ALGEBRAS DEFINED BY EQUATIONS

PETER M. HIGGINS AND MARCEL JACKSON

Abstract. We show that an elementary class of algebras is closed under the taking of homomorphic images and direct products if and only if the class consists of all algebras that satisfy a set of (generally simultaneous) equations. For classes of regular semigroups in particular this allows an interpretation of a universal algebraic nature that is formulated entirely in terms of the associative binary operation of the semigroup, which serves as an alternative to the approach via so called e-varieties. In particular we prove that classes of Inverse semigroups, Orthodox semigroups, and E-solid semigroups are equational in our sense.

1. Introduction

Groups may be characterized in terms of their binary operation alone as they form the class of semigroups that are both left and right simple, which is to say that a semigroup S is a group if and only if $aS = Sa = S$ for all $a \in S$. Since this introduces the theme of the paper let us observe that the given pair of conditions on S may be expressed by saying that the equations $ax = b$ and $ya = b$ are always solvable in S, meaning that the class $\mathcal{G}$ of all groups is defined within the class of semigroups by the equations:

$$\mathcal{G}\colon (\forall a, b \in S)\,(\exists x, y \in S) : \ (ax = b) \wedge (ya = b) \tag{1}$$

A second observation is that $\mathcal{G}$ is a class of semigroups closed under the operations H and P, which are respectively the taking of homomorphic images, and the taking of direct products, but $\mathcal{G}$ is not closed under the taking of subsemigroups, so that $\mathcal{G}$ does not represent a semigroup variety. Many fundamental semigroup classes are $\{\mathsf{H}, \mathsf{P}\}$-closed classes in this way and we may easily identify natural equational bases, as we show in Section 2. In general we will use the phrase *equation system* in preference to simply *equation*, to allow for the fact that they are typically systems of simultaneous equations (grouped by conjunction) and that we allow arbitrary quantification. This also avoids confusion with the common use of “equation” synonymously with “identity” in the context of varieties. Nevertheless, we allow *equational basis* to refer to any family of equation systems that characterise a class.

In Section 3 we prove the fundamental model theoretic theorem underlying this idea that being that an elementary class $\mathcal{C}$ of algebras is $\{\mathsf{H}, \mathsf{P}\}$-closed if and only if $\mathcal{C}$ consists of all algebras for which there exist solutions to a

Key words and phrases. Equational classes; regular semigroups; model theory.
The second author was partially supported by ARC Future Fellowship FT120100666.

certain set of equation systems. The reverse direction is clear but the forward implication is a consequence of Lyndon's positivity theorem (see [11] or Corollary 8.3.5 of [8]). In Section 4 we find equational bases for the $\{\mathsf{H}, \mathsf{P}\}$-classes of Inverse semigroups, Orthodox semigroups, and E-solid semigroups (semigroups whose idempotent generated part is a union of groups). The final section is on equation systems that are universally solvable in any semigroup.

General background on semigroup theory will be assumed. We direct the reader to the books [3, 7, 9, 10] and our textual source for universal algebra is [1]. Standard location theorems for Green's relations and properties of regularity will be used without further comment. The symbol S stands for a semigroup and we denote the set of idempotents of S by $E(S)$ or sometimes simply by E. We write $V(A)$ to denote the set of inverses of members of the subset A of S. One fact drawn upon in Section 4 is that in a regular semigroup S, $V(E^n) = E^{n+1}$, from which it follows that the idempotent generated subsemigroup $\langle E\rangle$ of S, is itself regular (see [6]).

2. Examples of equational bases for $\{\mathsf{H}, \mathsf{P}\}$-classes of semigroups

In the following, S always denotes a semigroup, and unless otherwise stated, quantification is over elements of S.

Example 2.1. (i) $\mathcal{R}eg$, the class of all *regular semigroups* may be defined by the single equation $a = axa$, which is to say

$$\mathcal{R}eg\colon (\forall a)\,(\exists x) : \; axa = a. \tag{2}$$

(ii) $\mathcal{CR}$, the class of all *completely regular semigroup*s (unions of groups) has an equational basis in our sense given by:

$$\mathcal{CR}\ :(\forall a)\,(\exists x) : \; (a = axa) \wedge (x = xax) \wedge (ax = xa) \tag{3}$$

for if $S \in \mathcal{CR}$ then for any $a \in S$ we take x as the group inverse of a in order to satisfy the equation system (and indeed that solution x is then unique). Conversely, given that S satisfies this equation system we have that $x \in V(a)$ and $a \mathscr{H} x$ as $a = a^2x = xa^2$ and $x = x^2a = ax^2$, so that H_a is a group and therefore S is a union of groups.

(iii) It is easy to show that the class $\mathcal{SL}$ of all *semilattices of groups* may be defined by augmenting the equations (3) for $\mathcal{SL}$ as follows:

$$(\forall a, b)\,(\exists x) : \; (a = axa) \wedge (x = xax) \wedge (ax = xa) \wedge (axb = bax).$$

However this class may also be defined by just two equations:

$$\mathcal{SL} : \; (\forall a, b)\,(\exists x, y) : \; (a = axa) \wedge (ab = bya). \tag{4}$$

To see this, given that $S \in \mathcal{SL}$ then the first equation is satisfied by regularity. Now $H_{ab} = H_{ba} = H$, a group with identity e say. Hence

$be, ea \in H$ so we put $y = e(be)^{-1}ab(ea)^{-1}e$ (where inversion is in the group H) and then

$$bya = be(be)^{-1}ab(ea)^{-1}ea = e(ab)e = ab.$$

Conversely if S satisfies the given equations then S is certainly regular. Take $a \in S$, $e \in E(S)$. Then there exists $y \in S$ such that $ae = eya$, whence $eae = e^2ya = eya = ae$. Similarly there exists $z \in S$ such that $ea = aze$, whence $eae = aze^2 = aze = ea$. Therefore $ae = eae = ea$, which shows that idempotents are central and we conclude that S is a semilattice of groups.

We note also that Theorems 5.1 and 5.2 of [12] show that the second equation over S^1 characterises semigroups in which $\mathscr{H}$ is a congruence such that $S/\mathscr{H}$ is commutative.

(iv) The class $\mathcal{CS}$ of *completely simple semigroups* is defined by the $\mathcal{CR}$ equations along with one other:

$$\mathcal{CS} : (\forall a, b)\,(\exists x, y) : \quad (a = axa) \wedge (x = xax) \wedge (ax = xa) \wedge (a = abay) \quad (5)$$

To see this we note that any completely simple semigroup S must satisfy the $\mathcal{CR}$ equation system (3) and since for any $a, b \in S$ we have $a \mathscr{H} aba$, it follows that there are always solutions to the final equality in (5) as well. Conversely, given that S satisfies the equation system we already have that S is completely regular while the final equality in (5) implies that $a \leq_{\mathscr{J}} b$ is true for all $a, b \in S$. Hence S is completely simple.

We may sometimes abbreviate certain collections of equalities by expressions that are shorter and the meaning of which is clearer. However, if the equalities required are simultaneous, meaning that they contain common variables, these abbreviations may not suffice and the equations may need to be listed explicitly to convey the required duplication of variables between equations. However the equations $((\forall a)\,(\exists x) : ax = xa = a)$ may be shortened to $x = 1$ and similarly $((\forall a)\,(\exists x) : ax = xa = x)$ can be written as $x = 0$. When dealing with long strings it is sometimes convenient to write the equation $w = w^2$ as $w \in E$, although this is an abuse of notation as w is a word in a free semigroup pre-image of S while $E = E(S)$. We adopt the convention in our equations that letters taken from the front of the alphabet, a, b, c are *parameters*, which means they are quantified by a $\forall$ symbol, while x, y, z denote *variables*, meaning that they are quantified by the symbol $\exists$.

Example 2.2. (i) Let $\mathcal{IG}$ be the class of all semigroups S for which each element has a group inverse: $\mathcal{IG}$ can be captured as the conjunction of equational properties as described above:

$$\mathcal{IG} : \ (\forall a)\,(\exists x, y) : \ (x \in V(a)) \wedge (x \mathscr{H} y) \wedge (y \in E(S))$$

(ii) By the class $\mathcal{C}r$ of *cryptogroups* is meant those semigroups S that are completely regular and for which $\mathscr{H}$ is a congruence (so that $\mathcal{CS} \subseteq \mathcal{C}r$). The class $\mathcal{C}r$ is defined by the $\mathcal{CR}$ equations (3) together with the equation systems defined by $ab \mathscr{H} axb$ and $ba \mathscr{H} bax$:

$$\mathcal{C}r : (\forall a, b)\,(\exists x) :$$

$$(a = axa) \wedge (x = xax) \wedge (ax = xa) \wedge (ab \mathscr{H} axb) \wedge (ba \mathscr{H} bax) \quad (6)$$

For supposing that S is a cryptogroup then S is completely regular and since $\mathscr{H}$ is a congruence and $a \mathscr{H} ax$ (as $ax = xa$) it follows that the additional equations are also satisfied. Conversely if S satisfies our equations then S is certainly a union of groups. Suppose that $a \mathscr{H} c$ in S. Then $e = ax = xa$ is the idempotent in the class $H_a = H_c$. Similarly there is a solution $x = d$ say to the given equations so that $c = cdc, d = dcd$ and $cd = dc$, from which it follows that $cd = ax = e$. It then follows from our equations that for any $b \in S$ we have $ab \mathscr{H} eb \mathscr{H} cb$ and by symmetry we obtain that $\mathscr{H}$ is also a left congruence, and therefore $\mathscr{H}$ is a congruence on S, which is to say that S is a cryptogroup.

Example 2.3. (i) *Semigroups with a right identity* (resp. *right zero*) are defined by the equation

$$(\exists x)\,(\forall a) : \ ax = a \,(\text{resp. } ax = x) \quad (7)$$

We also have of course the left and the two-sided versions of these, the two-sided cases respectively being the classes of *Monoids* ($\mathcal{M}$), and *Semigroups with zero*. In accord with the comment above, we may express these respectively via the equations $x = 1$ and $x = 0$. We do however explicitly call attention to this equational basis for $\mathcal{M}$:

$$\mathcal{M}\colon (\exists x)\,(\forall a) : \ ax = xa = a \quad (8)$$

The order of the logical quantifiers $\forall$ and $\exists$ in the equation systems of (1) to (6) is $\forall \ldots \exists \ldots$ whereas in (7) and (8) the order is reversed. What is more $\mathcal{M}$ cannot be represented by an equational basis of the form $\forall^* \exists^*$ (meaning any, possibly 0, number of $\forall$ followed by any, possibly zero, number of $\exists$) because any class defined in that way is closed under the taking of ascending chains of algebras (this is the easy half of The Chang-Łos-Suszko Preservation Theorem; see [2, Theorem 5.2.6]). However, this is not true of $\mathcal{M}$ as may be seen by considering the semilattice represented by the infinite ascending chain $E = e_1 < e_2 < \ldots$. The initial sub-chain $E_n = \{e_1 < e_2 < \cdots < e_n\}$ is a monoid with identity element e_n. We thus have an infinite ascending chain of semigroups $E_1 \leq E_2 \leq \ldots$ for which E, the union of the chain, is a semilattice that lacks an identity element, which is to say that the $E_n\,(n \geq 1)$ are all monoids yet their union E is not.

(ii) Semigroups with a maximum $\mathscr{J}$-class. The two element null semigroup N has a maximum $\mathscr{J}$-class, but its square $N \times N$ does not, so the class of semigroups with maximum $\mathscr{J}$-class is not P-closed. It turns out that this example is the main obstacle to being $\{\mathsf{H}, \mathsf{P}\}$-closed, as routine arguments show that the following properties are equivalent for a semigroup S:
- S has a maximum $\mathscr{J}$-class J and $S/(S-J)$ is not a null semigroup;
- S has a maximum $\mathscr{J}$-class J and the Rees quotient $S/(S-J)$ is *not* isomorphic to N;
- S satisfies equation (9):

$$(\exists x)\,(\forall a)\,(\exists y, z) : \,(yxz = a), \tag{9}$$

Equation (9) has three quantifier alternations, and we now show that this is necessary. By the Chang-Łos-Suszko Preservation Theorem, it suffices to show that there is a subsemigroup chain $A_1 \leq A_2 \leq \ldots$ satisfying Equation (9) and such that $\bigcup_{j\geq 1} A_j$ fails Equation (9), and a subsemigroup chain $B_1 \leq B_2 \leq \ldots$ failing Equation (9), but such that $\bigcup_{j\geq 1} B_j$ satisfies Equation (9). For the semigroups A_i we may use the semigroups E_i of Example 2.3(i): the union E fails (9). For B_j, we begin by considering the denumerably generated combinatorial Brandt semigroup $\mathbf{B}_\omega$, whose set of elements is $\{0\}\cup\{(i,j) \mid i,j \in \omega = \{0,1,2,\ldots\}\}$ with 0 acting as a multiplicative zero element and with multiplication

$$(i,j)(k,\ell) = \begin{cases} (i,\ell) & \text{if } j = k \\ 0 & \text{otherwise.} \end{cases}$$

For each $i = 1, 2, \ldots,$ choose B_i to be the subsemigroup of $\mathbf{B}_\omega$ on the set $\{0\} \cup \{(j,k) \mid j,k \leq i-1\} \cup \{(i,i)\}$, with two maximal $\mathscr{J}$-classes. Then each B_i fails (9), yet the union is $\mathbf{B}_\omega$, which has a single maximum $\mathscr{J}$-class and satisfies (9).

(iii) The class $\mathcal{J}$ of *simple semigroups* (semigroups with a single $\mathscr{J}$-class) are, by definition, defined by the condition that for all $a, b \in S$ there are solutions $x, y \in S^1$ to the equation $a = xby$. In those circumstances however, by replacing b by aba we may find solutions $u, v \in S^1$ such that $a = (ua)b(av)$ so that $x = ua$ and $y = av$ furnish solutions $x, y \in S$ that satisfy our equation $a = xby$. In summary we have the following equational basis for $\mathcal{J}$:

$$\mathscr{J}\text{-simple semigroups} : (\forall a, b)\,(\exists x, y) : \, a = xby \tag{10}$$

(iv) A non-example: $\mathcal{D}$ the class of bisimple semigroups. The class of *$\mathscr{R}$-simple semigroups* ($aS^1 = bS^1$ for all $a, b \in S$) has the equational basis: $((\forall a, b \in S)\,(\exists x \in S) : \, a = bx)$; the left-right dual to this defines the class of *$\mathscr{L}$-simple semigroups* while the conjunction of these equations defines the class of *$\mathscr{H}$-simple semigroups*, which coincides with $\mathcal{G}$, the class of all groups. There is something to check here however. Satisfaction of the equation $ax = b$ certainly guarantees $\mathscr{R}$-simplicity. Conversely $\mathscr{R}$-simplicity ensures a solution of $ax = b$ over S^1. Clearly for *distinct* elements a, b, the defining equation takes solutions from S, so only the case where $a = b$ is problematic. Since the one-element group satisfies every equation we may assume that $|S| \geq 2$. Then for any $a \in S$ take $b \in S$ such that $b \neq a$. Then there exists $x, y \in S$ such that $ax = b$ and $by = z$. It follows that $t = xy \in S$ is such that $a = at$, giving a solution in S for the case $a = b$ also. However the distinction between S and S^1 is important when it comes to $\mathscr{D}$, for it is unique among the five Green's relations in that the class of $\mathscr{D}$-simple (*bisimple*) semigroups is closed under H but not P. Bisimple semigroups are defined by the following triple disjunction of equational bases:

$$(\forall a, b)\,(\exists t, u, v, x, y) :$$
$$\big((a = tu) \wedge (t = av) \wedge (t = xb) \wedge (b = yt)\big) \vee \big((a = xb) \wedge (b = ya)\big)$$
$$\vee \big((a = bx) \wedge (b = ay)\big) \quad (11)$$

For suppose that S satisfies (11) and let $a, b \in S$. If the first equation set in (11) applies to a and b then $a \mathrel{\mathscr{R}} t \mathrel{\mathscr{L}} b$, while the second and third sets imply that $a \mathrel{\mathscr{L}} b$ and $a \mathrel{\mathscr{R}} b$ respectively. In any event it follows that S is bisimple. Conversely let S be any bisimple semigroup and let $a, b \in S$. Then there exists $x \in S$ such that $a \mathrel{\mathscr{R}} x \mathrel{\mathscr{L}} b$ and so we may satisfy the first equation set in (11) for a and b unless $a = x$ or $x = b$. If we have that $a = x$, then $a \mathrel{\mathscr{L}} b$ and the second equation set in (11) is solvable. Dually, if $b = x$ then $a \mathrel{\mathscr{R}} b$ and the third equation set in (11) can be solved for a and b. Hence if S is bisimple then S satisfies (11). It follows that the class of bisimple semigroups is closed under the taking of homomorphic images but, (as we now show), not under the taking of direct products and this is the reason why there is no redundancy in the list of disjunctions in (11) for, as our main theorem of Section 3 shows, if a class of algebras is closed under H and P then all but one of these sets of equations in such a disjunction will be redundant.

Let X be a countable infinite set. The *Baer-Levi semigroup* B is the subsemigroup of the full transformation semigroup T_X consisting of all one-to-one mappings $\alpha : X \to X$ such that $|X \setminus X\alpha|$ is infinite. It is well known and easily verified that B is $\mathscr{R}$-simple, right cancellative, and idempotent free; in consequence B is $\mathscr{L}$-trivial. In particular it follows that there are no factorizations of the form $a = ta$ in S (as then $ta = t^2a$ whence $t = t^2$ by right cancellativity) or what is the same, $a \notin Ba$ for all $a \in B$. Therefore B is an example of a bisimple semigroup that satisfies the third equation set in (11) but not the other two sets. Its left-right dual, B^*, will by symmetry also be bisimple and satisfy the second equation set in (11) but not the other two. (As another example of a semigroup that is left simple, left cancellative, idempotent free and hence $\mathscr{R}$-trivial, take the semigroup S of all surjections on X for which every kernel class is infinite.) The semigroup $B \times B^*$ is then an example of a direct product of two $\mathscr{D}$-simple semigroups that is not itself $\mathscr{D}$-simple: indeed $B \times B^*$ is $\mathscr{D}$-trivial (but $\mathscr{J}$-simple), by virtue of the following observation.

Proposition 2.4. *Let S (resp. T) be a semigroup that satisfies the condition that for all $a \in S$, $a \notin Sa$ (resp. for all $b \in T$, $b \notin bT$). Then $S \times T$ is $\mathscr{D}$-trivial.*

Proof. We first check that $S \times T$ is right trivial. Suppose that $(a,b) \mathscr{R} (c,d)$ say. Then either $(a,b) = (c,d)$ or there exists $(x,y),(u,v) \in S \times T$ such that $(a,b)(x,y) = (ax,by) = (c,d)$ and $(c,d)(u,v) = (cu,dv) = (axu,byv) = (a,b)$. But then we have $t = yv \in T$ and $b = bt$, contradicting that $b \notin bT$. Therefore it follows that $S \times T$ is right trivial. By symmetry it follows that $S \times T$ is also left trivial and hence $S \times T$ is $\mathscr{D}$-trivial. □

(v) Let $\mathcal{RG}$ denote the class of *right groups*, by which we mean semigroups that are right simple ($aS = S$) and left cancellative ($(ab = ac) \to (b = c)$). Another characterization of $\mathcal{RG}$ is the class of semigroups for which there is always a *unique* solution to the equation $ax = b$ ($a, b \in S$). The solvability of the equation $ax = b$ however does not in itself imply uniqueness: the Baer-Levi semigroup is an example of a right simple, right cancellative semigroup in which the equation $ax = b$ always has infinitely many solutions. However right groups are also characterized as those semigroups that are right simple and contain an idempotent (see [7] for details) and as such the class is determined by the availability of solutions to a pair of equations:

$$\mathcal{RG}\colon (\forall a, b)\,(\exists x, y) : \,(ax = b) \wedge (y = y^2) \tag{12}$$

(vi) Any *variety* $\mathcal{V}$ of semigroups (a class closed under the operators H, P, and S, the taking of subalgebras) is, by Birkhoff's theorem, defined by some countable set of identities, which are equations that may be expressed without the use of the existential symbol $\exists$. The following easy proposition is indicative of the kind of result that our approach leads to: it shows for example that the equation systems holding in a variety $\mathcal{V}$ are precisely those holding on the denumerably generated $\mathcal{V}$-free algebra (precisely as is the case for identities).

Proposition 2.5. *Let $\mathcal{K}$ be a class of algebras in a countable signature that is defined by a family of equation systems. Then $\mathcal{K}$ is a variety if and only if $\mathcal{K}$ contains the denumerably generated $\mathsf{HSP}(\mathcal{K})$-free algebras.*

Proof. The forward direction is trivial. Now assume that the denumerably generated $\mathsf{HSP}(\mathcal{K})$-free algebras lie in $\mathscr{K}$. Then as $\mathsf{H}(\mathcal{K}) = \mathcal{K}$ it follows that $\mathscr{K}$ contains all countably generated members of $\mathsf{HSP}(\mathcal{K})$. However as $\mathcal{K}$ is defined by a family of equation systems it is an elementary class (definable in first order logic) and hence is determined by its countably generated members. As these coincide with $\mathsf{HSP}(\mathcal{K})$ it follows that $\mathsf{HSP}(\mathcal{K}) = \mathcal{K}$. □

(vii) The dual idea to that which arises in (vi) is of a class defined by an equation set that is free of the symbol $\forall$. For example the class:

$$\mathcal{I}d\colon (\exists\, x) :\, x = x^2 \tag{13}$$

is the class of all semigroups S with idempotents, which is to say that $E(S) \neq \emptyset$. We note that $\mathcal{I}d$ is the minimum semigroup class of this kind as any semigroup S with an idempotent e satisfies every equation $p = q$ that is free of the $\forall$ quantifier as is seen by acting the substitution $x \to e$ on each variable x of $p = q$. We return to this topic in Section 5.

3. The equational representation theorem for $\{\mathsf{H}, \mathsf{P}\}$-classes

A formula of the predicate calculus is in *prenex form* if it is written as a string of quantifiers (referred to as the *prefix*) followed by a quantifier-free part (referred to as the *matrix*). An *equation system*, as informally described in Section 1, is a sentence in prenex form, whose matrix is a conjunction of atomic formulas. Familiar examples include identities (universally quantified equation systems) and primitive positive sentences (extensionally quantified equation systems). When the quantifiers are all universal, we also refer to a $\forall_1$ equation system, while a primitive positive sentence will be referred to as a $\exists_1$ equation system. Inductively, a $\forall_{i+1}$ equation system is an equation system of the form $(\forall x_1 \dots \forall x_{n_{i+1}})\phi(x_1, \dots, x_{n_{i+1}})$ where $\phi(x_1, \dots, x_{n_{i+1}})$ is an $\exists_i$ equation system, with a dual definition for an $\exists_{i+1}$ system.

We note that in the case of $\forall_1$ equation systems, we may use the property

$$\big((\forall x)\, \phi(x) \wedge \psi(x)\big) \leftrightarrow \big((\forall x)\, \phi(x)\big) \wedge \big((\forall x)\, \psi(x)\big)$$

in order to remove conjunctions (in favour of sets of quantified atomic formulas), however this is not in general true once existentially quantified variables are present. Equation systems are exactly the *positive* (that is, negation-free) *Horn sentences* (sentences with at most one *positive literal* or *atom*).

The following result is an extension of Lyndon's Positivity Theorem (see [11], or [8, Corollary 8.3.5]), and applies in all signatures, including those involving relations. The result is in the style of the many classical preservation theorems of model theory, though is not to be found in standard references such as [2] and [8], nor in other surveys such as [13]. In the case of relations, by a surjective homomorphic image, we mean there is a homomorphism that maps its domain onto the co-domain; it does not necessarily map the relations on the domain structure onto those of the co-domain structure.

Theorem 3.1. *An elementary class equals the class of models of some family of equation systems if and only if it is closed under taking homomorphic images of direct products. If the elementary class is the model class of a single sentence, then it is a class of models of a single equation system.*

Proof. One direction is easy: equation systems are preserved under direct products and under homomorphic images. We now must show that if $\mathcal{K}$ is an elementary class closed under taking surjective homomorphic images and direct products, then it can be axiomatised by a family of equation systems. Our method of proof will automatically derive the second statement in the theorem. The hard work is performed by Lyndon's positivity theorem, which states that a sentence closed under taking surjective homomorphic images is equivalent to a positive sentence. Thus there is no loss of generality in assuming that $\mathcal{K} = \{\mathsf{H}, \mathsf{P}\}(\mathcal{K})$ is the class of models of a set Σ of positive sentences. Our remaining task is to show that disjunctions can be removed from these sentences.

Consider a sentence from Σ; we assume it is a $\forall_t$, for, if we are given a $\exists_t$ sentence, we may augment Σ with the initial condition $(\forall a)$, where a is a symbol that does not appear elsewhere in Σ, and so replace Σ with an equivalent $\forall_{t+1}$ sentence. Therefore we may take it that the quantifier Q_t is $\forall$ if t is odd and $\exists$ if t is even. We may write our sentence as:

$$\begin{aligned}\rho = (\forall x_{1,1} \ldots \forall x_{1,n_1})(\exists x_{2,1} \ldots \exists x_{2,n_2}) \ldots (Q_k x_{k,1} \ldots Q_k\, x_{k,n_k})\\ \rho(x_{1,1}, \ldots x_{1,n_1}, x_{2,1}, \ldots, x_{2,n_2}, \ldots, x_{k,1} \ldots, x_{k,n_k}).\end{aligned}$$

Moreover, there is no loss of generality in assuming that k is even, as we may, if necessary, append a final $(\exists x)$ quantifier, where the symbol x does not appear in the matrix ρ, giving an equivalent sentence. We assume that the matrix of ρ is written as a finite conjunction of disjunctions; say $\bigwedge_{1 \le i \le m} \gamma_i$, where each γ_i is a finite disjunction:

$$\gamma_i = \alpha_{i,1} \vee \cdots \vee \alpha_{i,r_i}$$

where each $\alpha_{i,j}$ is an atomic formula involving some subset of the full set of variables $x_{1,1}, \ldots, x_{k,n_k}$. If $r_i = 1$ for $i = 1, \ldots, m$ then there is nothing to prove. Otherwise, if there is i such that $r_i \ge 2$, we shall show that there is a $j \in \{1, \ldots, r_i\}$ such that the conjunct γ_i may be replaced by the single atomic formula $\alpha_{i,j}$. Repeating this for each conjunct will see us arrive at the desired $\vee$-free sentence. The quantifiers remain unchanged throughout.

Without loss of generality then, we may assume that $r_i \ge 2$ for some i. For each $j = 1, \ldots, r_i$ let ρ_j be the result of replacing γ_i by $\alpha_{i,j}$ in ρ. Note that $\rho_j \vdash \rho$ so that the class of models satisfied by $(\Sigma \cup \{\rho_j\}) \setminus \{\rho\}$ is a subclass of $\mathcal{K}$. We wish to show that there is some j such that the reverse containment holds.

Assume by way of contradiction that no such j exists. In this case, for each $j \in \{1, \ldots, r_i\}$ there is a model $\boldsymbol{M}_j \in \mathcal{K}$ such that ρ_j fails in $\boldsymbol{M}_j$. We will now use the fact that $\boldsymbol{M} := \Pi_{1\le j\le r_i}\boldsymbol{M}_j \in \mathcal{K}$ and so $\boldsymbol{M} \models \rho$ in order to produce the required contradiction.

For each j: as $\boldsymbol{M}_j \not\models \rho_j$ there is an n_1-tuple $a_{1,1,j}, \dots, a_{1,n_1,j}$ such that

$$\boldsymbol{M}_j \not\models (\exists x_{2,1} \dots \exists x_{2,n_2}) \dots (Q_k x_{k,1} \dots Q_k x_{k,n_k}) \\ \rho_j(a_{1,1,j}, \dots, a_{1,n_1}, x_{2,1}, \dots, x_{2,n_2}, \dots, x_{k,1}, \dots, x_{k,n_k}).$$

Equivalently

$$\boldsymbol{M}_j \models (\forall x_{2,1} \dots x_{2,n_2}) \dots (Q'_k x_{k,1} \dots Q'_k x_{k,n_k}) \\ \neg\rho_j(a_{1,1,j}, \dots, a_{1,n_1,j}, x_{2,1}, \dots, x_{2,n_2}, \dots, x_{k,1}, \dots, x_{k,n_k}) \quad (14)$$

where again Q'_k denotes the opposite quantifier to Q_k.

Now let $a_{1,1}, \dots, a_{1,n_1} \in \boldsymbol{M}$ be the r_i-tuples formed from these violating tuples from each $\boldsymbol{M}_j$, which is to say that

$$a_{1,l}(j) = a_{1,l,j} \ (1 \le l \le n_1) \quad (15)$$

Now $\boldsymbol{M} \vDash \rho$, and so there exist elements $a_{2,1}, \dots, a_{2,n_2} \in \boldsymbol{M}$ such that

$$\boldsymbol{M} \vDash (\forall x_{3,1} \dots \forall x_{3,n_3}) \dots (Q_k x_{k,1} \dots Q_k x_{k,n_k}) \\ \rho(a_{1,1}, \dots, a_{1,n_1}, a_{2,1}, \dots, a_{2,n_2}, x_{3,1}, \dots, x_{3,n_3}, \dots, x_{k,1}, \dots, x_{k,n_k}) \quad (16)$$

We continue inductively in this way and assume that for some $t \ge 1$, for each $j \in \{1, \dots, r_i\}$ there exists elements

$$a_{1,1,j}, \dots, a_{1,n_1,j}, \dots, a_{2t-1,1,j}, \dots, a_{2t-1,n_{2t-1},j} \in \boldsymbol{M}_j$$

such that

$$\boldsymbol{M}_j \vDash (\forall x_{2t,1} \dots \forall x_{2t,n_{2t}}) \dots (Q_k x_{k,1} \dots Q_k x_{k,n_k}) \\ \neg\rho_j(a_{1,1,j}, \dots, a_{1,n_1,j}, \dots, a_{2t-1,1,j}, \dots, a_{2t-1,n_{2t-1},j}, \\ x_{2t,1}, \dots, x_{2t,n_{2t}}, \dots, x_{k,1}, \dots, x_{k,n_k}) \quad (17)$$

And with

$$a_{m,l}(j) = a_{m,l,j} \ (1 \le l \le n_m)\,(1 \le m \le 2t-1) \quad (18)$$

there exist elements $a_{2t,1}, \dots, a_{2t,n_{2t}} \in \boldsymbol{M}$ such that with

$$\boldsymbol{M} \vDash (\forall x_{2t+1,1} \dots \forall x_{2t+1,n_{2t+1}}) \dots (Q_k x_{k,1} \dots Q_{k,n_k}) \\ \rho(a_{1,1}, \dots, a_{1,n_1}, \dots, a_{2t,1}, \dots, a_{2t,n_{2t}}, x_{2t+1,1}, \dots \\ \dots, x_{2t+1,n_{2t+1}} \dots, x_{k,1}, \dots, x_{k,n_k}) \quad (19)$$

The base $t = 1$ case of (17), (18), and (19) is given by (14), (15) and (16) respectively. We now verify that we may increment each of the three parts of the inductive hypothesis, they being (17), (18), and (19), from t to $t + 1$ and thereby continue the induction.

We use (19) to project from $\boldsymbol{M}$ to each $\boldsymbol{M}_j$ by making substitutions in (17):

$$x_{2t,1} \mapsto a_{2t,1,j} = a_{2t,1}(j), \dots, x_{2t,n_{2t}} \mapsto a_{2t,n_2,j} = a_{2t,n_2}(j) \quad (20)$$

Then from (17) we have:

$$\begin{aligned}\boldsymbol{M}_j \vDash (\exists x_{2t+1,1} \dots \exists x_{2t+1,n_{2t+1}}) \dots (Q_k x_{k,1} \dots Q_k x_{k,n_k})\\ \neg\rho_j(a_{1,1,j}, \dots, a_{1,n_1,j}, \dots, a_{2t,1,j}, \dots, a_{2t,n_{2t},j},\\ x_{2t+1,1}, \dots, x_{2t+1,n_{2t+1}}, \dots, x_{k,1}, \dots, x_{k,n_k})\end{aligned} \tag{21}$$

Substituting witnesses $x_{2t+1,l,j} \mapsto a_{2t+1,l,j}$ $(1 \leq l \leq n_{2t+1})$ in (21) then increments (17) from t to $t+1$:

$$\begin{aligned}\boldsymbol{M}_j \vDash (\forall x_{2t+2,1} \dots \forall x_{2t+2,n_{2t+2}}) \dots (Q_k x_{k,1} \dots Q_k x_{k,n_k})\\ \neg\rho_j(a_{1,1,j}, \dots, a_{1,n_1,j}, \dots, a_{2t+1,1,j}, \dots, a_{2t+1,n_{2t+1},j},\\ x_{2t+2,1}, \dots, x_{2t+2,n_{2t+2}}, \dots, x_{k,1}, \dots, x_{k,n_k})\end{aligned} \tag{22}$$

Next we put $a_{2t+1,l}(j) = a_{2t+1,l,j}$ $(1 \leq l \leq n_{2t+1})$, which, together with (20), increments (18) from t to $t+1$. Finally, by (19) we may substitute in $\boldsymbol{M}$:

$$x_{2t+1,1} \mapsto a_{2t+1,1}, \dots, x_{2t+1,n_{2t+1}} \mapsto a_{2t+1,n_{2t+1}},$$

and call up witnesses:

$$x_{2t+2,1} \mapsto a_{2t+2,1}, \dots, x_{2t+2,n_{2t+2}} \mapsto a_{2t+2,n_{2t+2}}$$

such that

$$\begin{aligned}\boldsymbol{M} \vDash (\forall x_{2t+3,1} \dots \forall x_{2t+3,n_{2t+3}}) \dots (Q_k x_{k,1} \dots Q_{k,n_k})\\ \rho(a_{1,1}, \dots, a_{1,n_1}, \dots, a_{2t+2,1}, \dots, a_{2t+2,n_{2t+2}},\\ x_{2t+3,1}, \dots, x_{2t+3,n_{2t+3}}, \dots, x_{k,1}, \dots, x_{k,n_k})\end{aligned} \tag{23}$$

which increments (19) from t to $t+1$, and so the induction continues. This recursive procedure eventually yields a tuple

$$\bar{a} = (a_{1,1}, \dots, a_{1,n_1}, \dots, a_{k,1}, \dots, a_{k,n_k}) \tag{24}$$

such that $\boldsymbol{M} \vDash \rho(\bar{a})$ but for each j, $(1 \leq j \leq r_i)$, $\boldsymbol{M}_j \vDash \neg\rho_j(\bar{a}_j)$, where $\bar{a}_j$ represents the tuple obtained from (24) by projecting onto the jth coordinate:

$$\bar{a}_j = (a_{1,1,j}, \dots, a_{1,n_1,j}, \dots, a_{k,1,j}, \dots, a_{k,n_k,j}).$$

Now for all $i' = 1, \dots, k$, we have that $\gamma_{i'}(\bar{a})$ is true in $\boldsymbol{M}$ and also $\gamma_{i'}(\bar{a}_j)$ holds in each $\boldsymbol{M}_j$. Now $\boldsymbol{M}_j \vDash \neg\rho_j(\bar{a}_j)$ and for $i' \neq i$ the conjunct $\gamma_{i'}$ appears in ρ_j; as we have noted, $\boldsymbol{M}_j \vDash \gamma_{i'}(\bar{a}_j)$, and so it follows that $\alpha_{i,j}(\bar{a}_j)$ must be false in $\boldsymbol{M}_j$. But as $\gamma_i(\bar{a})$ is true in $\boldsymbol{M}$, there must exist $j \in \{1, \dots, r_i\}$ with $\alpha_{i,j}(\bar{a})$ true. But then, we obtain the contradiction that $\alpha_{i,j}(\bar{a}_j)$ is true in $\boldsymbol{M}_j$. Arrival at this contradiction completes the proof. □

Because the class operators H and P are related in composition by $\mathsf{PH} \leq \mathsf{HP}$, Theorem 3.1 can be re-expressed as stating that an elementary class K is definable by an equation system if and only if $K = \mathsf{HP}(K)$. If we wish to drop the assumption that K is an elementary class, we need more care.

Elementary classes are those closed under taking ultraproducts and elementary embeddings, however in the presence of H and P, we may ignore ultraproducts because they are particular cases of applications by HP. Thus Theorem 3.1 can be rephrased as "*a class K is the class of models of some equation systems if and only if it is closed under* E, H *and* P", where E denotes closure under taking elementary embeddings. In addition to the aforementioned containment $\mathsf{PH} \leq \mathsf{HP}$, it is possible to show that $\mathsf{HE} \leq \mathsf{EHP}$, which points toward the composite EHP as being a single closure operator equivalent to iterated closure under combinations of E, H and P. Unfortunately the authors are not aware of a useful containment between PE and EHP. We refer simply to $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-closed classes and even $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-classes, though $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}^*$-closed may be more technically correct. An interesting consequence of Theorem 3.1 is that all equationally defined classes arise as reducts of varieties. This is of course well-known for inverse semigroups and groups (as semigroups), but is not otherwise immediately obvious for other $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-classes.

The class of reducts of a variety is always closed under ultraproducts and direct products, but in general need not be closed under taking homomorphic images, nor subalgebras, nor even elementary embeddings. There are plentiful easy examples demonstrating the failure of the first two of these closure properties. For the case of elementary embeddings, we observe that real vector spaces form a variety (with vector addition as binary and $\mathbb{R}$-many unary operations for scalar multiplication). The class of reducts to the empty signature has no countably infinite members, and hence is not an elementary class. When the class of reducts of members of a variety is closed under H and E (as they are for groups and inverse semigroups), then Theorem 3.1 shows that the class is definable by the equation systems. We now show a converse to this statement.

Theorem 3.2. *Let $\mathscr{L}$ be a signature and $\mathcal{K}$ an $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-closed class of $\mathscr{L}$-structures. Then $\mathcal{K}$ is the class of reducts of a variety $\mathcal{V}$ in a signature extending $\mathscr{L}$. If $\mathcal{K}$ is finitely axiomatisable in first order logic, then $\mathcal{V}$ can be chosen to be finitely based, and of finite signature.*

Proof. As $\mathcal{K}$ is closed under taking homomorphic images, direct products and elementary embeddings, it can be axiomatised by a family of equation systems by Theorem 3.1. If $\mathcal{K}$ is finitely axiomatisable in first order logic, then the Completeness Theorem for first order logic ensures that it can equivalently be axiomatised by a finite family of equation systems. We now explain how to replace each equation system in the family of equation systems by an identity in some extended signature. A finite number of new operations is added for each equation system, so that if the family of equation systems defining $\mathcal{K}$ is finite, then so also will the resulting variety (after all equation systems are replaced) be finite.

Consider a signature $\mathscr{L}$ and a sentence $(\forall\vec{x})\,(\exists y)\,\phi(\vec{x},y)$ in $\mathscr{L}$, where $\vec{x}$ abbreviates $x_1,\dots,x_n$ (for some n, possibly 0) and where $\phi(\vec{x},y)$ may contain other quantified variables that are not displayed. Let f be a new n-ary operation symbol. It easily verified that the models of $(\forall\vec{x})\,(\exists y)\,\phi(\vec{x},y)$ are precisely the $\mathscr{L}$-reducts of models of $(\forall\vec{x})\,\phi(\vec{x},f(x_1,\dots,x_n))$. Indeed, if $\mathbf{M}\models(\forall\vec{x})\,(\exists y)\,\phi(\vec{x},y)$ then we may expand the signature $\mathscr{L}$ of $\mathbf{M}$ to include f by defining f at each tuple $\vec{a}\in M^n$ to be any witness x to $(\exists x)\,\phi(\vec{a},x)$. This expansion of $\mathbf{M}$ is not necessarily unique, but all such expansions are models of $(\forall\vec{x})\,\phi(\vec{x},f(x_1,\dots,x_n))$. Conversely any model $\mathbf{N}$ of $(\forall\vec{x})\,\phi(\vec{x},f(x_1,\dots,x_n))$ has its $\mathscr{L}$-reduct a model of $(\forall\vec{x})\,(\exists y)\,\phi(\vec{x},y)$, as the value of f at any tuple $\vec{a}\in N^n$ provides the witness x required in $(\exists x)\,\phi(\vec{a},x)$. This process is known as *Skolemisation*; see [2, §3.3] or [8, §3.1] for example. An application Skolemisation to an equation system produces an equation system with one fewer existential quantifiers. Repeated applications eventually leads to an equation system without any existential quantifiers. Such an equation system is a finite set of identities. □

As a first example, Skolemising the defining equation $(\exists x)(\forall a)\,xa=ax=a$ for monoids (as semigroups), we introduce a nullary operation e to replace x to obtain $(\forall a)\,ea=ae=a$, the familiar definition as a semigroup with constant. As a second example, we consider the result of applying Skolemisation to the definition (1) for the class of groups as semigroups. The given sentence is $(\forall a\forall b)(\exists x\exists y)\,ax=b\wedge ya=b$. Skolemising once (using the symbol $\backslash$ for the introduced Skolem function) we obtain $(\forall a\forall b)(\exists y)\,a(a\backslash b)=b\wedge ya=b$, and then a second time (using $/$) we obtain $(\forall a\forall b)\,a(a\backslash b)=b\ \ \&\ \ (b/a)a=b$ (note that b/a might have more consistently been written as a/b, however it is immediately clear that the required value is the element ba^{-1}). Thus groups (as semigroups) are the class of reducts of the variety with two additional binary operations $\backslash,/$ defined by the identities $a(a\backslash b)=b$ and $(b/a)a=b$ in addition to associativity of the semigroup multiplication.

4. Equational bases for e-varieties of semigroups

The theory of semigroup e-varieties was devised by T.E. Hall [6] and others in order create an interface between the theory of regular semigroups and universal algebra. The theory endows $\{\mathsf{H},\mathsf{P}\}$-closed classes of regular semigroups with the structure of a varietal type through introduction of unary operations corresponding to choices of inverses for elements of the regular semigroups involved. From Section 3 we know that we may, at least in principle, represent such classes by equational bases using only the associative binary operation with which the semigroup is naturally endowed. We shall henceforth refer to these as $\{\mathsf{E},\mathsf{H},\mathsf{P}\}$-bases, indicating that they determine a defining set of equations for a class that is closed under E, H and P.

Three fundamental classes that form e-varieties are the classes $\mathcal{I}$ of all *inverse semigroups*, $\mathcal{O}$ of *orthodox semigroups*, and $\mathcal{ES}$ of so-called *E-solid* *semigroups*. In this section we obtain finite equational bases for these benchmark e-varieties. The bases each involve choosing inverses for two arbitrary members $a, b \in S$ and adjoining a second set of equations that ensure that all products of the associated idempotents of a certain length (length 3 for $\mathcal{ES}$ and for $\mathcal{O}$, and 2 for $\mathcal{I}$) have a particular property associated with the class (are group members for $\mathcal{ES}$, are idempotents for $\mathcal{O}$, and commute with one another in the case of $\mathcal{I}$).

We begin with the equational basis problem for the class $\mathcal{ES}$ of all *E-solid semigroups*, which are those regular semigroups S that satisfy the *solidity condition* that for idempotents $e, f, g \in E(S)$

$$e \mathscr{L} f \mathscr{R} g \to \exists h \in E(S) : e \mathscr{R} h \mathscr{L} f \tag{25}$$

Note that $\mathcal{O} \subseteq \mathcal{ES}$ and $\mathcal{ES} \cap \mathcal{IG} = \mathcal{CR}$. We make use of the fact, taken from [5], that a semigroup S is E-solid if and only if S is a regular semigroup for which the idempotent-generated subsemigroup $\langle E(S)\rangle$ is a union of groups.

Theorem 4.1. *The class $\mathcal{ES}$ of all E-solid semigroups S is the $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-class with two-part $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-basis:*

$$(\forall a, b)\, (\exists x, y) : (x \in V(a),\, y \in V(b)) \tag{26}$$

In addition to (26) *the basis includes:*

$$(\forall p = g_1g_2g_3 :\, g_j \in F = \{ax, xa, by, yb\}\, (\exists z \in S)(z \in V(p),\, zp = pz) \tag{27}$$

Proof. We have in mind that it is implicit that equation system (27) really abbreviates a family of 4^3 equation systems, each incorporating (26) within:

$$(\forall a \forall b)\, (\exists x \exists y \exists z) :\, x \in V(a) \wedge y \in V(b) \wedge z \in V(p) \wedge zp = pz$$

where p denotes the product $g_1g_2g_3$ for $g_1, g_2, g_3 \in \{ax, xa, by, yb\}$.

If S is E-solid then S is regular and so satisfies the equation system (26). Moreover, since $\langle E(S)\rangle$ is a union of groups, it follows that for *any* product of idempotents $p = \Pi_{i=i}^k f_i$ ($f_i \in E(S)$), we may take some $z_p \in E^{k+1} \cap V(p)$ to be the group inverse of p in H_p and so satisfy (27). Note that since each of the g_j in (27) is idempotent, (27) implies that the product of three or fewer members of F has a group inverse.

Conversely, suppose that S satisfies the equation systems (26) and (27); by (26) S is regular. Take $a, b \in E(S)$ and take $x, y \in S$ so that (26) and (27) are satisfied. Then $xa \mathscr{L} a \mathscr{R} ax$ and $p = xa \cdot ax = xax = x$, and so by (27) H_x is a group. It follows that $ax \cdot xa \in H_a \cap E(S)$ and so $ax \cdot xa = a$, and similarly $by \cdot yb = b$. Now suppose that $a, f, b \in E(S)$ are such that $a \mathscr{L} f \mathscr{R} b$. We then have $xa \mathscr{L} f \mathscr{R} b$ so that $xa \cdot b = xa \cdot by \cdot yb \in E(S)$. Therefore H_{xab} is a group. We thus have the following $\mathscr{D}$-class structure for $D = D_a$ in which the $\mathscr{H}$-classes of each given entry is a group, for we now observe by (27) that $yb \cdot xa$ is a group element located as shown.

x	xa		xab	
ax	a	...		
	f		b	by
	$ybxa$		yb	y

Moreover $a \mathscr{L} yb \cdot xa \mathscr{R} yb$, which lets us conclude, again by (27), that

$$ab \mathscr{H} a \cdot yb = ax \cdot xa \cdot yb \in E(S) \text{ and so } H_{ab} \text{ is a group.}$$

Therefore the idempotent $h \in H_{ab}$ fulfills the requirements of the E-solid condition (25) and so S is E-solid. □

Theorem 4.2. *The class $\mathcal{O}$ of orthodox semigroups S is the $\{\mathsf{E},\mathsf{H},\mathsf{P}\}$-class with two-part $\{\mathsf{E},\mathsf{H},\mathsf{P}\}$-basis:*

$$(\forall a, b)\,(\exists x, y) : (x \in V(a),\ y \in V(b)) \tag{28}$$

In addition to (28) *the basis includes:*

$$(\forall p = g_1g_2g_3 :\ g_j \in F = \{ax, xa, by, yb\})\,(p \in E(S)) \tag{29}$$

Proof. Clearly, since the members of F are idempotents, the equations (28) and (29) have solutions in an orthodox semigroup S. Conversely suppose that S satisfies (28) and (29). Let $a \in E(S)$ and take any $b \in V(a)$. We shall show that $b \in E(S)$. Since in any regular semigroup $V(E) = E^2$, it follows from this that $E = E^2$, which is to say that S is orthodox. With x, y satisfying the equations of (28) and (29) we have $x = xax = xa \cdot ax \in E$. Then $ax \cdot xa \in E(S) \cap H_a$ so that $ax \cdot xa = a$. We have the following $\mathscr{D}$-class diagram, the remaining entries of which are explained below.

x	xa			$xaby$
ax	a	...	ab	aby
	ba		b	by
	yba		yb	y

Since $ba \in E(S)$ we have $xa{\cdot}by \in E(S)$ placed as shown; hence $by{\cdot}xa = ba$. Similarly since $ab \in E(S)$ we have $yb{\cdot}a = yb{\cdot}ax{\cdot}xa \in E(S)$ as shown. Hence $a \cdot yb = ax \cdot xa \cdot yb \in E(S)$ so that $a \cdot yb = ab$. Putting these factorizations together gives:

$$b = b \cdot ab = b \cdot ayb = ba \cdot yb = by \cdot xa \cdot yb \in E(S),$$

which completes the proof. □

Theorem 4.3. *The class $\mathcal{I}$ of inverse semigroups S is the $\{\mathsf{E},\mathsf{H},\mathsf{P}\}$-class with two-part $\{\mathsf{E},\mathsf{H},\mathsf{P}\}$-basis:*

$$(\forall a, b)\,(\exists x, y) : (x \in V(a),\ y \in V(b)) \tag{30}$$

In addition to (30) *the basis includes:*

$$(\forall p = g_1g_2 :\ g_j \in F = \{ax, xa, by, yb\})\,(g_1g_2 = g_2g_1) \tag{31}$$

Proof. Clearly any inverse semigroup satisfies (30) and (31) so only the converse is in question. By (30) S is regular so let us take $a, b \in E(S)$. (We require only three of the six equations specified by (31)). From (31) we obtain $x = xax = xa{\cdot}ax = ax{\cdot}xa$, whence $x = ax = xa$ and so $x = axa = a$. In the same way $b = y$. Putting $g_1 = ax$, $g_2 = yb$ we then obtain from (31) that $ab = ax \cdot yb = yb \cdot ax = ba$. Since S is regular and every pair of idempotents of S commute we have proved that S an inverse semigroup. □

Example 4.4. The products of length 3 in Theorems 4.2 and 4.1 cannot be replaced by products of length 2 as in Theorem 4.3.

Proof. We demonstrate this by finding an assignment of an inverse to each member of a regular semigroup S in such a way that S satisfies the equation system (28) and the length two version of equation system (29):

$$(\forall p = g_1 g_2 : g_j \in F = \{ax, xa, by, yb\})\,(p \in E(S)) \tag{32}$$

However, S is not E-solid (and so also not orthodox).

Let S be the 0-rectangular band with non-zero $\mathscr{D}$-class D defined by the first 'eggbox' of the following three diagrams, where an asterisk denotes an idempotent:

*	*		*
*	*		
	*		*
		*	

	1		
1			
			1
		1	

2	2		2
2	2		
	2		2
		2	

We see that S is regular but not E-solid since, for example, the entry at position $(2, 4)$ is not idempotent despite the presence of idempotents at positions $(2, 2)$, $(3, 2)$, and $(3, 4)$. We shall write (i, j) $(1 \leq i, j \leq 4)$ to denote the element of D in that corresponding position in the diagram. Let $a \in D$. We shall choose $x \in V(a)$ writing this selection in the form $a \to x$. We assign $(1, 1) \leftrightarrow (2, 2)$, $(1, 2) \to (1, 2), (2, 1) \to (2, 1)$, $(3, 3) \leftrightarrow (4, 4)$, $(3, 4) \to (3, 4)$, $(4, 3) \to (4, 3)$; $(1, 3) \leftrightarrow (4, 2)$, $(1, 4) \leftrightarrow (3, 2)$; $(2, 3) \leftrightarrow (4, 1)$, $(2, 4) \leftrightarrow (3, 1)$.

In the second diagram the idempotent products of the form ax or xa (under the previous assignment) are indicated by the numeral 1, which are all starred in the first diagram.

In the third diagram the non-zero products of pairs of idempotents of the form ax or yb $(a, b \in S)$ are indicated by the numeral 2. Since each instance of the numeral 2 lies in a starred square in the first diagram, it follows that, with the given assignment of inverses, S satisfies the equations of (28) and (32) but S is not E-solid. Therefore the equational basis given by (28) and (32) does not imply E-solidity. □

5. Universally Satisfied Equations

Given an $\{\mathsf{E}, \mathsf{H}, \mathsf{P}\}$-class $\mathcal{C}$ there are two natural tasks arising. The first is the determination of an equational basis for $\mathcal{C}$, which was the subject of Section 4. In this section we examine the other side of the coin, which is the question of finding all equations satisfied by $\mathcal{C}$. Here we shall solve the latter problem for one class only, that being the class generated by $P = (\mathbb{Z}^+, +)$ and for equations of the type $\forall \dots \exists$. As a corollary we obtain a description of the class of equations without parameters solvable in every semigroup.

We shall denote a typical semigroup equation as $e : p = q$ where $p, q \in F_{A\cup X}$, the free semigroup on $A \cup X$, where A and X are disjoint countably infinite sets. Elements of A will follow instances of the $\forall$ quantifier while those drawn from X will follow the $\exists$ symbol. We shall denote the number of instances of the letter $y \in A \cup X$ in a word $w \in F_{A\cup X}$ by $|w|_y$, with the length of w simply denoted by $|w|$. Define the *content* of w as the set

$$c(w) = \{y \in A \cup X : |w|_y \geq 1\}.$$

Definition 5.1. An equation $e : p = q$ $(p, q \in F_{A\cup X})$ is *semigroup universal* if e is satisfied by every semigroup S.

In this section we adopt the abbreviation that an equation is *universal* if it is a semigroup-universal equation. This is is not to be confused with "universally quantified equation", which in this article is referred to as an "identity".

We will make use of elementary results on subsemigroups of P. Our source here is Chapter 2 Section 4 of the book by Grillet [4] who therein gives the original sources of these and other related facts on numerical semigroups.

Theorem 5.2 (Proposition II.4.1 and Corollary 4.2 of [4])**.**

(i) *Let S be a subsemigroup of $P = (\mathbb{Z}^+, +)$. Then there exists a unique integer $d \geq 1$ such that S consists of multiples of d and S contains all sufficiently large multiples of d.*

(ii) *A subsemigroup S of $(\mathbb{Z}, +)$ either contains only non-negative integers, or only non-positive integers, or is a subgroup of $(\mathbb{Z}, +)$. In the latter case, $S = d\mathbb{Z}$ for some $d \geq 0$.*

Corollary 5.3. *Let $m_1, \dots, m_n \in \mathbb{Z}$ with $gcd(m_1, \dots, m_n) = d$. Define*

$$S = S(m_1, \dots, m_n) = \{t_1m_1 + \cdots + t_mm_n, t_i \geq 1, (1 \leq i \leq n)\}.$$

Then S is a subsemigroup of $(\mathbb{Z}, +)$. Moreover if all the integers m_i are positive then S is a subsemigroup of $P = (\mathbb{Z}^+, +)$ and further d is the unique integer such that $S \subseteq d\mathbb{Z}^+$ and there exists $k \in \mathbb{Z}^+$ such that $\{da : a \geq k\} \subseteq S$.

Proof. Clearly S is a subsemigroup of P. In the case where all the $m_i \geq 1$, by Theorem 5.2(i), there exists a unique positive integer d_1 that has both the properties that $S \subseteq d_1\mathbb{Z}^+$ and there exists a positive integer k such that $d_1a \in S$ for all $a \geq k$. On the other hand $d|s$ for all $s \in S$ so that $S \subseteq d\mathbb{Z}^+$. It follows that $d|kd_1$ and $d|(k + 1)d_1$, whence d is a factor of their difference, and so $d|d_1$. On the other hand, for any $1 \leq i \leq n$ we may write $m_1 + \cdots + m_n = pd_1$ and $m_1 + \cdots + 2m_i + \cdots + m_n = qd_1$ for some $p < q \in \mathbb{Z}^+$. Then we have $m_i = (q - p)d_1$ and so $d_1|m_i$ for all $1 \leq i \leq n$. Therefore $d_1|d = \gcd\{m_1, \dots, m_n\}$. We conclude that $d_1 = d$. □

Definition 5.4. Let e be of the form $(\forall a_1, \ldots, a_m)\,(\exists x_1, \ldots, x_n) : p = q$. Let $r_i = |p|_{x_i}$, $s_i = |q|_{x_i}$, $p_j = |p|_{a_j}, q_j = |q|_{a_j}$. We shall write $r_i - s_i$ as m_i and $q_i - p_i$ as n_j; let d stand for $\gcd\{m_1, \ldots, m_n\}$ and d' for $\gcd\{n_1, \ldots, n_m\}$.

Theorem 5.5. *Let $e : p = q$ be an equation written in the notation of Definition 5.4. Then $p = q$ holds in P if and only if for any given positive integers a_i $(1 \leq i \leq m)$ there exist positive integers t_i $(1 \leq i \leq n)$ (depending on the a_i) such that:*

$$\sum_{i=1}^{n} t_i m_i = \sum_{i=i}^{m} a_i n_i,$$

which is equivalent to the statement that $S(n_1, \ldots, n_m) \subseteq S(m_1, \ldots, m_n)$.

Proof. In P, after a substitution $x_i \to t_i$, our equation $p = q$ takes on the form:

$$(p_1a_1+\cdots+p_ma_m)+(r_1t_1+\cdots+r_nt_n) = (q_1a_1+\cdots+q_ma_m)+(s_1t_1+\cdots+s_nt_n)$$

$$\Leftrightarrow \sum_{i=1}^{m}(p_i - q_i)a_i + \sum_{i=1}^{n}(r_i - s_i)t_i = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} t_i m_i = \sum_{i=i}^{m} a_i n_i$$

The integers $r_i - s_i = m_i$ and $q_i - p_i = n_i$, which are fixed and may be negative, are determined by the equation e. It follows that $p = q$ will be solvable in P if and only if every linear combination of the n_i in positive integers is a linear combination in positive integers of the m_i, which is equivalent to the statement of the theorem. □

Theorem 5.6. *Let $e : p = q$ be an equation written in the notation of Definition 5.4. Let $d = gcd(m_1, \ldots, m_n)$ and $d' = gcd(n_1, \ldots, n_m)$. Suppose for some $x, y \in c(pq)$, $|p|_x < |q|_x$ and $|p|_y > |q|_y$. Then $e : p = q$ is solvable in P if and only if $d|d'$.*

Remark 5.7. If $n_1 = \cdots = n_m = 0$, we put $d' = 0$ and then $d|d'$ for all $d \in \mathbb{Z}^+$. We note that in this case, $S(n_1, \ldots, n_m) = \{0\} = d'\mathbb{Z}$.

Proof. By hypothesis, some of the integers m_i are positive and some are negative, from which it follows from Theorem 5.2(ii) that $S_1 := S(m_1, \ldots, m_n) = d_1\mathbb{Z}$ for some $d_1 \geq 1$. Since $d_1 \in S_1$ it follows that $d|d_1$. On the other hand, by the argument in the proof of Corollary 5.3, $d_1|d$ also and therefore $d_1 = d$ and so $S_1 = d\mathbb{Z}$.

For $d' \neq 0$, $d'|n_i$ for all $1 \leq i \leq m$; in any event it follows that $S(n_1, \ldots, n_m) \subseteq d'\mathbb{Z}$. Now suppose that $d|d'$ so that $d' = dy$ say. Then

$$S(n_1, \ldots, n_m) \subseteq d'\mathbb{Z} = dy\mathbb{Z} \subseteq d\mathbb{Z} = S(m_1, \ldots, m_n),$$

and so by Theorem 5.5, e is solvable in P. Conversely suppose that e may be solved in P. By Corollary 5.3, $S(n_1, \ldots, n_m)$ contains a set of the form $\{ad' : a \geq k\}$ for some fixed positive integer k. Take p to be a prime with $p \geq k+d$.

Then $pd' \in S(n_1, \ldots, n_m)$. By Theorem 5.5, $pd' \in S(m_1, \ldots, m_n) = d\mathbb{Z}$. We then have $d|pd'$. However d and p are relatively prime (as $p > d$) and so $d|d'$, thus completing the proof. □

Corollary 5.8. *Let $e : p = q$ be an equation without parameters. Let $c(pq) = \{x_1, \ldots, x_n\}$. Then e is universal if and only if either:*

(i) *$|p|_{x_i} = |q|_{x_i}$ for all $i = 1, 2, \ldots, n$ or*
(ii) *for some $x, y \in c(pq)$, $|p|_x < |q|_x$ and $|p|_y > |q|_y$.*

Proof. If (i) applies to e then for any semigroup S take $s \in S$ and substitute $x_i \to s$ $(1 \leq i \leq n)$ in e; this yields $s^{|p|} = s^{|q|}$, which is true in S as $|p| = |q|$ and therefore e is universal. Next suppose that (ii) holds. Then e is equivalent in P to the equation $e' : pa = qa$ where $a \in A$ is a parameter. Since $d' = 0$, the condition $d|d'$ holds whence it follows from Theorem 5.6 that e' is solvable in P and therefore e is likewise. Now taking any $s \in S$ we have $\langle s \rangle$ is a homomorphic image of P and hence e is solvable in $\langle s \rangle$. Since e has no parameters, any solution of e in $\langle s \rangle$ is also a solution of e in S and therefore e is universal.

Conversely suppose that neither conditions (i) nor (ii) hold for e. Without loss we may assume that $|p| \leq |q|$. Suppose that for some i $(1 \leq i \leq n)$ $|p|_{x_i} > |q|_{x_i}$. Then, since $|p| \leq |q|$ it would follow that for some other subscript j we would find that $|p|_{x_j} < |q|_{x_j}$, contradicting the assumption that condition (ii) does not hold. Therefore $|p|_{x_i} \leq |q|_{x_i}$ for all $1 \leq i \leq n$. Moreover, since condition (i) does not hold either, for at least one subscript i the previous inequality is strict. Any substitution $x_i \to t_i$ $(t_i \in \mathbb{Z}^+)$ in P therefore yields respective positive integers p' and q' say with $p' < q'$. In particular $p' \neq q'$ and so that e cannot be satisfed in P and therefore e is not universal. □

Example 5.9. Let us consider

$$e : \ x_1^9 x_2^{23} a_1^2 a_2^{13} a_3 = x_1^{30} x_2^8 a_1^{11} a_2^7 a_3^{10}.$$

Here $m_1 = 9 - 30 = -21$, $m_2 = 23 - 8 = 15$, and $d = \gcd(m_1, m_2) = 3$; $n_1 = 11 - 2 = 9$, $n_2 = 7 - 13 = -6$, $n_3 = 10 - 1 = 9$, and $d' = \gcd(9, -6, 9) = 3$. Then $d = d'$ so $d|d'$ and by Theorem 5.6, e is solvable. Particular selections for a_1, a_2, and a_3 lead to solvable linear diophantine equations.

Example 5.10. The equation

$$e : \ x^{13} y^{24} a^2 b^5 = x^{10} y^{16} a^{13} b^{19}$$

is an instance in which each of the variables (x and y) occurs more often on the left than they do on the right so that Theorem 5.6 does not apply. Nonetheless Theorem 5.5 shows us that e is solvable in P. We have $m_1 = 13 - 10 = 3, m_2 = 24 - 16 = 8$, $n_1 = 13 - 2 = 11$, $n_2 = 19 - 5 = 14$. We note that $34 = (6 \times 3) + (2 \times 8)$, $35 = (1 \times 3) + (4 \times 8)$, $36 = (4 \times 3) + (3 \times 8)$. It follows that

$$\{k \in \mathbb{Z} : k \geq 34\} \subseteq \{3t_1 + 8t_2 : t_1, t_2 \geq 1\} = T.$$

Now $\{11t_1 + 14t_2 : t_1, t_2 \geq 1\} \cap \{k \in \mathbb{Z} : k \leq 33\} = \{11 + 14 = 25\}$ and $(3 \times 3) + (2 \times 8) = 25 \in T$ also. Hence $\{n_1t_1 + n_2t_2 : t_1, t_2 \geq 1\} \subseteq \{m_1t_1 + m_2t_2 : t_1, t_2 \geq 1\}$ and so e is solvable in P. For a particular instance we put $a = 2$ and $b = 3$ giving the diophantine equation:

$$3t_1 + 8t_2 = (2 \times 11) + (3 \times 14) = 64;$$

hence $2t_2 \equiv 1 \pmod 3$ so $t_2 = 2 + 3t$, giving $3t_1 + 8(2 + 3t) = 64$, whence $3t_1 = 48 - 24t$ and so $t_1 = 16 - 8t$. Since $t_1, t_2 \geq 1$ there are two solutions given by $t = 0, 1$ which are respectively $t_1 = 16$, $t_2 = 2$ and $t_1 = 8$, $t_2 = 5$. Substituting $x = 16$ and $y = 2$ yields a common value of 275 for both sides of the equation e, while putting $x = 8$ and $y = 5$ gives 243 on each side.

References

[1] Burris S. and H.P. Sankappanavar, *A Course in Universal Algebra*, Springer-Verlag, 1981.
[2] Chang C.C. and H.J. Keisler, *Model Theory*, Studies in Logic and the Foundations of Mathematics Vol. 73, North Holland, 3rd ed., 1990.
[3] Clifford A.H. and G.B. Preston, *The algebraic theory of semigroups*, Vol. 1, Amer. Math. Soc. (1961), Vol. II, (1967).
[4] Grillet, P.A., *Commutative semigroups*, Springer Science + Business Media (2001).
[5] Hall, T.E., *On regular semigroups*, J. of Algebra, Vol. 24(1), (1973), 1–24.
[6] Hall, T.E., *Identities for existence varieties of regular semigroups,* Bull. Austral. Math. Soc.40 (1989), 59–77.
[7] Higgins, P.M., *Techniques of semigroup theory*, OUP (1992).
[8] Hodges, W., *A Shorter Model Theory*, Cambridge University Press, 2002.
[9] Howie, J.M., *Fundamentals of Semigroup Theory*, London Math. Soc. monographs, Clarendon Press, OUP (1995).
[10] Lallement, G., *Semigroups and Combinatorial Applications*, Wiley and Sons (1979).
[11] Lyndon, R., *Properties preserved under homomorphism*, Pacific J. Math. 9 (1959), 143–54.
[12] Nagy, A., *Special Classes of Semigroups*, Springer (2001).
[13] Rosen, E., *Some aspects of model theory and finite structures*, Bull. Symb. Logic 8 (2002), 380–403.

Department of Mathematical Sciences, University of Essex, UK
E-mail address: peteh@essex.ac.uk

Department of Mathematics and Statistics, La Trobe University, VIC 3086, Australia
E-mail address: m.g.jackson@latrobe.edu.au