) $\exists g \cdot x$ implies that $\exists g^{-1} \cdot (g \cdot x)$ and $g^{-1} \cdot (g \cdot x) = x$.

If G acts partially on X, the action determines an equivalence relation on X just as in the case of a usual group action.

The following provides sufficient conditions for the action of an inverse monoid on a set to be replaceable by a group whose partial action yields the same equivalence relation.

Proposition 8.1 *Let $\phi\colon S \to I(X)$ be an injective monoid homomorphism from the inverse monoid S that maps the zero of S to the zero of $I(X)$, and let $\sim$ be the equivalence relation induced on X by S. Suppose that $\tau\colon S^* \to G(S)$ is the universal group of S, and that S is strongly E^*-unitary. Put $G = G(S)$. For each $g \in G$ and $x \in X$ define*

$$\exists g \cdot x$$

iff there exists $s \in S$ such that $\tau(s) = g$ and $x \in \mathrm{dom}(\phi(s))$, in which case we put

$$g \cdot x = \phi(s)(x).$$

Then G acts partially on the set X, and this partial action induces the equivalence relation $\sim$ on X.

Proof We show first that $g \cdot x$ is well-defined. Suppose that $\tau(s) = \tau(t) = g$ and that $x \in \mathrm{dom}(\phi(s))$ and $x \in \mathrm{dom}(\phi(t))$. I claim first that $st^{-1} \neq 0$. Suppose to the contrary that $st^{-1} = 0$. Then $s^{-1}st^{-1}t = 0$. Since ϕ is a homomorphism mapping the zero of S to the zero of $I(X)$ this implies that $\phi(s^{-1}s)\phi(t^{-1}t)$ is the empty function. But by assumption we have that $x \in \mathrm{dom}(\phi(s^{-1}s))$ and $x \in \mathrm{dom}(\phi(t^{-1}t))$, which implies that $\phi(s^{-1}s)\phi(t^{-1}t)$ is not the empty function. It follows that $st^{-1} \neq 0$. Hence

$$\tau(st^{-1}) = \tau(s)\tau(t^{-1}) = \tau(s)\tau(t)^{-1} = gg^{-1} = 1.$$

By assumption, we have that st^{-1} is a non-zero idempotent. Let $\phi(s)(x) = y$ and $\phi(t)(x) = z$. Then

$$\phi(st^{-1})(z) = \phi(s)\phi(t)^{-1}(z) = \phi(s)(x) = y.$$

But since st^{-1} is an idempotent, we must have that $z = y$. Thus our definition of $g \cdot x$ is well-defined.

We now have to check that the three axioms for a partial action hold. Axiom (PA1) is immediate. To show that (PA2) holds, suppose that $\exists g \cdot (h \cdot x)$. Let $\tau(s) = g$ and $\tau(t) = h$. Then $\phi(s)(\phi(t)(x))$ is defined. But ϕ is a homomorphism and so $\phi(st)(x)$ is defined. Clearly $st \neq 0$ and so $\tau(st) = \tau(s)\tau(t) = gh$. It follows that $\exists (gh)(x)$ and it equals $g \cdot (h \cdot x)$. Finally, we show that $\exists g \cdot x$ implies that $\exists g^{-1} \cdot (g \cdot x)$ and $g^{-1} \cdot (g \cdot x) = x$. Let $\tau(s) = g$. Then $g \cdot x = \phi(s)(x)$. Clearly $\phi(s)^{-1}(\phi(s)(x))$ is defined. But $\tau(s^{-1}) = \tau(s)^{-1} = g^{-1}$. Thus $\exists g^{-1} \cdot (g \cdot x)$. The result is now clear.

We now calculate the equivalence relation the partial action of G determines on X. Suppose that $g \cdot x = y$. Then if $\tau(s) = g$ we have that $\phi(s)(x) = y$ and

so $x \sim y$. Conversely, suppose that $x \sim y$. Then there exists $s \in S$ such that $\phi(s)(x) = y$. Put $g = \tau(s)$. Then $g \cdot x = y$. ■

The theorem above puts the spotlight on those inverse monoids that are strongly E^*-unitary. It is known that the problem of whether an inverse semigroup is strongly E^*-unitary or not is undecidable [22].

Examples

1. When the operator domain consists only of unary operations, the clause monoid and the linear clause monoid are the same. If there are n unary operations, then the linear clause monoid is isomorphic to the polycyclic monoid on n generators by Proposition 4.1. Such monoids are strongly E^*-unitary [17].

2. Consider the clause monoid in the case where there is a single binary operation symbol $\oplus$. Observe that

$$[z \oplus z, z \oplus z] \leq [x \oplus y, y \oplus x],$$

where $[z \oplus z, z \oplus z]$ is a non-zero idempotent, and $[x \oplus y, y \oplus x]$ is not an idempotent. Thus in this case the clause monoid is not E^*-unitary.

The two examples above tell us that if we are looking for strongly E^*-unitary clause semigroups, we should concentrate on the linear clause monoids. We are interested in such monoids because of Proposition 8.1. As a first step we have the following.

Proposition 8.2 *The linear clause monoid is E^*-unitary.*

Proof Let $[u, u] \leq [s, t]$ in $\mathcal{LCM}_\Omega$. Then there is an $f \in \mathcal{C}$ is such that $f(s) = f(t) = u$. Thus the result follows by Lemma 5.8(ii). ■

The question is: can we strengthen the above result to *strongly* E^*-unitary? An E^*-unitary semigroup without a zero is automatically strongly E^*-unitary. So we consider this case first.

Let S be an inverse semigroup with zero. We say that the *zero is removable* if $st = 0$ implies that $s = 0$ or $t = 0$. If the zero is removable, then $S^* = S \setminus \{0\}$ is itself an inverse semigroup. Observe that the zero is removable in S iff for all non-zero idempotents e and f the idempotent ef is non-zero.

Example If the operator domain consists of at least two symbols, then the corresponding linear clause monoid does not have a removable zero. For example, if there are two unary operation symbols then the linear clause monoid is the polycyclic monoid on two generators and it is easy to check that its zero is not removable.

Because of the above example, the following result is the best we can hope for in general.

Proposition 8.3 *The linear clause monoid over an operator domain consisting of a single operation has a removable zero.*

Proof It is enough to prove the following: let s and t be two terms on $T_\Omega(X)$ such that $\mathbf{v}(s) \cap \mathbf{v}(t) = \emptyset$ and such that neither s nor t contains any repeated variables; then s and t are unifiable. But this is true by the Remark following Lemma 5.1. ■

The proof of the following is now immediate by Propositions 8.2 and 8.3.

Theorem 8.4 *The linear clause monoid over an operator domain consisting of a single function symbol is E-unitary.* ■

The following question is open but natural.

Question Are the linear clause monoids over operator domains consisting of at least two function symbols always strongly E^*-unitary?

Examples

1. The universal group of the (linear) clause monoid over an operator domain having n unary function symbols is the free group on n generators [17].
2. The universal group of the linear clause clause monoid over an operator domain consisting of a single function symbol of arity n is the Thompson group $V_{n,1}$ by Section 6 and [4, 21]. In this case, the linear clause monoids are actually *F-inverse* [15].
3. The universal group of the geometry monoid is called the *geometry group* by Dehornoy. The equation $x \oplus (y \oplus z) \approx (x \oplus y) \oplus z$ is linear and generates the variety of semigroups. The geometry group in this case is the Thompson group F [8, 10].

Lemma 8.5 *Let S be a dominating inverse submonoid of an E^*-unitary inverse monoid T. Then S and T have the same universal groups.*

Proof Let $\phi, \phi' \colon T^* \to G$ be two prehomomorphisms to a group G that agree on S^*. Then they are equal. Let $t \in T^*$. By assumption, there exists $s \in S^*$ such that $t \leq s$. Now $\phi(t) = \phi(s)$ and $\phi'(t) = \phi'(s)$ and $\phi(s) = \phi'(s)$. Thus $\phi = \phi'$.

Let $\theta \colon S^* \to G$ be a prehomomorphism to a group G. Then there is a unique prehomomorphism $\phi \colon T^* \to G$ such that ϕ restricted to S^* is θ. Uniqueness follows by the above once we have proved existence. Let $t \in T^*$. Then by assumption, $t \leq s$ for some s in S^*. Define $\phi(t) = \theta(s)$. We show first that ϕ is a well-defined function. Suppose that $t \leq a, b$ where $a, b \in S^*$. Then $a^{-1}b$ and ab^{-1} are non-zero idempotents since T is E^*-unitary. Put $s' = aa^{-1}b \in S$. Then

$$aa^{-1}b = a(a^{-1}b) = a(a^{-1}b)^{-1} = ab^{-1}a = (ab^{-1})a.$$

Thus $s' \leq a, b$. Hence $\theta(a) = \theta(b)$. It is now straightforward to check that ϕ is a prehomomorphism.

Let γ: $S^* \to G(S)$ be the universal prehomomorphism to the universal group of S. Let Γ: $T^* \to G(S)$ be the unique extension of γ to T^*. We prove that it is universal for prehomomorphisms from T^* to groups.

Let α: $T^* \to H$ be a prehomomorphism to a group H. Then $(\alpha|S^*)$ the restriction of α to S^* is a prehomomorphism to H. Thus there exists a unique group homomorphism β: $G(S) \to H$ such that $\beta\gamma = (\alpha|S^*)$. Now observe that $\beta\Gamma = \alpha$; to prove this, let $t \in T^*$. Then $t \leq s$ for some $s \in S$. Then by definition $\Gamma(t) = \gamma(s)$. Hence

$$(\beta\Gamma)(t) = \beta(\Gamma(t)) = \beta(\gamma(s)) = (\alpha|S)(s) = \alpha(t).$$

To finish off, we have to prove that β is unique such that $\Gamma\beta = \alpha$. Let β': $G(S) \to H$ be a group homomorphism such that $\beta'\Gamma = \alpha$. To prove that $\beta' = \beta$ it is enough to show that $\beta'\gamma = (\alpha|S^*)$, but this is almost immediate from the definitions. ■

By Propositions 7.8 and 8.2 and Lemma 8.5, we now have the following.

Corollary 8.6 *Let V be a linear variety. Then $\mathcal{LI}(\mathsf{V})$ and $\mathcal{D}(\mathsf{V})$ have the same universal groups.* ■

An inverse Ω-algebra is said to be *injective* if each of the Ω-operations is an injective function. By Proposition 3.7, the inverse Ω-algebras $\mathcal{I}(\mathsf{V})$ are injective and so, therefore, are the $\mathcal{LI}(\mathsf{V})$ when the variety V is linear.

Proposition 8.7 *Let S be an (injective) inverse Ω-algebra which is E-unitary. Then the universal group $G(S)$ is an (injective) group Ω-algebra.*

Proof I shall prove the result for the case where Ω consists of a single function of arity 2. The general case is proved similarly. Let $\otimes$ denote the Ω-algebra operation on our inverse semigroup S satisfying the conditions of the proposition. We define an operation $\bullet$ on $G(S)$ as follows: if $g, h \in G(S)$ where $g = \sigma(s)$ and $t = \sigma(t)$, the define $g \bullet h = \sigma(s \otimes t)$. We show that this operation is well-defined. Suppose that $a \,\sigma\, b$ and $c \,\sigma\, d$. Now $(a \otimes c)^{-1}(b \otimes d) = a^{-1}b \otimes c^{-1}d$ and $a^{-1}b$ and $c^{-1}d$ are both idempotents. Thus $(a \otimes c)^{-1}(b \otimes d)$ is an idempotent. Similarly $(a \otimes c)(b \otimes d)^{-1}$ is an idempotent. Hence $a \otimes c \,\sigma\, b \otimes d$, as required. Thus the operation is well-defined. That the group becomes a group Ω-algebra is now straightforward to prove.

Now suppose that S is injective and that

$$\sigma(s) \bullet \sigma(t) = \sigma(u) \bullet \sigma(v).$$

Then

$$s \otimes t \,\sigma\, u \otimes v.$$

Hence $s^{-1}u \otimes t^{-1}v$ and $su^{-1} \otimes tv^{-1}$ are both idempotents. By injectivity, $s^{-1}u, t^{-1}v, su^{-1}, tv^{-1}$ are all idempotents. Thus $\sigma(s) = \sigma(u)$ and $\sigma(t) = \sigma(v)$. Hence $G(S)$ is an injective Ω-group. ■

If Ω is an operator domain consisting of one operation symbol of arity n, then an inverse Ω-algebra (respectively group Ω-algebra) will be called an *inverse n-algebra* (respectively *group n-algebra*).

By Theorem 8.4, Corollary 8.6 and Proposition 8.7 we deduce the following.

Theorem 8.8 *The geometry group of a linear variety over an operator domain consisting of a single operation of arity n is an injective group n-algebra.* ■

Remark Consider the above theorem in the case $n = 2$, and the linear variety of semigroups. The Thompson group F is known to be the geometry monoid of this variety. By the theorem above it is therefore equipped with an injective binary operation which is also a semigroup homomorphism from $F \times F$ to itself. It is this binary operation which is used by Brown in computing the homology of F [5].

References

[1] S. Abramsky, A structural approach to reversible computation, in *Proceedings of the International Workshop on Logic and Complexity in Computer Science* (editors D. Beauquier, Y. Matyasevich) LACL, 2001, 1–16.

[2] F. Baader, T. Nipkow, *Term rewriting and all that*, CUP, 1999.

[3] M. Barr, C. Wells, *Category theory for computing science*, Prentice Hall, 1990.

[4] J.-C. Birget, The groups of Richard Thompson and complexity, *Inter. J. Alg. Computation* **14** (2004), 569–626.

[5] K. S. Brown, The homology of Richard Thompson's group F, Preprint, 2004.

[6] P. M. Cohn, *Universal algebra*, D. Reidel, Dordrecht, 1981.

[7] P. Dehornoy, Structural monoids associated to equational varieties, *Proc. Amer. Math. Soc.* **117** (1993), 293–304.

[8] P. Dehornoy, The structure group for the associativity identity, *J. Pure and Appl. Alg.* **111** (1996), 59–82.

[9] P. Dehornoy, *Braids and self-distributivity*, Birkhäuser, 2000.

[10] P. Dehornoy, Geometric presentations for Thompson's groups, To appear in *J. Pure and Appl. Alg.*.

[11] C. Ehresmann, *Oeuvres complètes et commentées*, (ed A. C. Ehresmann) Supplements to *Cahiers de Topologie et Géométrie Différentielle* Amiens, 1980-83.

[12] J.-Y. Girard, The geometry of interaction III: accommodating the additives, in *Advances in linear logic* (eds J.-Y. Girard, Y. Lafont, L. Regnier) Cambridge University Press, 1995.

[13] Th. Ihringer, *Allgemeine Algebra*, Teubner Studienbücher, Stuttgart, 1988.

[14] J. Kellendonk, M. V. Lawson, Partial actions of groups, *Inter. J. of Alg. and Computation* **14** (2004), 87–114.

[15] M. V. Lawson, *Inverse semigroups: the theory of partial symmetries*, World Scientific, 1998.

[16] M. V. Lawson, Constructing inverse semigroups from category actions, *J. of Pure and Applied Alg.* **137** (1999), 57–101.

[17] M. V. Lawson, The structure of 0-E-unitary inverse monoids, *Proc. Edin. Math. Soc.* **42** (1999), 497–520.

[18] M. V. Lawson, E^*-unitary inverse semigroups, in *Semigroups, algorithms, automata and languages* (eds G. M. S. Gomes, J.-E. Pin, P. V. Silva) World Scientific, 2002, 195–214.

[19] M. V. Lawson, Constructing ordered groupoids, To appear in *Cahiers de topologie et géométrie différentielle catégoriques*.

[20] J. Meakin, M. Sapir, Congruences on free monoids and submonoids of polycyclic monoids, *J. Austral. Math. Soc. (Series A)* **54** (1993), 236-253.

[21] E. A. Scott, A construction which can be used to produce finitely presented infinite simple groups, *J. Alg.* **90** (1984), 294–322.

[22] B. Steinberg, The uniform word problem for groups and finite Rees quotients of E-unitary inverse semigroups, *J. Alg.* **266** (2003), 1–13.