MT5823 SEMIGROUP THEORY

J. D. MITCHELL

Contents

List of Figures

Course Information

- **Lecturer:** James Mitchell, room 308.
- **Class hour:** 9am to 10am Mondays (odd), Wednesdays and Fridays.
- **Recommended texts:** Fundamentals of Semigroup Theory, J. M. Howie, Clarendon Press, Oxford, 1995.
- **Email:** jdm3@st-and.ac.uk
- **Course aims:**
  - the definition and basic properties of semigroups;
  - typical algebraic properties, such as subsemigroups, binary relations, congruences and homomorphisms, idempotents, regularity, inverses, Green's equivalences and Green's structure theory;
  - different ways of defining semigroups, such as by presentation or as transformations;
  - semigroup constructions, i.e. ways of building new semigroups from known ones;
  - the internal structure of a semigroup, most importantly Green's relations;
  - special kinds of semigroups. For example, zero semigroups, rectangular bands, completely simple semigroup, inverse and Clifford;
  - some important theorems in semigroup theory.
- **Assessment:** One 2 hour exam - 75%, project - 25%.

1. Definition and first examples

Definition 1.1. (**Semigroup**). A ***semigroup*** is a set S with a binary operation (usually denoted as multiplication) satisfying:

- **Closure:** $xy \in S$ for all $x, y \in S$;
- **Associativity:** $(xy)z = x(yz)$ for all $x, y, z \in S$.

An immediate consequence of associativity is that the arrangement of brackets in a product of elements of S does not matter. For example,

$$x((yz)(tu)) = ((xy)z)(tu).$$

Consequently, we may omit the brackets altogether and simply write $xyztu$. We can also use the power notation

$$x^n = \underbrace{xx \dots x}_{n},$$

but only for positive integers n.

Example 1.2. (Groups). Every group is a semigroup.

Example 1.3. (Semigroups of numbers). Any of the following sets of numbers: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_n$, is a semigroup both with respect to addition and also with respect to multiplication.

Example 1.4. (The empty and trivial semigroups). The empty set $\varnothing$ is a semigroup under the operation $\varnothing$; called the ***empty semigroup***. It is the unique semigroup where the set and the operation are the same!

The one element semigroup $S = \{e\}$ with $ee = e$ is called the ***trivial semigroup***.

Example 1.5. (The free semigroup). Let A be a non-empty set (called an ***alphabet***). A ***word*** over A is a finite sequence $w = (a_1, a_2, \dots, a_m)$ with $m > 0$ of elements of A. It is usual to simply write $w = a_1 a_2 \dots a_m$. The number m is called the ***length*** of w; denoted $|w|$.

The set of all non-empty words over the alphabet A with operation concatenation, that is,

$$(a_1 a_2 \dots a_m)(b_1 b_2 \dots b_n) = a_1 a_2 \dots a_m b_1 b_2 \dots b_n,$$

is called the ***free semigroup*** over A and is denoted by A^+.

For example, if $A = \{a, b\}$, then $A^+ = \{a, b, a^2, b^2, ab, ba, a^3, b^3, \dots\}$. The product (concatenation) of words ab and ba is just $abba = ab^2a$.

Those of you familiar with other algebraic objects (groups, rings, etc.) will know that the free semigroup should have certain other properties to merit the name 'free'. We will return to this later in Theorem 3.8.

Definition 1.6. (Left and right zero). If S is a semigroup and $x \in S$ is such that $xy = x$ for all $y \in S$, then x is called a ***left zero***. ***Right zeros*** are defined analogously, and an element of S that is both a left and a right zero is called a ***zero*** or a ***2-sided zero***.

Example 1.7. (Left zero semigroups). Let S be any set (finite or infinite). Then define multiplication on S such that $xy = x$ for all $x, y \in S$.

If $x, y, z \in S$, then

$$(xy)z = xz = x = xy = x(yz)$$

and so S is a semigroup, called a ***left zero semigroup*** or a ***semigroup of left zeros***.

If the multiplication is defined by $xy = y$, then we obtain a ***right zero semigroup***.

If a semigroup S contains a left zero x and a right zero y, then $x = y$, and x is the zero element of S.

Example 1.8. (Zero semigroups). Let S be any set (again finite or infinite). Then define multiplication on $S \cup \{0\}$ (assuming without loss of generality that $0 \notin S$) by

$$xy = 0$$

for all $x, y \in S \cup \{0\}$. Then

$$(xy)z = 0z = 0 = x0 = x(yz)$$

and so $S \cup \{0\}$ is a semigroup, called a ***zero semigroup***.

Example 1.9. (Rectangular bands). Let I and Λ be any sets. The operation

$$(i,\lambda)(j,\mu) = (i,\mu)$$

on $I \times \Lambda$ is associative because

$$(i,\lambda)[(j,\mu)(k,\nu)] = (i,\lambda)(j,\nu) = (i,\nu) = (i,\mu)(k,\nu) = [(i,\lambda)(j,\mu)](k,\nu).$$

Definition 1.10. (Commutative semigroup). A semigroup S is ***commutative*** if $xy = yx$ for all $x, y \in S$.

Example 1.11. (A non-commutative semigroup). Let X be any non-empty set and let S_X denote the ***symmetric group*** on X (the group of all permutations of X under composition of functions).

Definition 1.12. (Left and right identity). If S is a semigroup and $e \in S$ is such that $es = s$ for all $s \in S$, then e is called a ***left identity***. ***Right identity*** are defined analogously, and an element of S that is both a left and a right identity is called an ***identity*** or a ***2-sided identity***.

Definition 1.13. (Monoid). A semigroup with an identity is called a ***monoid***.

Example 1.14. (The full transformation monoid). Let X be a non-empty set, and let T_X be the set of all functions from X into X. The operation on T_X is again composition of functions. Since composition of functions is associative, T_X is a semigroup. It is called the ***full transformation semigroup*** on X.

If $X = \{1, 2, \ldots, n\}$ (a finite set) we write T_n instead of $T_{\{1,2,\ldots,n\}}$.

We write functions to the right of their arguments (i.e. xf instead of $f(x)$). An immediate consequence of this convention is that the rule for composition of functions becomes

$$x(f \circ g) = (xf)g.$$

That is, functions multiply from left to right. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 4 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 1 & 4 & 4 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 4 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 1 & 1 & 5 \end{pmatrix}.$$

It follows that composition of functions is not commutative. Hence T_X, at least when $|X| \geq 5$.

The identity of T_X is the identity function id_X defined by $(x)\mathrm{id}_X = x$ for all $x \in X$.

Example 1.15. (Partial transformation monoid). A partial function is a function from a subset of X into X, i.e. it need not be defined on the whole of X. The composition is defined in the same way as for full functions:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & - & 4 & - \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & 5 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & - & - & 1 & - \end{pmatrix}.$$

The set of all partial functions on X under composition; denoted by P_X and again P_n when $X = \{1, \ldots, n\}$.

The function

$$e = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

is the identity for both T_n and P_n.

Lemma 1.16 (**Identities are unique**). *Let S be a semigroup with an identity $e \in S$. Then e is the unique element in S with the property that $es = se = s$ for all $s \in S$.*

Proof. Suppose that $f \in S$ is an identity. Then $e = ef$ since f is an identity and $ef = f$ since e is an identity. Hence $e = f$. □

Example 1.17. (Free monoid). Let A be any set, and let A^+ be the free semigroup over A. Form a new semigroup by taking the set $A^+ \cup \{\epsilon\}$ where ϵ denotes the empty word (of length 0). This semigroup is a monoid with identity ϵ called the ***free monoid*** and denoted A^*.

Definition 1.18. (Adjoining an identity or zero). Let S be a semigroup and let $1 \notin S$. Then denote by S^1 the set $S \cup \{1\}$ with extended multiplication

$$1s = s1 = s$$

for all $s \in S \cup \{1\}$. [1] The monoid S^1 is called a ***semigroup with adjoined identity***.

If $0 \notin S$, then we denote by S^0 the set $S \cup \{0\}$ with extended multiplication

$$0s = s0 = 0$$

for all $s \in S \cup \{0\}$. The semigroup with zero S^0 is called a ***semigroup with adjoined zero***.

For example, adjoining an identity to A^+ we obtain A^*; adjoining an identity to a semigroup of right zeros does not give a semigroup of right zeros.

Definition 1.19. (Idempotent). An ***idempotent*** is an element $e \in S$ satisfying $e^2 = e$.

For example, the identity of a monoid, or group is an idempotent; every element of a left zero semigroup is an idempotent; the unique idempotent in zero semigroup is the zero; the free semigroup has no idempotents. Idempotents other than the identity (if there is one) are not necessarily unique. Every finite semigroup contains an idempotent.

2. Subsemigroups and generators

Definition 2.1. (Subsemigroup). A ***subsemigroup*** of a semigroup S is any subset T of S which is closed under the operation of S; we write $T \leq S$.

For example, the symmetric group S_n is a subsemigroup of the full transformation monoid T_n, which is a subsemigroup of the partial transformation monoid P_n. Another example, $A^+ \leq A^*$.

Lemma 2.2 (**Intersection of subsemigroups is a subsemigroup**). *Let S be a semigroup, and let S_i $(i \in I)$ be any family of subsemigroups of S. Then*

$$T = \bigcap_{i \in I} S_i,$$

is a subsemigroup of S.

Proof. Let $x, y \in T$. Then $x, y \in S_i$ and so $xy \in S_i$ for all $i \in I$. Hence $xy \in T$. □

The intersection of two subsemigroups of a semigroup S may be empty; unlike for groups where the intersection always contains the identity.

Example 2.3. (Semigroup that is a disjoint union of subsemigroups). Let $(\mathbb{R}, +)$ denote the real numbers under addition. Then $(\mathbb{R}, +)$ is a group. The sets

$$P = \{x \in \mathbb{R} : x > 0\}$$

and

$$N = \{x \in \mathbb{R} : x \leq 0\}$$

are subsemigroups and $P \cap N = \varnothing$.

Example 2.4. If e and f are two distinct idempotents of a semigroup S, then both $\{e\}$ and $\{f\}$ are subsemigroups, and clearly $\{e\} \cap \{f\} = \varnothing$.

As in other areas of algebra, it is possible to generate subsemigroups using just a few elements. This gives us a method for constructing new semigroups from the known ones.

Definition 2.5. (Subsemigroup generated by a set). Let S be a semigroup and let X be a non-empty subset of S. If T_i $(i \in I)$ is the collection of all of the subsemigroups of S containing X, then

$$\langle X \rangle = \bigcap_{i \in I} T_i$$

is called to ***the subsemigroup generated by*** X.

The above definition gives us no clue as to what the elements of $\langle X \rangle$ are. In practice, finding all the subsemigroups of a semigroup containing a given set could be extremely difficult.

Theorem 2.6. *Let S be a semigroup, and let X be a non-empty subset of S. Then the following hold:*

(i) *$\langle X \rangle$ is the smallest (with respect to containment) subsemigroup of S containing X;*

[1] Some authors use S^1 to denote S if S is a monoid and the construction given above otherwise. That is, no identity is adjoined if S already has one.

(ii) $\langle X \rangle$ *consists precisely of all products of elements of* X *of finite length:*

$$\langle X \rangle = \{x_1 \cdots x_n : n \geq 1,\ x_i \in X\}.$$

Proof. **(i).** Let $\{T_i : i \in I\}$ be the collection of all subsemigroups of S containing X. If T is any subsemigroup of S containing X, then $T = T_j$ for some $j \in I$. Hence $\bigcap_{i \in I} T_i \subseteq T$ and so $\langle X \rangle \subseteq T$.

(ii). Suppose that $Y = \{x_1 \cdots x_n : n \geq 1,\ x_i \in X\}$. Then, clearly, $X \subseteq Y$ (products of length 1). The set Y is also a subsemigroup of S, and since $\langle X \rangle$ is the least subsemigroup containing X, it follows that $\langle X \rangle \subseteq Y$.

On the other hand, if $x_1 \cdots x_n \in Y$, then since $x_1, \ldots, x_n \in X$ and $\langle X \rangle$ is a subsemigroup, it follows that $x_1 \cdots x_n \in \langle X \rangle$. □

If a subset X of a semigroup S generates S, then X is called a ***generating set for*** S.

Example 2.7. Every semigroup generates itself, i.e. $S = \langle S \rangle$.

The semigroup $(\mathbb{N}, +)$ is generated by the element 1, that is, $\langle 1 \rangle = \mathbb{N}$. In the additive semigroup of integers $\mathbb{Z} = \langle -1, 1 \rangle$.

If S is a semigroup of left zeros, then $\langle X \rangle = X$ for every subset X of S.

The free semigroup A^+ is generated by the set A. Furthermore, if $X \subseteq A^+$ is a generating set for A^+ if and only if $A \subseteq X$.

Let $I \times \Lambda$ be a rectangular band (see Example 1.9). Let $i_0 \in I$ and $\lambda_0 \in \Lambda$ be arbitrary, and let

$$X = \{(i_0, \lambda) : \lambda \in \Lambda\} \cup \{(i, \lambda_0) : i \in I\}.$$

If $(i, \lambda) \in I \times \Lambda$ is arbitrary, then

$$(i, \lambda) = (i, \lambda_0)(i_0, \lambda) \in \langle X \rangle$$

and so $I \times \Lambda = \langle X \rangle$.

Definition 2.8. (Finitely generated semigroup). A semigroup S is ***finitely generated*** if there exists a finite set $X \subseteq S$ such that $\langle X \rangle = S$.

Example 2.9. Since $\langle S \rangle = S$ for all semigroups S, every finite semigroup is finitely generated. The additive semigroups $\mathbb{N}$ and $\mathbb{Z}$ are finitely generated.

Any infinite semigroup of left zeros is not finitely generated.

The free semigroup A^+ is finitely generated if and only if A is finite.

The rectangular band $I \times \Lambda$ is finitely generated if and only if it is finite.

Theorem 2.6 can be used to compute the elements of the subsemigroup generated by a set in the finite case. We just need a systematic procedure for obtaining all products of generators.

Algorithm 1 Generating the subsemigroup $\langle x_1, x_2, \ldots, x_n \rangle$

```
T ← {t_1 = x_1, t_2 = x_2, ..., t_n = x_n} (the list of elements so far)
l ← n (subscript of the last element of the list T)
i ← 0 (the current position in the list)
repeat
    i ← i + 1
    for j ∈ {1, 2, ..., n} do
        t ← t_i x_j
        if t ∉ T then
            l ← l + 1
            t_l ← t
        end if
    end for
until i = l
return T
```

Of course, computing $x_j t_i$ instead of $t_i x_j$ would give the same answer.

Example 2.10. Let S be the subsemigroup of the full transformation semigroup T_3 generated by

$$a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \end{pmatrix}.$$

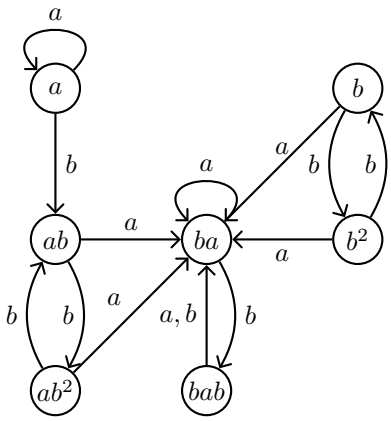

FIGURE 1. The right Cayley graph of the semigroup in Example 2.10.

We compute $\langle a, b\rangle$ using Algorithm 1:

$$
\begin{array}{llll}
a = \begin{pmatrix}1&2&3\\2&2&3\end{pmatrix} & \textbf{(new)} & b = \begin{pmatrix}1&2&3\\2&1&2\end{pmatrix} & \textbf{(new)}\\
a^2 = \begin{pmatrix}1&2&3\\2&2&3\end{pmatrix} = a & \textbf{(old)} & ab = \begin{pmatrix}1&2&3\\1&1&2\end{pmatrix} & \textbf{(new)}\\
ba = \begin{pmatrix}1&2&3\\1&1&1\end{pmatrix} & \textbf{(new)} & b^2 = \begin{pmatrix}1&2&3\\1&2&2\end{pmatrix} & \textbf{(new)}\\
aba = \begin{pmatrix}1&2&3\\2&2&2\end{pmatrix} = ba & \textbf{(old)} & ab^2 = \begin{pmatrix}1&2&3\\2&2&1\end{pmatrix} & \textbf{(new)}\\
ba^2 = \begin{pmatrix}1&2&3\\2&2&2\end{pmatrix} = ba & \textbf{(old)} & bab = \begin{pmatrix}1&2&3\\1&1&1\end{pmatrix} & \textbf{(new)}\\
b^2a = \begin{pmatrix}1&2&3\\2&2&2\end{pmatrix} = ba & \textbf{(old)} & b^3 = \begin{pmatrix}1&2&3\\2&1&2\end{pmatrix} = b & \textbf{(old)}\\
ab^2a = \begin{pmatrix}1&2&3\\2&2&2\end{pmatrix} = ba & \textbf{(old)} & ab^3 = \begin{pmatrix}1&2&3\\1&1&2\end{pmatrix} = ab & \textbf{(old)}\\
baba = \begin{pmatrix}1&2&3\\2&2&2\end{pmatrix} = ba & \textbf{(old)} & bab^2 = \begin{pmatrix}1&2&3\\2&2&2\end{pmatrix} = ab & \textbf{(old)}.
\end{array}
$$

Hence the elements of $\langle a, b\rangle$ are $\{a, b, ab, ba, b^2, ab^2, bab\}$.

Definition 2.11. (Cayley graph). Let S be a semigroup and let X be a generating set for S. Then the ***right Cayley graph*** of S with respect to X is defined to be the digraph with vertices S and edge $(u, v) \in S \times S$ labelled $x \in X$ whenever $ux = v$. The ***left Cayley graphs*** of semigroup with respect to X is defined analogously.

It is straightforward to draw the right Cayley graph of the semigroup $\langle a, b\rangle$ in Example 2.10 using the information produced by the algorithm; see Figure 1.

Proposition 2.12 (A generating set for the full transformation monoid). *The full transformation monoid T_n is generated by the permutations*

$$(1\ 2), \quad (1\ 2\ \cdots\ n),$$

and the transformation

$$\begin{pmatrix}1&2&3&\cdots&n\\2&2&3&\cdots&n\end{pmatrix}.$$

Proof. We require the following notation. Let $i, j \in \{1, \ldots, n\}$, $i \neq j$, and define $||i\ j|| \in T_n$ by

$$(x)||i\ j|| = \begin{cases} j & \text{if } x = i\\ x & \text{if } x \neq i.\end{cases}$$

With this notation the non-permutation in the statement is $||1\ 2||$.

Recall that $S_n = \langle (1\ 2), \quad (1\ 2\ \cdots\ n)\rangle$, and so we must show that every $f \in T_n \setminus S_n$ belongs to $\langle ||1\ 2||, (1\ 2), (1\ 2 \cdots\ n)\rangle$.

Step 1. It is straightforward to verify that

$$(1\ i)||1\ 2||(1\ i) = ||i\ 2|| \quad i \neq 1, 2,$$

$$(2\ j)||1\ 2||(2\ j) = ||1\ j|| \quad j \neq 1, 2,$$

$$(1\ i)(2\ j)||1\ 2||(2\ j)(1\ i) = ||i\ j||, \quad i, j \neq 1, 2,$$

and

$$(i\ j)||i\ j||(i\ j) = ||j\ i||, \quad \text{for all } i, j.$$

(The last of these equalities is required in order to generate elements of the form $||j\ 1||$ and $||2\ i||$.)

It follows that for all i, j

$$||i\ j|| \in \langle ||1\ 2||, (1\ 2), (1\ 2 \cdots\ n) \rangle.$$

Step 2. *Let $f \in T_n \setminus S_n$ be arbitrary. Then there exists $g \in T_n$ and $i, j \in \{1, 2, \ldots, n\}$ with* $\text{rank}(g) = \text{rank}(f) + 1$ *and $f = ||ij|| \circ g$.*

Let $i, j \in \{1, 2, \ldots, n\}$ such that $i \neq j$ and $if = jf$ (f is not injective!), and let $k \in \{1, 2, \ldots, n\} \setminus \text{im}(f)$ (f is not surjective!) and define g by

$$xg = \begin{cases} k & x = i \\ xf & x \neq i. \end{cases}$$

Now, $\text{im}(g) = \text{im}(f) \cup \{k\}$ and so $\text{rank}(g) = \text{rank}(f) + 1$, as required. Finally, $(x)||ij|| \neq i$ for all x and so $f = ||ij|| \circ g$.

Step 3. Let $f_0 = f \in T_n$ be arbitrary. Then either $f_0 \in S_n$ or $f_0 \in T_n \setminus S_n$. In the first case, $f_0 \in \langle (1\ 2), (1\ 2 \cdots n) \rangle$ and we are finished. In the second case, $f_0 \in T_n \setminus S_n$ and so, by Step 2, we can write

$$f_0 = ||i_0\ j_0||\ f_1$$

for some i_0, j_0 and f_1 satisfying $\text{rank}(f_1) = \text{rank}(f_0) + 1$. If $\text{rank}(f_0) = n - k$ for some k, then repeatedly applying Step 2 we obtain

$$f_0 = ||i_0\ j_0||\ f_1 = ||i_0\ j_0||\ ||i_1\ j_1||\ f_2 = \cdots = ||i_0\ j_0||\ ||i_1\ j_1|| \cdots ||i_{k-1}\ j_{k-1}||\ f_k$$

where $f_k \in S_n$ (since $\text{rank}(f_k) = n$). It follows that $f \in \langle ||1\ 2||, (1\ 2), (1\ 2 \cdots n) \rangle$. □

3. Homomorphisms

You may already have encountered homomorphisms of groups, graphs, or rings. The notion of a semigroup homomorphism is analogous.

Definition 3.1. (**Homomorphism**). Let S and T be semigroups. Then a ***semigroup homomorphism*** is any function $f : S \longrightarrow T$ such that

$$(xy)f = (x)f\,(y)f$$

for all $x, y \in S$.

Example 3.2. If S is any semigroup, then the identity function $1_S : S \longrightarrow S$ is a homomorphism.

If $e \in S$ is an idempotent, then the constant function $s \mapsto e$ for all $s \in S$ is a homomorphism.

Two semigroups S and T are said to be ***isomorphic*** if there exists an isomorphism $f : S \longrightarrow T$; we denote this by $S \cong T$. If S and T are isomorphic, then they differ in the names of their elements only. We normally study semigroups *up to isomorphism*, meaning that we will not distinguish between isomorphic semigroups.

Lemma 3.3. *If $f : S \longrightarrow T$ is a homomorphism, then the image* $\text{im}(f)$ *of f is a subsemigroup of T. Moreover, if f is injective, then $S \cong \text{im}(f)$.*

Proof. As an exercise. □

Theorem 3.4 (**Analogue of Cayley's Theorem**). *Every semigroup is isomorphic to a subsemigroup of a full transformation semigroup.*

Proof. Let $X = S^1$ (S with an identity adjoined) and for all $s \in S$ define $\tau_s : X \longrightarrow X$ by $(x)\tau_s = xs$. Then define $\Psi : S \longrightarrow T_X$ by $(s)\Psi = \tau_s$. If Ψ is a monomorphism, then the result follows by Lemma 3.3.

Homomorphism. By definition $(st)\Psi = \tau_{st}$. But

$$(x)\tau_{st} = xst = (xs)\tau_t = ((x\tau_s)\tau_t).$$

It follows that

$$(st)\Psi = \tau_{st} = \tau_s\tau_t = (s)\Psi(t)\Psi,$$

as required.

Injective. Assume $(s)\Psi = (t)\Psi$ for some $s, t \in S$. We want to show that $s = t$. From the assumption, we deduce that $\tau_s = \tau_t$ and so $xs = xt$ for all $x \in X$. In particular, $1s = 1t$ and so $s = t$. □

Example 3.5. Let $S = \{x, y, z\}$ be the three element semigroup of left zeros. Then, according to the above theorem, take $X = S^1 = \{1, x, y, z\}$. The functions corresponding to x, y and z are

$$\tau_x = \begin{pmatrix} 1 & x & y & z \\ x & x & y & z \end{pmatrix} \tau_y = \begin{pmatrix} 1 & x & y & z \\ y & x & y & z \end{pmatrix} \tau_z = \begin{pmatrix} 1 & x & y & z \\ z & x & y & z \end{pmatrix}.$$

Example 3.6. Let $S = \{x, y, z, 0\}$ be a zero semigroup. Then, by Theorem 3.4, $X = S^1 = \{x, y, z, 0, 1\}$. The functions corresponding τ_s, $s \in S$ given in the proof of Theorem 3.4 are

$$\tau_x = \begin{pmatrix} x & y & z & 0 & 1 \\ 0 & 0 & 0 & 0 & x \end{pmatrix} \tau_y = \begin{pmatrix} x & y & z & 0 & 1 \\ 0 & 0 & 0 & 0 & y \end{pmatrix} \tau_z = \begin{pmatrix} x & y & z & 0 & 1 \\ 0 & 0 & 0 & 0 & z \end{pmatrix} \tau_0 = \begin{pmatrix} x & y & z & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

If $X = S$ was used instead of $X = S^1$ in Theorem 3.4, then in the last example

$$\tau_x = \tau_y = \tau_z = \tau_0 = \begin{pmatrix} x & y & z & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

It follows that Ψ is very far from being injective, and so $S \ncong \mathrm{im}(\Psi)$.

Definition 3.7. (Homomorphic image). If S and T are semigroups and $f : S \longrightarrow T$ is a surjective homomorphism (an ***epimorphism***), then we say that T is a ***homomorphic image*** of S.

Theorem 3.4 states that every semigroup is **isomorphic** to a subsemigroup of some full transformation monoid T_X. Free semigroups play a role similar to the full transformation semigroups with respect to homomorphic images instead of subsemigroups.

Theorem 3.8 (**Free semigroups are free**). *Let S be a semigroup, let A be an alphabet, and let $f : A \longrightarrow S$ be any function. Then there exists a unique homomorphism $\phi : A^+ \longrightarrow S$ such that $(a)f = (a)\phi$ for all $a \in A$. In other words, the following diagram commutes:*

$$\begin{array}{ccc} A & \xrightarrow{f} & S \\ \downarrow & \nearrow_{\phi} & \\ A^+ & & \end{array}$$

Proof. Recall that A^+ is the set of all the non-empty finite words over A with juxtaposition as the operation.

Let $a_1a_2 \cdots a_m \in A^+$ where $a_1, a_2, \ldots, a_m \in A$. Then define $\phi : A^+ \longrightarrow S$ by

$$(a_1a_2 \cdots a_m)\phi = (a_1f)(a_2f) \cdots (a_mf).$$

Homomorphism. Let $a_1a_2 \ldots a_m, b_1b_2 \ldots b_n \in A^+$ be arbitrary. Then

$$\begin{aligned} & ((a_1a_2 \ldots a_m)(b_1b_2 \ldots b_n))\phi = (a_1a_2 \ldots a_mb_1b_2 \ldots b_n)\phi \\ = \quad & (a_1f)(a_2f) \ldots (a_mf)(b_1f)(b_2f) \ldots (b_nf) = (a_1a_2 \ldots a_m)\phi(b_1b_2 \ldots b_n)\phi, \end{aligned}$$

and so ϕ is a homomorphism. It is clear that $a\phi = af$ for all $a \in A$.

Uniqueness of ϕ. Let $\psi : A^+ \longrightarrow S$ be any homomorphism such that $(a)\psi = (a)f$ for all $a \in A$. Then for any $a_1a_2 \cdots a_m \in A^+$

$$(a_1a_2 \ldots a_m)\psi = (a_1\psi)(a_2\psi) \ldots (a_m\psi) = (a_1f)(a_2f) \ldots (a_mf) = (a_1a_2 \ldots a_m)\phi,$$

and so $\psi = \phi$. □

Corollary 3.9 (**Every semigroup is a homomorphic image of a free semigroup**). *Let S be a semigroup, let A be an alphabet, and let $f : A \longrightarrow S$ be a function such that $S = \langle \mathrm{im}(f) \rangle$. Then the unique homomorphism $\phi : A^+ \longrightarrow S$ such that $(a)f = (a)\phi$ for all $a \in A$ is surjective.*

Proof. Let $s \in S$ be arbitrary. Then $s = x_1 x_2 \dots x_n$ for some $x_1, \dots, x_n \in \mathrm{im}(f)$ and $x_i = a_i f$ ($a_i \in A$, $i = 1, \dots, n$). Hence

$$s = (a_1 f)(a_2 f) \dots (a_n f) = (a_1 a_2 \dots a_n)\phi,$$

and so ϕ is surjective. □

4. Binary relations

The purpose of this section is to introduce some basic definitions and properties of binary relations.

Definition 4.1. (**Binary relation**). A ***binary relation*** ρ on a set X is just a subset of $X \times X$. If $\rho \subseteq X \times X$, then we might write $x \rho y$ rather than $(x, y) \in \rho$.

Example 4.2. The set $\{(1, 2)\}$ is a binary relation on $\{1, 2, 3\}$.

Define $\rho \subseteq \mathbb{N} \times \mathbb{N}$ by $(x, y) \in \rho$ whenever x divides y. Clearly $(2, 4) \in \rho$ and $(2, 5) \notin \rho$.

The usual relations $\leq$, $=$ are binary relations on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and so on.

Since binary relations are sets we take intersections, unions, differences, and compose them. That is, if ρ, σ are binary relations on X, then we can find $\rho \cap \sigma$, $\rho \cup \sigma$, $\rho \setminus \sigma$, and $\rho \circ \sigma$:

$$\rho \circ \sigma = \rho\sigma = \{(x, y) \in X \times X : (\exists z \in X)((x, z) \in \rho \text{ and } (z, y) \in \sigma)\}.$$

Lemma 4.3. *The set B_X of all binary relations on X is a monoid with zero with operation composition of relations.*

Proof. It suffices to prove that composition of relations is associative

$$\begin{aligned}
(x, y) \in (\rho\sigma)\tau \quad &\text{if and only if} \quad (\exists z \in X)((x, z) \in \rho\sigma \text{ and } (z, y) \in \tau) \\
&\text{if and only if} \quad (\exists z \in X)(\exists u \in X)((x, u) \in \rho \text{ and } (u, z) \in \sigma \text{ and } (z, y) \in \tau) \\
&\text{if and only if} \quad (\exists z \in X)(\exists u \in X)((x, u) \in \rho \text{ and } (u, z) \in \sigma \text{ and } (z, y) \in \tau) \\
&\text{if and only if} \quad (\exists u \in X)((x, u) \in \rho \text{ and } (u, y) \in \sigma\tau) \\
&\text{if and only if} \quad (x, y) \in \rho(\sigma\tau).
\end{aligned}$$

It is straightforward to check that the relation

$$\Delta = \Delta_X = \{(x, x) : x \in X\}$$

is the identity of B_X; this relation is also sometimes referred to as the **diagonal**. The zero element of B_X is the empty relation $\varnothing$. □

A function $\rho : X \longrightarrow X$ (either partial or full) on X is a binary relation on X satisfying

$$|\{y \in X : (x, y) \in \rho\}| \leq 1$$

for all $x \in \mathrm{dom}(\rho)$. Composition of functions is a special case of composition of relations. In other words,

$$T_X \leq P_X \leq B_X.$$

Definition 4.4. (**Inverse of a binary relation**). The ***inverse*** of a binary relation $\rho \in B_X$ is defined to be $\rho^{-1} = \{(x, y) : (y, x) \in \rho\}$.

It is important to note that $\rho^{-1}\rho$ and $\rho\rho^{-1}$ are not necessarily equal to the identity Δ_X, or in other words that B_X is not a group.

Definition 4.5. Let ρ be a relation on X. We say that ρ is:

- **(R)** ***Reflexive:*** if $(x, x) \in \rho$ for all $x \in X$;
- **(S)** ***Symmetric:*** if $(x, y) \in \rho$ implies $(y, x) \in \rho$ for all $x, y \in X$;
- **(AS)** ***Antisymmetric:*** if $(x, y) \in \rho$ and $(y, x) \in \rho$ imply $x = y$ for all $x, y \in X$;
- **(T)** ***Transitive:*** if $(x, y) \in \rho$ and $(y, z) \in \rho$ imply $(x, z) \in \rho$ for all $x, y, z \in X$.

Definition 4.6. (**Equivalence relation**). An ***equivalence relation*** is a reflexive, symmetric, and transitive binary relation. A ***partial order*** is a reflexive, antisymmetric, and transitive binary relation.

A ***partition*** of a set X is any collection of disjoint subsets of X whose union is X. If $\rho \subseteq X \times X$ is an equivalence relation, then the sets $x/\rho = \{y \in X : (x, y) \in \rho\}$, $x \in X$, form a partition of X, and $x/\rho = y/\rho$ if and only if $(x, y) \in \rho$. If $x \in X$, then x/ρ is called the ***equivalence class*** of x. The set of all equivalence classes is

$$X/\rho = \{x/\rho_x : x \in X\}.$$

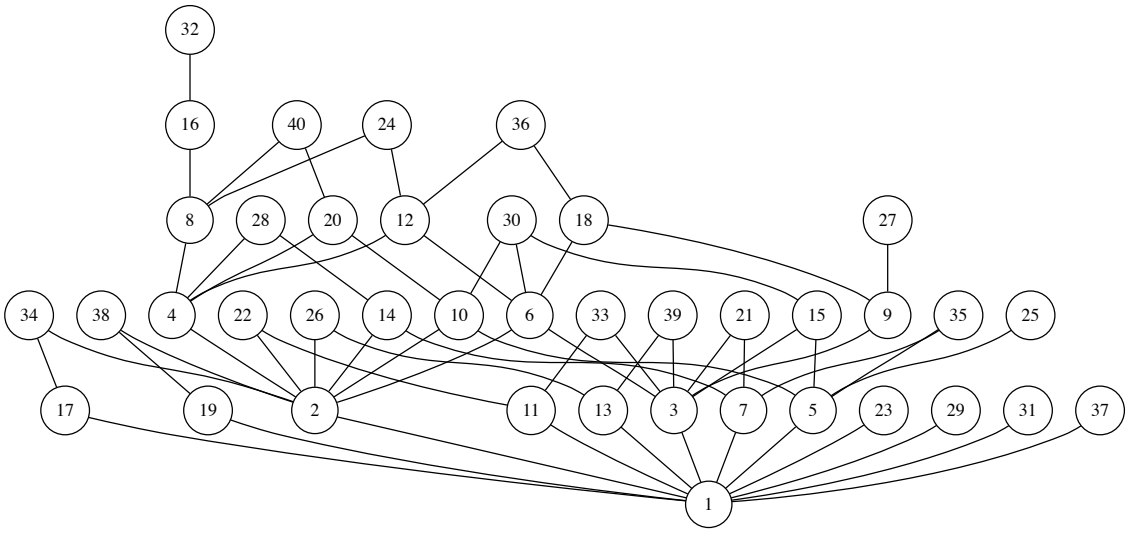

FIGURE 2. A Hassé diagram of the divisibility relation from Example 4.10.

Example 4.7. (**Equivalence relations**). Let $X = \{1, \ldots, 10\}$, and let ρ be the relation defined by

$$(x, y) \in \rho \text{ if and only if } x \equiv y \mod 3.$$

Clearly ρ is an equivalence relation and its equivalence classes are $\{1, 4, 7, 10\}$, $\{2, 5, 8\}$ and $\{3, 6, 9\}$.

Definition 4.8. (**Kernel of a function**). Let X and Y be sets and let $f : X \longrightarrow Y$ be any function. Then the ***kernel*** of f is the equivalence relation

$$\ker(f) = \{(x, y) \in X \times X : (x)f = (y)f\}.$$

The equivalence classes of $\ker(f)$ are called the ***kernel classes*** of f.

It is straightforward to see that the equivalence classes of $\ker(f)$ are in one-one correspondence with the elements of $\mathrm{im}(f)$, and so there are precisely $\mathrm{rank}(f)$ kernel classes. In Section 5 we will see that kernel classes of a semigroup homomorphism are analogous to cosets of the kernel of a group homomorphism.

Example 4.9. Let

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 1 & 2 & 7 & 5 & 3 & 1 \end{pmatrix}.$$

Then $\ker(f) = \{\{1, 2, 7\}, \{3\}, \{4\}, \{5\}, \{6\}\}$. Note that $\mathrm{rank}(f) = 5$ and there are 5 kernel classes.

Order relations are often visualised by means of ***Hassé*** diagrams.

Example 4.10. (**Hassé diagram**). Let $X = \{1, \ldots, 40\}$, and let $\rho = \{(x, y) \in X \times X : x \mid y\}$. Then ρ is a partial order, and its Hassé diagram is shown in Figure 2.

5. Congruences and quotients

In this section, we will consider a notion analogous to that of a normal subgroup of a group, or an ideal of a ring.

Definition 5.1. (**Congruence**). Let S be a semigroup. Then an equivalence relation $\rho \subseteq S \times S$ is called a ***left congruence*** if $(sx, sy) \in \rho$ for all $(x, y) \in \rho$ and for all $s \in S$. ***Right congruences*** are defined analogously, and a (two-sided) ***congruence*** is left congruence that is also a right congruence.

The equivalence classes of a congruence are called ***congruence classes***.

Proposition 5.2 (**Kernel of a homomorphism is a congruence**). *Let S and T be semigroups and let $f : S \longrightarrow T$ be a homomorphism. Then* $\ker(f) = \{(x, y) \in S \times S : (x)f = (y)f\}$ *is a congruence.*

Proof. Let $(x, y) \in \ker(f)$ and let $s \in S$ be arbitrary. Then $(x)f = (y)f$ and so

$$(x)f\,(s)f = (y)f\,(s)f.$$

Since f is a homomorphism, it follows that $(xs)f = (ys)f$ and so $(xs, ys) \in \ker(f)$. A similar argument shows that $(sx, sy) \in \ker(f)$, and so $\ker(f)$ is a congruence. □

Congruences have a natural multiplication on their equivalence classes.

Theorem 5.3 (**Quotient semigroups are well-defined**). *Let ρ be a congruence on a semigroup S. Then the set*

$$S/\rho = \{x/\rho : x \in S\}$$

of all congruence classes of ρ forms a semigroup under the multiplication defined by

$$(x/\rho)(y/\rho) = (xy)/\rho$$

*[S/ρ is called the **quotient** of S by ρ.]*

Proof. The set S/ρ is clearly closed under the given operation, and if $x/\rho, y/\rho, z/\rho \in S/\rho$, then

$$((x/\rho)(y/\rho))(z/\rho) = ((xy)/\rho)(z/\rho) = ((xy)z)/\rho = (x(yz))/\rho = (x/\rho)((y/\rho)(z/\rho)),$$

and so the multiplication is associative.

It remains to show that the multiplication is well defined, i.e. that it does not depend on the choice of the representatives x and y of equivalence classes. We want to prove that if $x/\rho = x'/\rho$ and $y/\rho = y'/\rho$, then $xy/\rho = x'y'/\rho$. In other words, we want to prove that if $(x, x'), (y, y') \in \rho$, then $(xy, x'y') \in \rho$. But $(x, x') \in \rho$ implies that $(xy, x'y) \in \rho$ (ρ is a right congruence!) and $(y, y') \in \rho$ implies that $(x'y, x'y') \in \rho$. Thus by transitivity $(xy, x'y') \in \rho$. □

Quotients of semigroups differ from the analogous notions you might have met in groups, rings, or vector spaces. In these examples, the notions of normal subgroups, ideals, and null spaces are used to define quotients, respectively. Quotients of semigroups are defined in terms of a binary relation. As it turns out, the semigroup approach is more common in general algebra.

Theorem 5.4 (**The First Isomorphism Theorem**). *Let S and T be semigroups. Then the following hold:*

(i) *If ρ is a congruence on S, then $f : S \longrightarrow S/\rho$ defined by $(x)f = x/\rho$ is a surjective homomorphism.*
(ii) *If $f : S \longrightarrow T$ is a homomorphism, then $S/\ker(f) \cong \mathrm{im}(f)$.*

Proof. **(i).** Let $x/\rho \in S/\rho$ be arbitrary. Then $(x)f = x/\rho$ and so f is surjective. By the definition of the multiplication in S/ρ (Theorem 5.3), if $x, y \in S$, then

$$(xy)f = (xy)/\rho = (x/\rho)(y/\rho) = (x)f\,(y)f.$$

Hence f is a homomorphism.

(ii). We saw in Proposition 5.2 that $\ker(f)$ is a congruence and in Lemma 3.3 that $\mathrm{im}(f)$ is a subsemigroup of T. Let $\phi : S/\ker(f) \longrightarrow \mathrm{im}(f)$ be defined by

$$(x/\ker(f))\phi = (x)f.$$

There are 4 things to check: ϕ is well-defined, injective, surjective and a homomorphism.

Well-defined. Let $x, y \in S$ such that $x/\ker(f) = y/\ker(f)$. We must show that the value of ϕ on $x/\ker(f) = y/\ker(f)$ does not depend on the choice of x and y. In other words, we want to show that $(x/\ker(f))\phi = (y/\ker(f))\phi$. But $x/\ker(f) = y/\ker(f)$ implies that $(x, y) \in \ker(f)$ and so $xf = yf$. Thus $(x/\ker(f))\phi = xf = yf = (y/\ker(f))\phi$.

Injective. Let $x, y \in S$ such that $(x/\ker(f))\phi = (y/\ker(f))\phi$. Then $xf = yf$ implies that $(x, y) \in \ker(f)$ and so $x/\ker(f) = y/\ker(f)$.

Surjective. If $xf \in \mathrm{im}(f)$, then $(x/\ker(f))\phi = xf$ and so ϕ is surjective.

Homomorphism. Let $x, y \in S$ be arbitrary. Then

$$(x/\ker(f))\phi\,(y/\ker(f))\phi = (x)f\,(y)f = (xy)f = (xy/\ker(f))\phi,$$

as required. □

The following proposition is a reformulation of Corollary 3.9.

Proposition 5.5. *Every semigroup is isomorphic to the quotient of a free semigroup.*

Proof. Let S be any semigroup, and let A be any generating set for S. Then, by Corollary 3.9, if $f : A \longrightarrow S$ is defined by $(a)f = a$ for all $a \in A$, then there exists a (unique) surjective homomorphism $\phi : A^+ \longrightarrow S$ such that $(a)f = (a)\phi$ for all $a \in A$. Thus, by Theorem 5.4, $A^+/\ker(\phi) \cong \operatorname{im}(\phi) = S$. □

If S is any semigroup, then Theorem 3.4 shows that there exists a monomorphism from S into T_{S^1}. If S is not too large, then it is straightforward to find an explicit subsemigroup of T_{S^1} isomorphic to S. On the other hand, Proposition 5.5 states that S is a quotient of a free semigroup by a congruence ρ.

Theorem 5.6 (**3rd Isomorphism Theorem**). *Let ρ and σ be congruences of a semigroup S such that $\rho \subseteq \sigma$. Then*

$$\sigma/\rho = \{(x/\rho, y/\rho) \in S/\rho \times S/\rho : (x, y) \in \sigma\}$$

is a congruence on S/ρ and that $(S/\rho)/(\sigma/\rho) \cong S/\sigma$. In other words, the following diagram commutes:

$$\begin{array}{ccc} S & \xrightarrow{\alpha} & S/\rho \\ \downarrow{\scriptstyle \gamma} & \swarrow{\scriptstyle \beta} & \\ S/\sigma & & \end{array}$$

where α, β, and γ are the natural homomorphisms from Theorem 5.4(i) defined by $(s)\alpha = s/\rho$, $(s/\rho)\beta = s/(\sigma/\rho)$, and $(s)\gamma = s/\sigma$ for all $s \in S$.

Proof. This is Problem **4-4**. □

This section is somewhat devoid of examples of congruences, we will end this section by showing how to succinctly define a congruence on a semigroup, and how to compute its classes using an algorithm similar to Algorithm 1.

Proposition 5.7. *Let S be a semigroup and let $\rho, \sigma \subseteq S \times S$. Then*

(i) *if ρ and σ are equivalence relations, then $\rho \cap \sigma$ is an equivalence relation;*
(ii) *if ρ and σ are congruences, then $\rho \cap \sigma$ is a congruence.*

Proof. **(i).** This is Problem **3-3**.

(ii). Suppose that $(x, y) \in \rho \cap \sigma$ and $s \in S$. Then $(x, y) \in \rho$ and so $(xs, ys), (sx, sy) \in \rho$, and similarly $(xs, ys), (sx, sy) \in \sigma$. Hence $(xs, ys), (sx, sy) \in \rho \cap \sigma$ and so $\rho \cap \sigma$ is a congruence. □

It follows from Proposition 5.7 that if $\rho \subseteq S \times S$, then we can define the ***least equivalence relation containing*** ρ to be the intersection of every equivalence relation on S containing ρ. The ***least congruence containing*** ρ is defined analogously, and denoted $\rho^\sharp$. We defined a subsemigroup of a semigroup S generated by a set X to be the intersection of all the subsemigroups of S containing X. Similarly, we refer to the congruence $\rho^\sharp$ ***congruence generated by*** ρ. This definition is not very useful when trying to figure out what the congruence generated by a given set of pairs is however.

Proposition 5.8. *Let X be a set and let $\rho \subseteq X \times X$. Then (x, y) belongs to the least equivalence relation on X containing ρ if and only if $x = y$ or there exists a sequence $x = x_1, x_2, \ldots, x_n = y$ such that $(x_i, x_{i+1}) \in \rho$ or $(x_{i+1}, x_i) \in \rho$ for all i.*

Proof. Suppose that $\sigma \subseteq S \times S$ is such that $(x, y) \in \sigma$ if and only if $x = y$ or there exists a sequence $x = x_1, x_2, \ldots, x_n = y$ such that $(x_i, x_{i+1}) \in \rho$ or $(x_{i+1}, x_i) \in \rho$ for all i. It is straightforward to verify that σ is reflexive, symmetric, and transitive, and hence σ is an equivalence relation.

Let τ be any equivalence relation on X with $\rho \subseteq \tau$, and let $(x, y) \in \sigma$. Then there exists a sequence $x = x_1, x_2, \ldots, x_m = y$ such that $(x_i, x_{i+1}) \in \rho$ or $(x_{i+1}, x_i) \in \rho$ for all i. Since τ contains ρ and τ is transitive, it follows that $(x, y) \in \tau$. Hence $\sigma \subseteq \tau$ and so σ is the least equivalence relation on S containing ρ. □

Before we can describe the least congruence containing a given binary relation, we require the following definition.

Definition 5.9. (**Elementary sequence**). If S is a semigroup and $\rho \subseteq S \times S$, then there is an ***elementary sequence*** between $x, y \in S$ if $x = y$ or $x = s_1, s_2, \ldots, s_m = y$, for some $m \geq 2$, where $s_i = p_i u_i q_i$, $s_{i+1} = p_i v_i q_i$, $p_i, q_i \in S^1$ and (u_i, v_i) or $(v_i, u_i) \in \rho$ for all i.

Theorem 5.10. *Let S be a semigroup, let $\rho \subseteq S \times S$, and let $\rho^\sharp$ be the least congruence of S containing ρ. Then the following hold:*

(i) $(x, y) \in \rho^\sharp$ *if and only if there is an elementary sequence between* x *and* y*;*
(ii) $\rho^\sharp$ *is the least equivalence relation on* S *containing*

$$\{(puq, pvq) \in S \times S : p, q \in S^1,\ (u, v) \in \rho\}.$$

Proof. **(i).** Suppose that $\sigma \subseteq S \times S$ is such that $(x, y) \in \sigma$ if and only if there is an elementary sequence between x and y. We will show that $\rho^\sharp = \sigma$, by showing that σ is the least congruence containing ρ.

Certainly, $\rho \subseteq \sigma$, and it is straightforward to check that σ is a congruence on S.

Suppose that τ is any congruence on S with $\rho \subseteq \tau$, and suppose that $(x, y) \in \sigma$. Then there exists an elementary sequence $x = s_1, s_2, \ldots, s_m = y$. Hence, for every i, there exist $p_i, q_i \in S^1$ and $(u_i, v_i) \in \rho$ such that $s_i = p_i u_i q_i$ and $s_{i+1} = p_i v_i q_i$. Since $(u_i, v_i) \in \rho \subseteq \tau$ it follows that $(s_i, s_{i+1}) = (p_i u_i q_i, p_i v_i q_i) \in \tau$ for all i. But τ is transitive, and so $(x, y) \in \tau$. Hence $\sigma \subseteq \tau$ and so σ is the least congruence on S containing ρ.

(ii). Let σ denote the least equivalence relation on S containing

$$\zeta = \{(puq, pvq) \in S \times S : p, q \in S^1,\ (u, v) \in \rho\}.$$

If $p, q \in S^1$ are arbitrary and $(u, v) \in \rho$, then $(xuy, xvy) \in \rho^\sharp$, since $\rho^\sharp$ is a congruence. Hence $\rho^\sharp$ is an equivalence relation containing ζ and so $\sigma \subseteq \rho^\sharp$.

For the converse, if $(x, y) \in \rho^\sharp$, then, by part (i), there exists a sequence $x = s_1, s_2, \ldots, s_n = y$ such that, for every i, $s_i = p_i u_i q_i$ and $s_{i+1} = p_i v_i q_i$ for some $p_i, q_i \in S^1$ and where $(u_i, v_i) \in \rho$ or $(v_i, u_i) \in \rho$. In other words, $(s_i, s_{i+1}) \in \zeta$ or $(s_{i+1}, s_i) \in \zeta$ for every i. Hence, by Proposition 5.8, $(x, y) \in \sigma$. and so $\sigma = \rho^\sharp$. □

Theorem 5.10(ii) suggests a means of computing the least congruence containing a binary relation ρ on a semigroup S which we elaborate in the following example.

Example 5.11. Let $S = \{a, b, c, d\}$ be the semigroup with multiplication defined by the following ***Cayley*** or ***multiplication table***:

	a	b	c	d
a	a	b	c	c
b	b	c	a	a
c	c	a	b	b
d	c	a	b	b

.

The value of the product ab, say, is the element in the row labelled a and the column labelled b. For example, $ab = b$ and $ba = b$. It is possible to verify that $\langle d \rangle = \{d, d^2 = b, d^3 = a, d^4 = c\} = S$, and that S is commutative.

We left and right multiply every pair, similar to in Algorithm 1, to compute $\rho^\sharp$ when $\rho = \{(a, b)\}$:

$$\begin{array}{lcll} (ad, bd) & = & (c, a) & \textbf{(new)} \\ (cd, ad) & = & (b, c) & \textbf{(new)} \\ (bd, cd) & = & (a, b) & \textbf{(old)} \end{array}$$

and at the end of the algorithm the table is:

a	b	c	d
1	2	3	4
1	1	3	4
1	1	1	4

Hence $\rho^\sharp = \{\{a, b, c\}, \{d\}\}$.

Even for small examples, this approach is extremely tedious, and in general it has complexity $O(|S|^2|)$.

6. Presentations

Presentations are a means of defining semigroups as quotients of free semigroups. It can also be argued that they are a succinct, aesthetic, and computationally useful.

In this section we show how to succinctly define the congruence $\ker(\phi) = \{(x, y) \in A^+ \times A^+ : x\phi = y\phi\}$ in Proposition 5.5 such that $A^+ / \ker(\phi) \cong S$.

Example 6.1. Let $S = \{a, b, c\}$ be the semigroup with multiplication defined by the Cayley table:

$\cdot$	a	b	c
a	b	c	b
b	c	b	c
c	b	c	b

.

Since $\langle a \rangle = \{a, a^2 = b, a^3 = c\} = S$, S is monogenic, and

$$a^4 = c \cdot a = b = a^2.$$

So if $A = \{a\}$, then $\phi : A^+ \longrightarrow S$ from the proof of Proposition 5.5 is defined by $a^i \mapsto a^i$. Let $i, j \in \mathbb{N}$ such that $i < j$ and $(a^i, a^j) \in \ker(\phi)$. Then $a^i = (a^i)\phi = (a^j)\phi = a^j$ in S and so

$$a^i = a^{i-4}a^4 = a^{i-2} = a^{i-4} = \cdots = a^{(i \bmod 2)} = a^{(j \bmod 2)} = \cdots = a^{j-2} = a^j.$$

Hence

$$\begin{aligned}
\ker(\phi) &= \{(a^i, a^j) \in A^+ \times A^+ : (a^i)\phi = (a^j)\phi\} \\
&= \{(a^i, a^j) \in A^+ \times A^+ : i \equiv j \mod 2\} \\
&= \{(a, a), (a^2, a^2), \cdots, (a^4, a^2), (a^2, a^4), (a^4, a^6), \ldots, (a^5, a^3), (a^7, a^3), (a^5, a^7), \ldots\} \\
&= \{(a^2, a^4)\}^\sharp.
\end{aligned}$$

Definition 6.2. (**Semigroup presentation**). A ***semigroup presentation*** is a pair $\langle A \,|\, R \rangle$ where A is an alphabet and $R \subseteq A^+ \times A^+$ is a set of relations on A^+. The ***semigroup defined by the presentation*** $\langle A|R \rangle$ is any semigroup isomorphic to $A^+/R^\sharp$ where $R^\sharp \subseteq A^+ \times A^+$ is the least congruence on A^+ containing R.

The elements of A are called the ***generators*** and the elements of R are called the ***defining relations*** or ***relations*** of the presentation. Typically, we write $u = v$ instead of $(u, v) \in R$. The reasons behind this conventions should become clear later in this section. We usually write A and R just by listing their elements, and not as sets: for example, we write

$$\langle a, b \,|\, a^5 = a,\ b^6 = b,\ ab = ba \rangle$$

and not

$$\langle \{a, b\} \,|\, \{(a^5, a), (b^6, b), (ab, ba)\} \rangle.$$

The elements of the semigroup S defined by a presentation $\langle A|R \rangle$ are the congruence classes $w/R^\sharp$ where $w \in A^+$ and $R^\sharp$ is the least congruence on A^+ containing R. We may say that $w \in A^+$ ***represents*** the element $w/R^\sharp$ of S. However, beware that the symbol w will have two meanings, $w \in A^+$ and $w/R^\sharp \in S$, which may cause confusion. If in doubt write $w/R^\sharp$ instead of w wherever appropriate!

The intuitive idea behind the semigroup S defined by a presentation $\langle A \,|\, R \rangle$ is that S is generated by A, that products of these generators are governed by the defining relations R, and that S is in a certain sense the largest semigroup with these two properties.

Example 6.3. The semigroup S given in Example 6.1 is defined by the presentation $\langle a \,|\, a^2 = a^4 \rangle$, and also by the presentation

$$\langle a | a^2 = a^2, a = a, a^3 = a^3, a^4 = a^2, a^2 = a^4, a^5 = a^3, a^6 = a^2 \rangle.$$

The presentation

$$\langle a \,|\, a^{m+r} = a^m \rangle$$

defines the monogenic semigroup $S = \langle a \rangle$ where $m, r \in \mathbb{N}$ are the least values such that $a^{m+r} = a^m$.

The semigroup defined by the presentation

$$\langle A \,|\, \rangle$$

is the free semigroup A^+. The least congruence containing the empty set $\varnothing$ is $\Delta_{A^+} = \{(w, w) : w \in A^+\}$.

Theorem 6.4. *Let S be the semigroup defined by the presentation $\langle A \,|\, R \rangle$. If T is any semigroup such that there exists $f : A \longrightarrow T$ with $T = \langle (A)f \rangle$ and*

$$(a_1)f \cdots (a_m)f = (b_1)f \cdots (b_n)f$$

for every relation $a_1 \cdots a_m = b_1 \cdots b_n$ in R, then T is a homomorphic image of S.

Proof. Let ρ denote the least congruence on A^+ containing R, let $\gamma : A^+ \longrightarrow T$ be the unique surjective homomorphism extending $f : A \longrightarrow T$ (Theorem 3.8), and let σ denote the kernel of γ. Then, by the First Isomorphism Theorem (Theorem 5.4(ii)), $S \cong A^+/\rho$ and $T \cong A^+/\sigma$. We will show that $\rho \subseteq \sigma$ and so, by Theorem 5.6, it will follow that $\beta : A^+/\rho \longrightarrow A^+/\sigma$ defined by $(w/\rho)\beta = w/(\sigma/\rho)$ is a homomorphism. The diagram is:

$$\begin{array}{ccc} A^+ & \xrightarrow{\alpha} & A^+/\rho \cong S \\ \uparrow & \overset{\gamma}{\searrow} & \downarrow \beta \\ A & \longrightarrow & A^+/\sigma \cong T \end{array}$$

Suppose that $(u, v) \in \rho$. Then there exists an elementary sequence $u = w_1, w_2, \ldots, w_n = v$ where $w_i = p_i u_i q_i$ and $w_{i+1} = p_i v_i q_i$ where $p_i, q_i \in A^*$ and $(u_i, v_i) \in R$ or $(v_i, u_i) \in R$ for all i. If $u_i = a_1 \cdots a_k$ and $v_i = b_1 \cdots b_l$ where $a_1, \ldots, a_k, b_1, \ldots, b_l \in A$, then since $\gamma : A^+ \longrightarrow T$ is a homomorphism extending $f : A \longrightarrow T$ and by the assumption in the theorem, it follows that

$$(u_i)\gamma = (a_1)\gamma \cdots (a_k)\gamma = (a_1)f \cdots (a_k)f = (b_1)f \cdots (b_l)f = (b_1)\gamma \cdots (b_l)\gamma = (v_i)\gamma$$

and so $(u_i, v_i) \in \ker(\gamma) = \sigma$ for all i. Since σ is a congruence, it follows that $(u, v) \in \sigma$. □

We refer to semigroup T satisfying the hypothesis of Theorem 6.4 as ***satisfying the relations*** R ***defining*** S.

Theorem 6.5. *Every semigroup is defined by a presentation.*

Proof. Let S be any semigroup, and let $A_S = \{a_s : s \in S\}$ be any set of cardinality equal to that of S. Then we will show that S is defined by the presentation:

$$\langle A \,|\, a_s a_t = a_{st} \ (s, t \in S) \rangle.$$

Suppose that T is the semigroup defined by this presentation. Clearly, the generators S of S satisfy the relations in this presentation, and so S is a homomorphic image of T (Theorem 6.4). On the other hand, any word $w \in A^+$ can be reduced to a single letter by using the defining relations. It follows that the function $\phi : S \longrightarrow T$ defined by $(s)\phi = a_s$ is an isomorphism. □

Example 6.6. (**Free semilattices**). Let S be the semigroup defined by the presentation

$$\langle a_1, a_2, \ldots, a_n \,|\, a_i^2 = a_i, \ a_i a_j = a_j a_i \ (1 \leq i, j \leq n) \rangle. \tag{6.1}$$

What can we say about S?

By definition, every element $s \in S$ is a product of elements $a_1, a_2, \ldots, a_n$. Using the relations $a_i a_j = a_j a_i$ we deduce that

$$s = a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n}$$

where $\epsilon_1, \epsilon_2, \ldots, \epsilon_n \geq 0$ and at least one $\epsilon_i \neq 0$. Then applying the relations $a_i^2 = a_i$, $\epsilon_i \in \{0, 1\}$ and so $|S| \leq 2^n - 1$.

On the other hand, if $a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n} = a_1^{\delta_1} a_2^{\delta_2} \cdots a_n^{\delta_n}$, then $\epsilon_1 = \delta_1$, $\epsilon_2 = \delta_2, \ldots, \epsilon_n = \delta_n$ since none of the relations can change the number of occurrences of a_i from 0 to 1 or ***vice versa***. Thus $|S| = 2^n - 1$.

For example, if $n = 2$, then $S = \{a_1, a_2, a_1 a_2\}$.

Let X be any set. Then the set $\mathcal{P}(X) \setminus \{\varnothing\}$ of non-empty subsets of X is a semigroup with operation $\cup$; denoted SL_X (for reasons that may become apparent later). The set $\{\{x\} : x \in X\}$ is a generating set for SL_X.

If $n \in \mathbb{N}$, then we write SL_n to denote $SL_{\{1,\ldots,n\}}$. Let $\{a_1, a_2, \ldots, a_n\}$ be the generators of the semigroup S defined by the presentation in (6.1) and define $f : \{a_1, a_2, \ldots, a_n\} \longrightarrow \{\{x\} : x \in \{1, \ldots, n\}\}$ by $(a_i)f = \{i\}$. Then

$$(a_i)f \ (a_i)f = \{i\} \cup \{i\} = \{i\} = (a_i)f \quad \text{and} \quad (a_i)f \ (a_j)f = \{i, j\} = (a_j)f \ (a_i)f$$

for all i, j. Hence SL_n satisfies the relations $a_i^2 = a_i$ and $a_i a_j = a_j a_i$ from the presentation defining S. It follows, from Theorem 6.4, that SL_n is a homomorphic image of S. On the other hand, $|SL_n| = 2^n - 1 = |S|$ and so S is isomorphic to SL_n.

Definition 6.7. (**Monoid presentation**). The monoid defined by the monoid presentation

$$\langle A \,|\, R \rangle$$

is the semigroup defined by the semigroup presentation

$$\langle A, e \,|\, R, \ ae = ea = a \ (a \in A \cup \{e\}) \rangle$$

The reason for introducing monoid presentations is that it is convenient to miss out the implicit relations $ae = ea = a$ for all $a \in A \cup \{e\}$.

If $\langle A \,|\, R\rangle$ is a monoid presentation, then the monoid defined by it can be defined in terms of a quotient of a free monoid. Replace the free semigroup A^+ by the free monoid A^* in all of the preceding statements about semigroup presentations to achieve this definition.

Example 6.8. (**Bicyclic monoid**). The monoid presentation

$$\langle b, c \,|\, bc = 1\rangle$$

defines a monoid called the ***bicyclic monoid*** denoted B. From the presentation defining B it is clear that $B = \{c^i b^j : i, j \in \mathbb{N} \cup \{0\}\}$.

Define $\mathbf{b}, \mathbf{c} \in T_{\mathbb{N}}$ by

$$(x)\mathbf{b} = x + 1 \quad \text{and} \quad (x)\mathbf{c} = \begin{cases} \min \mathbb{N} & \text{if } x = \min \mathbb{N} \\ x - 1 & \text{if } x \neq \min \mathbb{N} \end{cases}.$$

Since $\mathbf{b} \circ \mathbf{c} = 1_{\mathbb{N}}$, the identity function on $\mathbb{N}$, the semigroup $\langle \mathbf{b}, \mathbf{c}\rangle$ satisfies the relation $bc = 1$ of the presentation defining B. Hence, by Theorem 6.4, $\phi : B \longrightarrow \langle \mathbf{b}, \mathbf{c}\rangle$ defined by $(c^i b^j)\phi = \mathbf{c}^i \mathbf{b}^j$.

Suppose that $c^i b^j = c^k b^l$ for some $i, j, k, l \in \mathbb{N}$. Then $\mathbf{c}^i \mathbf{b}^j = (c^i b^j)\phi = (c^k b^l)\phi = \mathbf{c}^k \mathbf{b}^l$ and so

$$\min \mathbb{N} + j = (\min \mathbb{N})\mathbf{c}^i \mathbf{b}^j = (\min \mathbb{N})\mathbf{c}^k \mathbf{b}^l = \min \mathbb{N} + l,$$

which implies that $j = l$.

Also, for a sufficiently large n,

$$n - i + j = (n)\mathbf{c}^i \mathbf{b}^j = (n)\mathbf{c}^k \mathbf{b}^l = n - k + l,$$

and so $i = k$. It follows that in B, $c^i b^j = c^k b^l$ if and only if $i = k$ and $j = l$.

Remember: In semigroup presentations there are no inverses of generators and no identity in the relations. In monoid presentations there are no inverses of generators in the relations. In group presentations you can have the lot!

7. Finitely presented semigroups

We have seen that presentations are means of defining semigroups. Their particular advantage over, say, transformation semigroups, is that one can define infinite semigroups by finite means (see Examples 6.3 and 6.8). However, although every semigroup has a presentation (because it is isomorphic to a quotient of a free semigroup; see Theorem 5.3), not every semigroup can be defined by a finite presentation (i.e. one in which both A and R are finite). Those that can are called ***finitely presented*** semigroups.

Example 7.1. The free semigroup A^+ over a finite set A is finitely presented; see Example 6.3.

Every finite semigroup S is finitely presented, this follows from the proof of Theorem 6.5.

Theorem 7.2. *Let S be a finitely presented semigroup with a finite presentation $\langle A|R\rangle$, and let $\langle B \,|\, Q\rangle$ be any presentation for S with B finite and Q infinite. Then there exists a finite subset $Q_0 \subseteq Q$ such that $\langle B \,|\, Q_0\rangle$ is also a presentation for S.*

Proof. The proof of this theorem is quite technical, and we omit the details. However, the basic idea is simple. Since Q is a set of defining relations for S, any relation that holds in S is a consequence of Q. In particular, every relation from R is a consequence of relations in Q. If $(u, v) \in R$, is a consequence of Q, then there is an elementary sequence between u and v over Q. Elementary sequences are by definition finite in length, and hence only use finitely many of the relations from Q. Since R is finite, it follows that there exists a finite subset $Q_0 \subseteq Q$ such that every relation from R is a consequence of Q_0. Since the relations R define S, and they are consequences of Q_0 it follows that Q_0 also defines S, and this proofs the theorem.

The technicality of the proof arises from the fact that R and Q may be defined over different alphabets, and so it is necessary to find a way of translating between the two alphabets. □

An immediate consequence of Theorem 7.2 the existence of a finite presentation for a semigroup is independent of the choice of a finite generating set.

Corollary 7.3. *Let S be a finitely presented semigroup. If B is any finite generating set for S, then there exists a finite $Q_0 \subseteq B^+ \times B^+$ such that the presentation $\langle B \,|\, Q_0\rangle$ defines S.*

Proof. It is enough to show that S has a (possibly infinite) presentation in terms of B, and to then apply Theorem 7.2. Since $\langle B \rangle = S$, there exists a surjective homomorphism $\phi : B^+ \longrightarrow S$, and so, for example, the presentation $\langle B \,|\, \ker(\phi) \rangle$ defines S. □

It is clear that a finitely presented semigroup must be finitely generated. However, not every finitely generated semigroup is finitely presented, as the following example shows.

Example 7.4. Let S be the semigroup defined by the presentation

$$\langle a, b \,|\, ab^i a = aba \quad (i \geq 2) \rangle.$$

We will prove that S is not finitely presented. Assume to the contrary that S is finitely presented. Then, by Theorem 7.2, there exists $m \geq 2$ such that S is defined by the presentation

$$\langle a, b \,|\, ab^i a = aba \quad (2 \leq i \leq m) \rangle;$$

we denote the set of defining relations in this presentation Q_0. In S, $ab^{m+1}a = aba$ and so must be a consequence of Q_0. However, it is not possible to apply any relation from Q_0 to $ab^{m+1}a$, a contradiction.

Let S be a finitely presented semigroup, and let $\langle A \,|\, R \rangle$ be a finite presentation for S. In principle, all the information about S is encoded in its presentation, and it would be natural to ask if there is a systematic way of recovering certain specific pieces of information about S. For example, we might ask if there is an algorithm that can check if two words $w_1, w_2 \in A^+$ represent the same element of S or not. It turns out that such an algorithm often does not exist; this is referred to as the undecidability of the word problem. Also, there is no general algorithm which determines whether S is trivial, or finite, or a group, or indeed to answer any sensible question about S.

8. Ideals

The concept of ideals came to semigroup theory from rings, and the notion is analogous.

Definition 8.1. (**Ideal**). Let S be a semigroup and let $I \subseteq S$. Then I is a a ***left ideal*** if $si \in I$ for all $s \in S$ and for all $i \in I$. ***Right ideals*** are defined analogously, and a ***(two-sided) ideal*** is both a left and a right ideal. An ideal is called ***proper*** if $I \neq S$.

Example 8.2. Every semigroup S is an ideal of itself. In a semigroup S with a zero element 0 the set $\{0\}$ is an ideal. If G is a group, then I is an ideal of G if and only if $I = G$ or $I = \varnothing$.

Let

$$K(n, r) = \{f \in T_n : \operatorname{rank}(f) = |\operatorname{im}(f)| \leq r\}.$$

In Problem **1-6** we showed that

$$\operatorname{rank}(fg) \leq \min\{\operatorname{rank}(f), \operatorname{rank}(g)\}$$

and so $K(n, r)$ is a (two-sided) ideal of T_n for all $r = 1, 2, \ldots, n$.

As with subsemigroups, we can generate ideals by subsets.

Theorem 8.3. *Let S be a semigroup, and let X be any non-empty subset of S. Then*

(i) $S^1X = \{sx : s \in S^1,\ x \in X\}$ *is the least left ideal of S containing X;*
(ii) $XS^1 = \{xs : s \in S^1,\ x \in X\}$ *is the least right ideal of S containing X;*
(iii) $S^1XS^1 = \{sxt : s, t \in S^1,\ x \in X\}$ *is the least (two-sided) ideal of S containing X.*

Proof. We prove part (i); parts (ii) and (iii) are proved using analogous arguments.

If $sx \in S^1X$ and $t \in S$ are arbitrary, then $t(sx) = (ts)x \in S^1X$. Hence S^1X is a left ideal. Since $1 \in S^1$, $X \subseteq S^1X$ and so S^1X is a left ideal containing X.

If I is any left ideal containing X, then $sx \in I$ for all $s \in S^1$ and for all $x \in X \subseteq I$. Therefore $S^1X \subseteq I$, and S^1X is the least ideal containing X. □

Similar to rings, the notion of ideals for semigroups give rise to certain quotients of semigroups.

Theorem 8.4. ***(Rees congruence).*** *Let S be a semigroup and let I be an ideal of S. Then the relation*

$$\rho_I = (I \times I) \cup \Delta_S = \{(x, y) \in S \times S : x, y \in I \text{ or } x = y\}$$

is called the ***Rees congruence*** *of I.*

Proof. Let $(x, y) \in \rho_I$ and let $s \in S$. If $x = y$, then $sx = sy$ and $xs = ys$ and so $(sx, sy), (xs, ys) \in \rho_I$. If $x \neq y$, then $x, y \in I$ and so $sx, sy, xs, ys \in I$ since I is an ideal. It follows that $(sx, sy), (xs, ys) \in \rho_I$ and so ρ_I is a congruence on S. □

It is straightforward to verify that ρ_I is an equivalence relation with classes I and $\{s\}$ for all $s \in S \setminus I$. Hence the number of classes in ρ_I is $|S \setminus I| + 1$.

Definition 8.5. (**Rees quotient**). If I is an ideal of a semigroup S, then the quotient semigroup S/ρ_I of S by the Rees congruence ρ_I is called the ***Rees quotient of*** S ***by*** I and it is denoted by S/I.

One very important difference between semigroups and rings is that in a semigroup not every congruence is induced by an ideal. For example, a group has no non-empty ideals, but has a non-trivial congruence for every non-trivial normal subgroup.

9. Green's Relations

Green's relations describe the fundamental structure of a semigroup. They partition the elements and allow us to analyse the way they interact. There are 5 Green's relations on a semigroup, which we define and discuss in the 4 subsections of this section.

9.1. Green's $\mathscr{L}$- and $\mathscr{R}$-relations.

Definition 9.1. (**Green's $\mathscr{L}$- and $\mathscr{R}$-relations**). Let S be a semigroup and let $x, y \in S$. Then x and y are ***$\mathscr{L}$-related*** if they generate the same principal left ideal, i.e. if $S^1x = S^1y$. Green's $\mathscr{R}$-relation is defined analogously.

Lemma 9.2. *Let S be a semigroup and let $x, y \in S$. Then the following hold:*

(i) *$x \mathscr{L} y$ if and only if there exist $u, v \in S^1$ such that $ux = y$ and $vy = x$;*
(ii) *$x \mathscr{R} y$ if and only if there exist $u, v \in S^1$ such that $xu = y$ and $yv = x$;*

Proof. We prove part (i), the proof of part (ii) is dual.

($\Rightarrow$) If $x \mathscr{L} y$, then $S^1x = S^1y$ and so $x = 1 \cdot x \in S^1y$ and $y = 1 \cdot y \in S^1x$. Hence there exist $u, v \in S^1$ such that $x = uy$ and $y = vx$, as required.

($\Leftarrow$) If there exist $u, v \in S^1$ such that $uy = x$ and $vx = y$, then $x \in S^1y$ and $y \in S^1x$. □

Example 9.3. Are the elements

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{pmatrix}, \quad g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 3 \end{pmatrix} \in T_4$$

$\mathscr{L}$-related or $\mathscr{R}$-related? Suppose that

$$u = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 4 \end{pmatrix} \quad v = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 3 & 1 & 2 \end{pmatrix}.$$

Then

$$uf = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 3 \end{pmatrix} = g$$

and

$$vg = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{pmatrix} = f$$

and so $f \mathscr{L} g$.

On the other hand, $(1)fh = (2)fh$ for all $h \in T_4$ but $(1)g = 2 \neq 3 = (2)g$ and so $fh \neq g$ for all $h \in T_4$. Hence $(f, g) \notin \mathscr{R}$.

Theorem 9.4. *Let $f, g \in T_n$. Then*

(i) *$f \mathscr{L} g$ if and only if $\operatorname{im}(f) = \operatorname{im}(g)$;*
(ii) *$f \mathscr{R} g$ if and only if $\ker(f) = \ker(g)$.*

Proof. **(i).** ($\Rightarrow$) If $f \mathscr{L} g$, then there exist $u, v \in T_n$ such that

$$vf = g \quad \text{and} \quad ug = f.$$

Hence $\operatorname{im}(f) = \operatorname{im}(ug) \subseteq \operatorname{im}(g)$ and similarly $\operatorname{im}(g) \subseteq \operatorname{im}(f)$.

($\Leftarrow$) Suppose that $f, g \in T_n$ and $\operatorname{im}(f) = \operatorname{im}(g)$. We want to find $u, v \in T_n$ such that $vf = g$ and $ug = f$. For every $i \in \operatorname{im}(f) = \operatorname{im}(g)$, we choose $a_i, b_i \in \{1, \ldots, n\}$ such that $a_i f = b_i g = i$ and define transformations $u, v \in T_n$ by $(x)u = a_i$ if $(x)g = i$, and $(x)v = b_i$ if $(x)f = i$. If $x \in \{1, \ldots, n\}$ is arbitrary and $(x)g = i$, then

$$(x)uf = (a_i)f = i = (x)g.$$

Similarly, $(x)vg = (x)f$ for all $x \in \{1, \ldots, n\}$, and so $uf = g$ and $vg = f$. Thus $f \mathscr{L} g$.

(ii). ($\Rightarrow$) If $f\mathscr{R}g$, then there exists $u, v \in T_n$ such that

$$fu = g \text{ and } gv = f.$$

Hence if $(x, y) \in \ker(f)$, then $(x)f = (y)f$ and so

$$(x)g = (x)fu = (y)fu = (y)g.$$

It follows that $(x, y) \in \ker(g)$ and so $\ker(f) \subseteq \ker(g)$. By symmetry, $\ker(g) \subseteq \ker(f)$.

($\Leftarrow$) We want to find $u, v \in T_n$ such that $fu = g$ and $gv = f$. Let $\{k_1, k_2, \ldots, k_m\}$ be a set containing exactly one element in each equivalence class of $\ker(f) = \ker(g)$. Let $u, v \in T_n$ be any mappings such that

$$(k_i f)u = (k_i)g$$

$$(k_i g)v = (k_i)f.$$

Now, if $x \in \{1, \ldots, n\}$ is arbitrary, then $(x, k_i) \in \ker(f) = \ker(g)$ for some k_i. Thus

$$(x)fu = (k_i)fu = (k_i)g = (x)g,$$

and likewise

$$(x)gv = (k_i)f = (x)f,$$

as required. □

Since $S^1x = S^1x$, it follows that $x\mathscr{L}x$ for all $x \in S$ and so $\mathscr{L}$ is reflexive. If $x\mathscr{L}y$, then $y\mathscr{L}x$ and so $\mathscr{L}$ is symmetric. If $x\mathscr{L}y$ and $y\mathscr{L}z$, then $S^1x = S^1y$ and $S^1y = S^1z$ and so $x\mathscr{L}z$. It follows that $\mathscr{L}$ and, likewise, $\mathscr{R}$ are equivalence relations. The equivalence class of an element $x \in S$ with respect to $\mathscr{L}$ is called its $\mathscr{L}$-***class***, and is usually denoted as L_x, rather than $x/\mathscr{L}$. The $\mathscr{R}$-***class*** of x is defined analogously, and is denoted by R_x.

Theorem 9.5. *The relation $\mathscr{L}$ is a right congruence and the relation $\mathscr{R}$ is a left congruence.*

Proof. We want to show that for all if $(x, y) \in \mathscr{L}$, then $(sx, sy) \in \mathscr{L}$ for all $s \in S$. Let $(x, y) \in \mathscr{L}$. Then there exists $u, v \in S^1$ such that $y = ux$ and $x = vy$. So, if $s \in S$, then

$$uxs = ys \text{ and } vys = xs.$$

Thus $(xs, ys) \in \mathscr{L}$ and so $\mathscr{L}$ is a right congruence.

The proof for $\mathscr{R}$ is analogous. □

In general, neither $\mathscr{L}$ nor $\mathscr{R}$ is a two-sided congruence.

Example 9.6. ($\mathscr{L}$ **is not always a left congruence**). If

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 1 & 3 \end{pmatrix} \text{ and } g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 2 & 1 \end{pmatrix},$$

then $\text{im}(f) = \text{im}(g) = \{1, 2, 3\}$. It follows, by Theorem 9.4, that $f\mathscr{L}g$. But if

$$h = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & 2 & 2 \end{pmatrix},$$

then $hf = h$ and

$$hg = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 3 & 3 & 3 \end{pmatrix}.$$

Therefore $\text{im}(hf) = \{2\} \neq \{3\} = \text{im}(hg)$ and so $(hf, hg) \notin \mathscr{L}$ and $\mathscr{L}$ is not a left congruence.

There is a strong connection between the relations $\mathscr{L}$ and $\mathscr{R}$ and the left and right Cayley graphs of the semigroup. Before we can state this connection precisely, we require the following notions regarding digraphs.

Definition 9.7. (**Path in a digraph**). Let Γ be a digraph. A ***path*** in Γ is a sequence $(v_1, v_2, \ldots, v_m)$ of vertices such that there is an edge from v_i to v_{i+1} for all $i = 1, \ldots, m-1$.

Example 9.8. If Γ is the right Cayley digraph shown in Figure 1, then (a, ab, ba) is a path from a to ba but there is no path from ba to a.

Definition 9.9. (**Strongly connected component**). The ***strongly connected component*** of a vertex v in Γ is the set of all vertices u such that there is a path from u to v, and a path from v to u.

It is straightforward to verify that "there is a path from u to v and a path from v to u" defines an equivalence relation on the vertices of a digraph. Hence the strongly connected components of Γ partition its vertices.

Theorem 9.10. *Let S be a semigroup. Then*

(i) *$x\mathscr{R}y$ if and only if x and y are in the same strongly connected component of the right Cayley graph of S (with respect to any generating set);*
(ii) *$x\mathscr{L}y$ if and only if x and y are in the same strongly connected component of the left Cayley graph of S (with respect to any generating set).*

Proof. We prove the statement for $\mathscr{R}$.

($\Rightarrow$) Since $x\mathscr{R}y$, there exists $s, t \in S^1$ such that $xs = y$ and $yt = x$. If A be a generating set for S, then $s = a_1a_2 \dots a_m$ and $t = b_1b_2 \dots b_n$ for some $a_1, a_2, \dots, a_m, b_1, b_2, \dots, b_n \in A$. The sequences

$$x, xa_1, xa_1a_2, \dots, xa_1a_2 \dots a_m = y$$

and

$$y, yb_1, yb_1b_2, \dots, yb_1b_2 \dots b_n = x$$

are paths between x and y in the right Cayley graph of S with respect to the generating set A.

($\Leftarrow$) Let $x = u_1, u_2, \dots, u_m = y$ and $y = v_1, v_2, \dots, v_n = x$ be paths between x and y in the right Cayley graph of S with respect to some generating set A. Let the elements of A labelling the edge from u_i to u_{i+1} be denoted by a_i ($i = 1, \dots, m-1$). Then

$$y = u_m = u_{m-1}a_{m-1} = u_{m-2}a_{m-2}a_{m-1} = \dots = u_1a_1 \dots a_{m-2}a_{m-1} = xs,$$

where $s = a_1 \dots a_{m-1}$. Similarly, by using the other path, we obtain $t \in S$ such that $x = yt$. Thus $x\mathscr{R}y$, as required. □

Example 9.11. If Γ is the right Cayley digraph shown in Figure 1, then the strongly connected components of Γ are $\{a\}$, $\{ab, ab^2\}$, $\{ba, bab\}$, and $\{b, b^2\}$. By Theorem 9.10, it follows that these are also the $\mathscr{R}$-classes of the semigroup from Example 2.10.

9.2. **Green's $\mathscr{H}$-relation.**

Definition 9.12. (**Green's $\mathscr{H}$-relation**). Let S be a semigroup. Then $s, t \in S$ are ***$\mathscr{H}$-related*** if and only if they are both $\mathscr{L}$-related and $\mathscr{R}$-related. In other words, $\mathscr{H} = \mathscr{L} \cap \mathscr{R}$.

By Proposition 5.7(i), since $\mathscr{H}$ is the intersection of equivalence relations, it is an equivalence relation. The equivalence class of $s \in S$ is called its ***$\mathscr{H}$-class*** and is denoted by H_s.

Example 9.13. Let $f, g \in T_n$. Then by Theorem 9.4, $f\mathscr{H}g$ if and only if $\operatorname{im}(f) = \operatorname{im}(g)$ and $\ker(f) = \ker(g)$. The equivalence classes of $\ker(f)$ are in one-one correspondence with the elements of $\operatorname{im}(f)$. Therefore, the number of elements in the $\mathscr{H}$-class of α is equal to the number of bijections from the set of equivalence classes of $\ker(f)$ onto $\operatorname{im}(f)$. It is easy to see that this number is $r!$, where $r = |\operatorname{im}(f)| = |\ker(f)|$.

Of course, $r!$ is also the order of the symmetric group S_r, and we shall see later that this is not a coincidence.

9.3. **Green's $\mathscr{D}$-relation.** The fourth of Green's relations is the following.

Definition 9.14. (**Green's $\mathscr{D}$-relation**). ***Green's $\mathscr{D}$-relation*** on a semigroup is the composition of its $\mathscr{L}$- and $\mathscr{R}$-relations, that is, $\mathscr{D} = \mathscr{L} \circ \mathscr{R}$.

From the above definition, it is not clear that $\mathscr{D}$ is an equivalence relation.

Theorem 9.15. *$\mathscr{L}$ and $\mathscr{R}$ commute, that is, $\mathscr{L} \circ \mathscr{R} = \mathscr{R} \circ \mathscr{L}$.*

Proof. We will show that $\mathscr{L} \circ \mathscr{R} \subseteq \mathscr{R} \circ \mathscr{L}$ and $\mathscr{R} \circ \mathscr{L} \subseteq \mathscr{L} \circ \mathscr{R}$. In fact, we will only show that $\mathscr{L} \circ \mathscr{R} \subseteq \mathscr{R} \circ \mathscr{L}$, that $\mathscr{R} \circ \mathscr{L} \subseteq \mathscr{L} \circ \mathscr{R}$ follows by symmetry.

Let $(x, y) \in \mathscr{L} \circ \mathscr{R}$. Then, by definition, there exists $z \in S$ such that $(x, z) \in \mathscr{L}$ and $(z, y) \in \mathscr{R}$. Hence there exist $s_1, s_2, t_1, t_2 \in S^1$ such that

$$z = s_1x, \quad x = s_2z, \quad y = zt_1, \quad z = yt_2.$$

If $u = xt_1$, then

$$ut_2 = xt_1t_2 = s_2zt_1t_2 = s_2yt_2 = s_2z = x$$

and so $x\mathscr{R}u$. On the other hand,

$$s_1u = s_1xt_1 = zt_1 = y$$
$$s_2y = s_2zt_1 = xt_1 = u$$

and so $y\mathscr{L}u$. It follows that $(x,y) \in \mathscr{R} \circ \mathscr{L}$ and so $\mathscr{L} \circ \mathscr{R} \subseteq \mathscr{R} \circ \mathscr{L}$. □

Corollary 9.16. *$\mathscr{D}$ is an equivalence relation.*

Proof. Recall Problem **3-5**:

> Let α, β be equivalence relations on a set X. Prove that $\alpha \circ \beta$ is an equivalence relation if and only if $\alpha \circ \beta = \beta \circ \alpha$.

It follows immediately from Theorem 9.15 that $\mathscr{D} = \mathscr{L} \circ \mathscr{R} = \mathscr{R} \circ \mathscr{L}$ is an equivalence relation. □

Theorem 9.17. *Let $f, g \in T_n$. Then $f\mathscr{D}g$ if and only if* $\operatorname{rank}(f) = \operatorname{rank}(g)$.

Proof. ($\Rightarrow$) Assume that $f\mathscr{D}g$. From the definition of $\mathscr{D}$, it follows that there exists $h \in T_n$ such that $f\mathscr{L}h$ and $h\mathscr{R}g$. By Theorem 9.4, $\operatorname{im}(f) = \operatorname{im}(h)$ and $\ker(h) = \ker(g)$, and in particular $\operatorname{rank}(f) = \operatorname{rank}(h) = \operatorname{rank}(g)$.

($\Leftarrow$) Assume that $\operatorname{rank}(f) = \operatorname{rank}(g)$. Then $|\operatorname{im}(f)| = |\operatorname{im}(g)|$ and $\ker(f)$ and $\ker(g)$ have the same number of kernel classes. If $K_1, \ldots, K_r$ are the kernel classes of $\ker(g)$ and if $\operatorname{im}(f) = \{i_1, \ldots, i_r\}$, then define a mapping k by

$$(x)k = i_j \ \ \forall\, x \in K_j.$$

We now have $\operatorname{im}(k) = \operatorname{im}(f)$, so that $f\mathscr{L}k$. Also $\ker(k) = \ker(g)$, so that $k\mathscr{R}g$. Therefore $(f,g) \in \mathscr{L} \circ \mathscr{R} = \mathscr{D}$. □

9.4. **Green's $\mathscr{J}$-relation.** The last of Green's relations is a two-sided analogue of the relations $\mathscr{L}$ and $\mathscr{R}$.

Definition 9.18. Let S be a semigroup. Then ***Green's $\mathscr{J}$-relation*** is defined such that $(x,y) \in \mathscr{J}$ if $S^1xS^1 = S^1yS^1$. Equivalently, $(x,y) \in \mathscr{J}$ if and only if there exist $s,t,u,v \in S^1$ such that $sxt = y$ and $uyv = x$.

It is easy to see that the following inclusions hold for an arbitrary semigroup hold:

$$\mathscr{H} \subseteq \mathscr{L},\ \mathscr{H} \subseteq \mathscr{R},\ \mathscr{L} \subseteq \mathscr{D},\ \mathscr{R} \subseteq \mathscr{D}. \tag{9.1}$$

Lemma 9.19. *$\mathscr{D} \subseteq \mathscr{J}$.*

Proof. Let $(x,y) \in \mathscr{D}$. Then there exists $z \in S$ such that $x\mathscr{L}z\mathscr{R}y$. Hence $S^1x = S^1z$ and $zS^1 = yS^1$. It follows that $S^1xS^1 = S^1zS^1 = S^1yS^1$ and so $(x,y) \in \mathscr{J}$. □

The relationships between the different Green's relations can be conveniently illustrated by the Hasse diagram in Figure 3. The example of the full transformation semigroup T_n shows that the four inclusions in (9.1) are proper. In Problem **6-3** we saw or will see that, in general, $\mathscr{D} \neq \mathscr{J}$.

In some sense, the most general structural information about a semigroup can be found from the coarsest relation $\mathscr{J}$. It might also be argued that the definition of $\mathscr{J}$ is more natural than that of $\mathscr{D}$. However, it is $\mathscr{D}$ that is the more interesting and more strongly related to the structure of the semigroup.

Definition 9.20. A semigroup S is said to be ***periodic*** if $\langle x \rangle$ is finite for all $x \in S$.

Note that every finite semigroup is periodic. Can you think of an example of an infinite periodic semigroup?

Theorem 9.21. *If S periodic (and, in particular, if S is finite), then $\mathscr{J} = \mathscr{D}$.*

Proof. We already know that $\mathscr{D} \subseteq \mathscr{J}$, and so it suffices to prove that $\mathscr{J} \subseteq \mathscr{D}$. Let $(x,y) \in \mathscr{J}$. (We want to show that $(x,y) \in \mathscr{D}$.) Then there exist $a,b,c,d \in S^1$ such that

$$x = ayb \text{ and } y = cxd.$$

We will start by proving that $x\mathscr{L}cx$. We know that

$$x = ayb = (ac)x(db) = (ac)^2x(db)^2 = \cdots = (ac)^ix(db)^i = \cdots.$$

Since S is periodic, $\langle ac \rangle$ is finite and so contains an idempotent (by Problem **2-10**), say $(ac)^i$. Then

$$x = (ac)^ix(db)^i = (ac)^i(ac)^ix(db)^i = (ac)^ix = (ac)^{i-1}acx,$$

Thus $c \cdot x = cx$ and $[(ac)^{i-1}a] \cdot cx = x$ and so $x\mathscr{L}cx$.

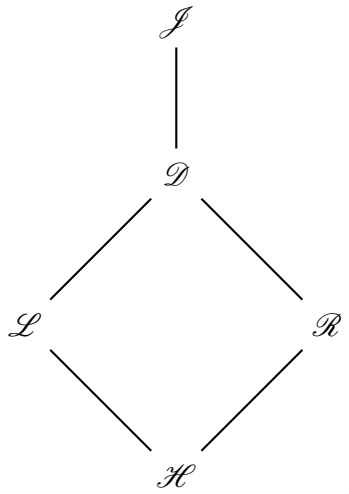

FIGURE 3. Hassé diagram of Green's equivalences.

We will now prove that $cx\mathscr{R}y$. As above

$$y = cxd = (ca)y(bd) = (ca)^2y(bd)^2 = \ldots = (ca)^jy(bd)^j = \ldots$$

and so there exists a power of bd that is an idempotent, say $(bd)^j$. Thus

$$\begin{aligned} cx = c(ac)^{j+1}x(db)^{j+1} = (ca)^{j+1}cxd(bd)^jb = (ca)^{j+1}y(bd)^jb \\ = (ca)^{j+1}y(bd)^{2j}b = (ca)^{j+1}y(bd)^{j+1}(bd)^{j-1}b = y(bd)^{j-1}b. \end{aligned}$$

Hence $cx = y \cdot (bd)^{j-1}b$ and $y = cx \cdot d$. It follows that $cx\mathscr{R}y$.

To conclude, $x\mathscr{L}cx\mathscr{R}y$ and so $x\mathscr{D}y$. □

10. Green's Lemma

Let S be a semigroup. Then $\mathscr{D} = \mathscr{L} \circ \mathscr{R} = \mathscr{R} \circ \mathscr{L}$, $\mathscr{D}$ is an equivalence relation, and $\mathscr{L}, \mathscr{R} \subseteq \mathscr{D}$. Let D be a $\mathscr{D}$-class of S and let $x \in D$ be arbitrary. Then $(x, y) \in \mathscr{L} \subseteq \mathscr{D}$ for all $y \in L_x$ and so $L_x \subseteq D$. It follows that the $\mathscr{L}$-classes of a $\mathscr{D}$-class D partition D, and so too do the $\mathscr{R}$-classes.

Lemma 10.1. *Let $a, b \in S$ be such that $a\mathscr{D}b$. Then $L_a \cap R_b \neq \varnothing$ and $L_a \cap R_b \subseteq D_a = D_b$ is an $\mathscr{H}$-class of S.*

Proof. From the definition of $\mathscr{D}$, it follows that there exists c such that $(a, c) \in \mathscr{L}$ and $(c, b) \in \mathscr{R}$. Hence $c \in L_a \cap R_b$ and $L_a \cap R_b \neq \varnothing$.

If $d \in H_c$, then $d\mathscr{L}c\mathscr{L}a$ and $d\mathscr{R}c\mathscr{R}b$ and so $H_c \subseteq L_a \cap R_b$. On the other hand, if $d \in L_a \cap R_b$, then □

The main theorem in this section, known as Green's Lemma, considers $\mathscr{L}$-, $\mathscr{R}$-, and $\mathscr{H}$-classes in a single $\mathscr{D}$-class, and shows that they have a great degree of uniformity. One formulation of this theorem is given below.

Theorem 10.2 (Green's Lemma)**.** *Let S be a semigroup and let $a, b \in S$ be such that $a\mathscr{R}b$. If $s, t \in S^1$ are such that $as = b$ and $bt = a$, then $\rho_s : L_a \longrightarrow L_b$ and $\rho_t : L_b \longrightarrow L_a$ defined by*

$$(x)\rho_s = xs \qquad \text{and} \qquad (y)\rho_t = yt$$

are mutually inverse bijections. Furthermore, if $x \in L_a$, then $xs \in R_x \cap L_b$.

Proof. We start by proving that $(x)\rho_s \in L_b$ and $(y)\rho_t \in L_a$ for all $x \in L_a$ and $y \in L_b$. If $x \in L_a$, then $(x, a) \in \mathscr{L}$ and so $(xs, as) = (xs, b) \in \mathscr{L}$ since, by Theorem 9.5, $\mathscr{L}$ is a right congruence. In other words, $(x)\rho_s = xs \in L_b$. Similarly, $(y)\rho_t \in L_a$ for all $y \in L_b$.

To show that ρ_s and ρ_t are mutually inverse bijections, it suffices to to prove that $x\rho_s\rho_t = x$ for all $x \in L_a$ and that $y\rho_t\rho_s = y$ for all $y \in L_b$. Let $x \in L_a$ be arbitrary. Then there exists $u \in S^1$ such that $ua = x$ and so

$$x\rho_s\rho_t = xst = u(as)t = ubt = ua = x.$$

Similarly, $y\rho_t\rho_s = y$ for all $y \in L_b$.

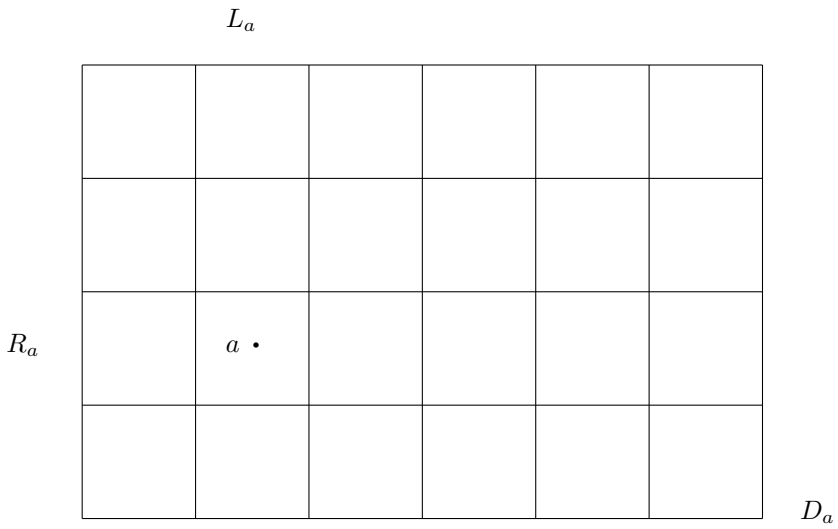

FIGURE 4. The egg box diagram of a $\mathscr{D}$-class

To conclude, if $x \in L_a$, then $xs = x\rho_s \in L_b$ from above and $xst = x\rho_s\rho_t = x$, which implies $xs \in R_x$. □

The following is the other handed version of Green's Lemma, the proof is similar to that of Theorem 10.2 and is omitted.

Theorem 10.3 (Green's Lemma)**.** *Let S be a semigroup and let $a, b \in S$ be such that $a\mathscr{L}b$. If $s, t \in S^1$ are such that $sa = b$ and $tb = a$, then $\lambda_s : R_a \longrightarrow R_b$ and $\lambda_t : R_b \longrightarrow R_a$ defined by*

$$(x)\lambda_s = sx \qquad \text{and} \qquad (y)\lambda_t = ty$$

are mutually inverse bijections. Furthermore, if $x \in R_a$, then $sx \in L_x \cap R_b$.

As a corollary of the existence of a bijection between any pair of $\mathscr{L}$-classes, or $\mathscr{R}$-classes, contained in a $\mathscr{D}$-class we obtain the following.

Corollary 10.4. *Let S be a semigroup and D be a $\mathscr{D}$-class in S. Then any two $\mathscr{L}$-classes in D have the same size, and any two $\mathscr{R}$-classes in D have the same size.*

The final result in this section shows that there is a strong connection between $\mathscr{H}$-classes and subgroups of a semigroup.

Lemma 10.5. *Let S be a semigroup and let $a, b \in S$. Then the following hold:*

(a) *if $a\mathscr{R}b$ and $s \in S^1$ is such that $as = b$, then $\rho_s : L_a \longrightarrow L_b$ defined by $(x)\rho_s = xs$ is a bijection and $(H_a)\rho_s = H_a s = H_b$;*
(b) *if $a\mathscr{L}b$ and $t \in S^1$ is such that $ta = b$, then $\lambda_t : R_a \longrightarrow R_b$ defined by $(x)\lambda_t = tx$ is a bijection and $(H_a)\lambda_t = tH_a = H_b$;*

Proof. It follows immediately from Green's Lemma that $\rho_s : L_a \longrightarrow L_b$ is a bijection. If $x \in H_a = L_a \cap R_a$, then $xs \in R_x \cap L_b = R_a \cap L_b = R_b \cap L_b = H_b$. Hence $(H_a)\rho_s \subseteq H_b$ and by symmetry, $(H_a)\rho_s = H_b$.
The proof of part (b) is similar. □

Corollary 10.6. *If S is a semigroup and $a\mathscr{D}b$ for some $a, b \in S$, then $|H_a| = |H_b|$.*

Proof. Since $a\mathscr{D}b$, there exists $c \in S$ such that $a\mathscr{R}c\mathscr{L}b$. If $s, t \in S^1$ are such that $as = c$ and $tc = b$, then, by Lemma 10.5(i), $\rho_s : L_a \longrightarrow L_c$ is a bijection and $(H_a)\rho_s = H_c$. Similarly, by Lemma 10.5(ii), $\lambda_t : R_c \longrightarrow R_b$ is a bijection and $(H_c)\lambda_t = H_b$. Therefore $\rho_s\lambda_t$ is a bijection from H_a to H_b, and so $|H_a| = |H_b|$. □

Green's Lemma allows us to represent a $\mathscr{D}$-class using an ***egg box*** diagram; see Figure 4. The rows represent $\mathscr{R}$-classes (and they are all of equal sizes), the columns represent $\mathscr{L}$-classes (again of equal sizes), and the little squares represent the $\mathscr{H}$-classes (of equal sizes).

Theorem 10.7. *If H is an $\mathscr{H}$-class of a semigroup S, then either $H^2 \cap H = \varnothing$ or H is a group.*

Proof. We will show that if $H^2 \cap H \neq \varnothing$, then H is a group. In Problem **1-8**, we showed that if H is a semigroup and $xH = Hx = H$ for all $x \in H$, then H is a group. The condition that S is a semigroup is, however, implied by the condition $xH = Hx = H$ for all $x \in H$. Hence it suffices to prove that if $H^2 \cap H \neq \varnothing$, then $xH = Hx = H$ for all $x \in H$.

Suppose that $H^2 \cap H \neq \varnothing$. Then there exist $a, b \in H$ such that $ab \in H$. It follows that $ab \mathscr{H} a$ and so, in particular, $a \mathscr{R} ab$ and $a \cdot b = ab$. Hence, by Lemma 10.5(i), $\rho_b : L_a \longrightarrow L_{ab}$ defined by $(y)\rho_b = yb$ is a bijection and $(H)\rho_b = (H_a)\rho_b = H_{ab} = H$. In particular, if $x \in H$ is arbitrary, then $xb = (x)\rho_b \in H$, or equivalently, $xb \mathscr{H} b$. It follows that $xb \mathscr{L} b$ and $x \cdot b = xb$. So, by Lemma 10.5(ii), $\lambda_x : R_b \longrightarrow R_{xb}$ defined by $(z)\lambda_x = xz$ is a bijection and $(H_b)\lambda_x = xH = H_{xb} = H$.

By symmetry, $Hx = H$ and so, by Problem **1-8**, H is a group. □

An $\mathscr{H}$-class that is a group is called a ***group $\mathscr{H}$-class***.

11. The structure of regular $\mathscr{D}$-classes

Definition 11.1. (**Regular element/semigroup.**). Let x be an element of a semigroup S. Then x is ***regular*** if there exists $y \in S$ such that $xyx = x$.

If every element of S is regular, then S is said to be a ***regular semigroup***.

Example 11.2. Let $f \in T_X$ be arbitrary. We will prove that f, and hence T_X, is regular. For each $x \in \mathrm{im}(f)$ choose $t_x \in X$ such that $(t_x)f = x$. Also choose an arbitrary $x_0 \in X$. Define a new mapping g by

$$(x)g = \begin{cases} t_x & \text{if } x \in \mathrm{im}(f) \\ x_0 & \text{if } x \notin \mathrm{im}(f). \end{cases}$$

Then

$$(xf)gf = t_{xf}f = xf,$$

and hence $fgf = f$. Therefore f is regular. But f was arbitrary, and so T_X is regular.

The next results shows that regularity is an invariant of a $\mathscr{D}$-class.

Theorem 11.3. *If a is a regular element of a semigroup S and $a \mathscr{D} b$, then b is regular.*

Proof. Since $a \mathscr{D} b$, there exists $c \in S$ such that $a \mathscr{L} c \mathscr{R} b$. We will start by proving that c is regular. Since a is regular, there exists $a' \in S$ such that $aa'a = a$. Also $a \mathscr{L} c$ and so there exist $u, v \in S^1$ such that $ua = c$ and $vc = a$. Thus

$$c = ua = uaa'a = ca'a = ca'vc,$$

and so c is regular.

By symmetry, since $b \mathscr{R} c$ and c is regular, it follows that b is regular too. □

It follows from Theorem 11.3 that either every element in a $\mathscr{D}$-class is regular, or none of the elements in the $\mathscr{D}$-class are regular.

Definition 11.4. (**Inverse of an element**). If $x, y \in S$ and $xyx = x$ and $yxy = y$, then y is called an ***inverse of*** x.

Theorem 11.5. *An element of a semigroup has an inverse if and only if it is regular.*

Proof. ($\Rightarrow$) This implication is obvious.

($\Leftarrow$) Assume that x is regular element of a semigroup. Then there exists $y \in S$ such that $x = xyx$. If $z = yxy$, then

$$xzx = xyxyx = xyx = x, \;\; zxz = y(xyx)yxy = y(xyx)y = yxy = z,$$

and so x and z are mutually inverse. □

The egg-box picture of a $\mathscr{D}$-class is very useful in locating the inverses of an element.

Theorem 11.6. *Let a be a regular element of a semigroup S, and let D_a be the $\mathscr{D}$-class of a.*

(a) *If a' is an inverse of a, then $a' \in D_a$ and the $\mathscr{H}$-classes $R_a \cap L_{a'}$ and $L_a \cap R_{a'}$ contain the idempotents aa' and $a'a$, respectively.*

(b) *If $b \in D_a$ and there exist idempotents $e \in R_a \cap L_b$ and $f \in R_b \cap L_a$, then H_b contains one and only one inverse a' of a that satisfies $aa' = e$ and $a'a = f$.*

Proof. **(a).** It is straightforward to verify that aa' and $a'a$ are idempotents. From $aa'a = a$, it follows that $a\mathscr{R}aa'$ and $a'a\mathscr{L}a$, and similarly from $a'aa' = a'$ it follows that $aa'\mathscr{L}a'$ and $a'a\mathscr{R}a'$. In particular, $a\mathscr{D}a'$, $aa' \in R_a \cap L_{a'}$, and $a'a \in L_a \cap R_{a'}$.

(b). By assumption, $a\mathscr{R}e$ and $a\mathscr{L}f$, and so there exist $s, t, u, v \in S^1$ such that

$$as = e,\ et = a,\ ua = f,\ vf = a.$$

If we define $a' = fse$, then

$$\begin{aligned} aa'a &= afsea = vf^2se^2t = vfset = aset = e^2t = et = a, \\ a'aa' &= fseafse = fsevf^2se = fsevfse = fsease = fse^3 = fse = a', \end{aligned}$$

and so a and a' are mutually inverse. Also we have

$$\begin{aligned} aa' &= afse = vf^2se = vfse = ase = e^2 = e, \\ a'a &= fsea = uase^2t = ue^3t = uet = ua = f. \end{aligned}$$

It follows from (a) that $a'\mathscr{L}aa' = e\mathscr{L}b$ and $a'\mathscr{R}a'a = f\mathscr{R}b$, in other words $a'\mathscr{H}b$.

Assume that H_b contains an inverse a''. It follows from (a) that $aa'' = e = aa'$ and that $a''a = f = a'a$. But then

$$a' = a'aa' = a''aa' = a''aa'' = a'',$$

as required. □

Theorem 11.7. *If D is a $\mathscr{D}$-class consisting of regular elements, then every $\mathscr{R}$-class and every $\mathscr{L}$-class of D contain an idempotent (and hence contains a group $\mathscr{H}$-class).*

Proof. Let $a \in D$ be arbitrary. Then a is regular, and so by Theorem 11.5 there is an inverse a' for a in D, and, by Theorem 11.6(a), the $\mathscr{R}$-class R_a contains the idempotent aa'.

The proof for $\mathscr{L}$-classes is analogous. □

Next we prove that the isomorphism type of a group $\mathscr{H}$-class is invariant within a single $\mathscr{D}$-class.

Theorem 11.8. *If H and K are two group $\mathscr{H}$-classes in the same (regular) $\mathscr{D}$-class, then H and K are isomorphic (as groups).*

Proof. Let e be the identity (idempotent) of H and let f be the identity of K. If $a \in L_f \cap R_e$ and $b \in R_f \cap L_e$, then $e \in R_a \cap L_b$ and $f \in R_b \cap L_a$ are idempotents. Thus there exists an inverse $a' \in H_b (= R_f \cap L_e)$ of a such that $aa' = e$ and $a'a = f$ by Theorem 11.6(b).

Since a' is an inverse of a,

$$a'e = a'aa' = a'.$$

Hence by Lemma 10.5(b), the mapping $\lambda_a : R_e = R_a \longrightarrow R_f = R_b$ defined by $x\lambda_{a'} = a'x$ is a bijection and $(H)\lambda_{a'} = (H_e)\lambda_{a'} = H_{a'} = H_b$.

Similarly, $a'\mathscr{R}f$ and $a'a = f$, and so, by Lemma 10.5(a), the mapping $\rho_a : L_e = L_b \longrightarrow L_f = L_a$ defined by $(x)\rho_a = xa$ is a bijection and $(H_{a'})\rho_a = H_f = K$.

Thus the mapping $\phi : H \longrightarrow K$ defined by $(x)\phi = a'xa$ is a bijection (being the composition of $\lambda_{a'}$ and ρ_a). It remains to prove that ϕ is a homomorphism. To this end, let $x, y \in H$ be arbitrary. Then

$$(x\phi)(y\phi) = a'xaa'ya = a'xeya = a'xya = (xy)\phi,$$

the second to last equality holds as e is the identity of H. □

Remark 11.9. The word 'regular' appears in Theorem 11.8 since if a $\mathscr{D}$-class is not regular, then it contains no regular elements. In particular, such a $\mathscr{D}$-class contains no idempotents and so it has no group $\mathscr{H}$-classes.

Example 11.10. In this example, we completely describe the Green's structure of the full transformation semigroup T_4. By Theorem 9.17 it has four $\mathscr{D}$-classes (which are also $\mathscr{J}$-classes by Theorem 9.21):

$$D_i = \{f \in T_4 : \operatorname{rank}(f) = i\}$$

where $i = 1, 2, 3, 4$. The $\mathscr{D}$-class D_4 consists precisely of permutations; hence $D_4 = S_4$, the symmetric group on 4 elements, and it is also an $\mathscr{R}$-, $\mathscr{L}$- and $\mathscr{H}$-class. The $\mathscr{D}$-class D_3 has four $\mathscr{L}$-classes, determined by the images 123, 124, 134 and 234. It also has six $\mathscr{R}$-classes, determined by the kernels 12|3|4, 13|2|4, 14|2|3, 23|1|4, 24|1|3 and 34|1|2. The $\mathscr{H}$-class determined by the image ijk and the kernel $rs|t|u$ contains

an idempotent if and only if i, j and k are in the different classes of the kernel (check this). For instance, the $\mathscr{H}$-class H determined by the image 123, and the kernel 14|2|3 contains the idempotent

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 1 \end{pmatrix}.$$

An arbitrary element of H is obtained by assigning to the elements 1, 2, 3 different images from the same set, and defining the image of 4 to be the same as that of 1. Thus H has 6 elements and is isomorphic to the symmetric group S_3. By Theorem 11.8 (or by repeating the above argument) every group $\mathscr{H}$-class in D_3 is isomorphic to S_3.

The $\mathscr{D}$-class D_2 has six $\mathscr{L}$-classes, determined by the images 12, 13, 14, 23, 24 and 34, and seven $\mathscr{R}$-classes, determined by the kernels 123|4, 124|3, 134|2, 234|1, 12|34, 13|24 and 14|23. Each $\mathscr{H}$-class has two elements, and those containing idempotents are cyclic groups of order 2. Finally, the $\mathscr{D}$-class D_1 contains the four constant mappings; it is also an $\mathscr{R}$-class, and it has four singleton $\mathscr{L}$-classes. The complete egg-box picture is shown in Figure 5.

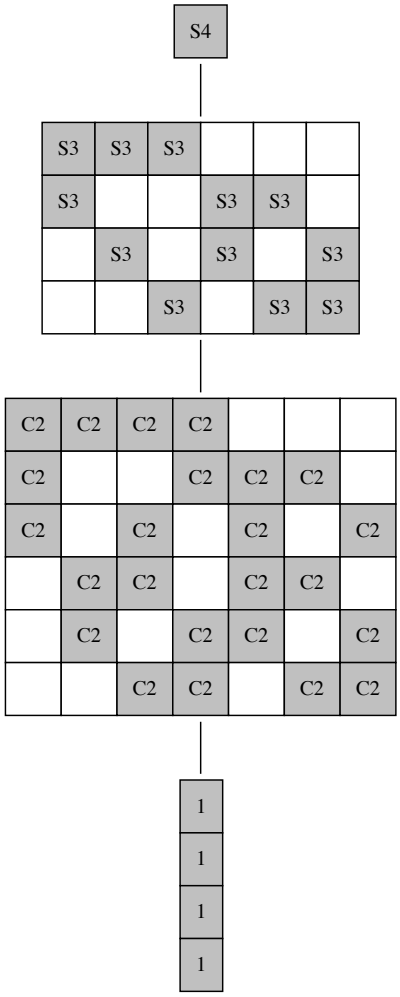

FIGURE 5. Egg box picture of T_4.

12. Inverse semigroups

We have seen that the notion of regularity, and the related notion of inverses, are an important bridge between semigroups and groups. Research has shown, however, that the class of regular semigroups is

very extensive, and that regular semigroups in general are still rather difficult to understand. As a result semigroup theorists have studied various classes of regular semigroups which are even closer to groups. Probably the most natural of these classes is that of inverse semigroups.

Definition 12.1. A semigroup S is said to be ***inverse*** if every element $s \in S$ has a unique (semigroup) inverse s^{-1}.

For example, every group is an inverse semigroup.

Before we give any further examples of inverse semigroups, let us recall some basic properties:

$$\begin{aligned} &xx^{-1}x = x,\ x^{-1}xx^{-1} = x^{-1},\\ &(x^{-1})^{-1} = x,\\ &x^2 = x \longrightarrow x^{-1} = x,\\ &(xx^{-1})^2 = xx^{-1}. \end{aligned}$$

The following theorem gives two further equivalent definitions for inverse semigroups.

Theorem 12.2. *Let S be a semigroup. Then the following conditions are equivalent:*

(a) *S is inverse;*

(b) *S is regular and its idempotents commute (i.e. if $e, f \in S$ are idempotents, then $ef = fe$);*

(c) *every $\mathscr{L}$-class and every $\mathscr{R}$-class of S contains exactly one idempotent.*

Proof. (a)⇒(b) Since S is inverse, it is regular. Let e and f be arbitrary two idempotents, and let $z = (ef)^{-1}$. Then

$$(ef)(fze)(ef) = ef^2ze^2f = efzef = ef(ef)^{-1}ef = ef$$
$$(fze)(ef)(fze) = fze^2f^2ze = f(zefz)e = f((ef)^{-1}ef(ef)^{-1})e = fze.$$

Therefore, fze is also an inverse of ef. By the uniqueness of inverses $z = fze$, and

$$z^2 = (fze)^2 = f(zefz)e = fze = z,$$

i.e. z is an idempotent. But then $z = z^{-1} = ef$, and so ef is an idempotent, and its own inverse. Similarly fe is an idempotent and so

$$(ef)(fe)(ef) = efef = ef,\ (fe)(ef)(fe) = fefe = fe,$$

i.e. fe is also an inverse of ef. Again, by the uniqueness of inverses, $ef = fe$.

(b)⇒(c) Assume that an $\mathscr{L}$-class L contains two idempotents e and f. By (Problem **5-8**) both e and f are right identities for L, and hence we have $e = ef = fe = f$. The statement for $\mathscr{R}$-classes is proved similarly.

(c)⇒(a) By Theorem 11.6(b), the inverses of an element $s \in S$ correspond to the pairs of idempotents $e \in R_s$, $f \in L_s$. Since both R_s and L_s contain precisely one idempotent, it follows that s has a unique inverse. □

Lemma 12.3. *Let S be an inverse semigroup and $x, y \in S$. Then $(xy)^{-1} = y^{-1}x^{-1}$.*

Proof. Since yy^{-1} and $x^{-1}x$ are idempotents, and since idempotents commute, we have

$$xyy^{-1}x^{-1}xy = xx^{-1}xyy^{-1}y = xy \quad \text{and} \quad y^{-1}x^{-1}xyy^{-1}x^{-1} = y^{-1}yy^{-1}x^{-1}xx^{-1} = y^{-1}x^{-1}.$$

Hence, the lemma follows by the uniqueness of inverses. □

Example 12.4. A commutative semigroup consisting of idempotents is called a ***semilattice***. Every semilattice S is an inverse semigroup since S satisfies Theorem 12.2(b). On the other hand, the idempotents of any inverse semigroup form a semilattice, since they are regular and they commute.

An example of a semilattice is the semigroup of all subsets of a set under union or intersection; see Example 6.6.

Example 12.5. A ***partial bijection*** is any injective partial mapping. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & - & - \end{pmatrix}$$

is a partial bijection.

It is easy to see that the set I_X of all partial bijections on a set X is a subsemigroup of the semigroup P_X of all partial mappings on X.

We are going to prove that I_X is inverse.

Let $f \in I_X$. Then f can be considered as a bijection $\mathrm{dom}(f) \longrightarrow \mathrm{im}(f)$ between its domain and its image. But then the inverse $f^{-1} : \mathrm{im}(f) \longrightarrow \mathrm{dom}(f)$ satisfies

$$ff^{-1}f = f \quad \text{and} \quad f^{-1}ff^{-1} = f^{-1}.$$

If g is any other mapping such that $fgf = f$ and $gfg = g$, then from the first equality we obtain $\mathrm{im}(f) \subseteq \mathrm{dom}(g)$, $\mathrm{dom}(f) \subseteq \mathrm{im}(g)$. The reverse inclusions can be obtained using the second equality, so that $\mathrm{dom}(g) = \mathrm{im}(f) = \mathrm{dom}(f^{-1})$ and $\mathrm{im}(g) = \mathrm{dom}(f) = \mathrm{im}(f^{-1})$. Finally, using $fgf = f$, we have that for every $j = if \in \mathrm{im}(f)$ we have $jg = ifg = ifgff^{-1} = iff^{-1} = i = jf^{-1}$, i.e. $g = f^{-1}$. In other words, every element of I_X has a unique inverse, and I_X is inverse.

The semigroup I_X is called the ***symmetric inverse semigroup*** on X; if $X = \{1, \ldots, n\}$ we write I_n rather than $I_{\{1,\ldots,n\}}$.

13. The Vagner-Preston Representation Theorem

One of the best examples of a theorem which states that inverse semigroups are intermediate objects between groups and general semigroups is the Vagner-Preston representation theorem. It parallels Cayley's Theorem for groups and its analogue for semigroups. Recall that Cayley's Theorem for groups asserts that every group is isomorphic to a subgroup of a symmetric group S_X. Similarly, Theorem 3.4 asserts that every semigroup is isomorphic to a subsemigroup of the full transformation semigroup T_X.

Theorem 13.1 (Vagner-Preston)**.** *Every inverse semigroup S is isomorphic to an (inverse) subsemigroup of a symmetric inverse semigroup I_X.*

Before proving the theorem we give some technical properties of inverse semigroups.

Lemma 13.2. *Let S be an inverse semigroup.*

(a) *If e and f are idempotents of S and $Se = Sf$, then $e = f$.*
(b) *If e and f are idempotents of S, then $Se \cap Sf = Sef$.*
(c) *if $a \in S$, then $Sa = Sa^{-1}a$.*

Proof. (a) From $e = ee \in Se = Sf$, we can write $e = xf$ for some $x \in S$, so that $ef = xff = xf = e$. By symmetry we have $fe = f$, but then $e = ef = fe = f$, because the idempotents commute in an inverse semigroup.

(b) Clearly $Sef \subseteq Sf$ and $Sef = Sfe \subseteq Se$, and so $Sef \subseteq Se \cap Sf$. For the converse inclusion let $x = ye = zf \in Se \cap Sf$. Then $x = zf = zff = xf = yef = yeef = xef \in Sef$.

(c) The result follows from $Sa^{-1}a \subseteq Sa = Saa^{-1}a \subseteq Sa^{-1}a$. □

Proof of Theorem 13.1. Let $X = S$, and consider the symmetric inverse semigroup I_X. For each $a \in S$ define a partial mapping τ_a, with domain Sa^{-1} and $x\tau_a = xa$ ($x \in Sa^{-1}$). To prove that this mapping is indeed in I_X we have to show that it is one-one on its domain. To this end let $x, y \in Sa^{-1}$, and write $x = x_1a^{-1}$, $y = y_1a^{-1}$. We have

$$\begin{aligned} x\tau_a = y\tau_a \quad &\Rightarrow \quad xa = ya \Rightarrow x_1a^{-1}a = y_1a^{-1}a \Rightarrow x_1a^{-1}aa^{-1} = y_1a^{-1}aa^{-1} \\ &\Rightarrow \quad x_1a^{-1} = y_1a^{-1} \Rightarrow x = y. \end{aligned}$$

Now define a mapping $\phi : S \longrightarrow I_X$ by $a\phi = \tau_a$. We are going to show that ϕ is a monomorphism, thus proving the theorem. It is easy to see that ϕ is one-one. Indeed if $a\phi = b\phi$ then, first of all we must have $Saa^{-1} = Sa^{-1} = \mathrm{dom}\,\tau_a = \mathrm{dom}\,\tau_b = Sb^{-1} = Sbb^{-1}$, and so $aa^{-1} = bb^{-1}$ by Lemma 13.2. But then

$$a = aa^{-1}a = (aa^{-1})\tau_a = (bb^{-1})\tau_a = (bb^{-1})\tau_b = bb^{-1}b = b.$$

Next we note that $(\tau_a)^{-1} = \tau_{a^{-1}}$; this follows from $x\tau_a\tau_a^{-1}\tau_a = xaa^{-1}a = xa = x\tau_a$ ($x \in Sa^{-1}$) and $x\tau_{a^{-1}}\tau_a\tau_{a^{-1}} = xa^{-1}aa^{-1} = xa^{-1} = x\tau_{a^{-1}}$ ($x \in Sa$), and the fact that I_X is an inverse semigroup.

Now consider two arbitrary $a, b \in S$. We want to show that $(ab)\phi = (a\phi)(b\phi)$, which is equivalent to proving $\tau_a\tau_b = \tau_{ab}$. First note that

$$\begin{aligned} \mathrm{dom}\,\tau_a\tau_b &= \mathrm{dom}\,\tau_a \cap (\mathrm{dom}\,\tau_b)\tau_a^{-1} = (\mathrm{im}\,\tau_a \cap \mathrm{dom}\,\tau_b)\tau_{a^{-1}} = (Sa^{-1}a \cap Sb^{-1})a^{-1} \\ &= (Sa^{-1}a \cap Sbb^{-1})a^{-1} = Sa^{-1}abb^{-1}a^{-1} = Saa^{-1}abb^{-1}a^{-1} \\ &= Sabb^{-1}a^{-1}aa^{-1} = Sabb^{-1}a^{-1} = S(ab)(ab)^{-1} = S(ab)^{-1} = \mathrm{dom}\,\tau_{ab}, \end{aligned}$$

by Lemma 13.2. Finally, for all x from this domain we have $x\tau_a\tau_b = xab = x\tau_{ab}$, as required. □

It is worth remarking that Vagner's theorem ***does not*** assert that every subsemigroup of I_X is inverse. For example, the subsemigroup of I_2 generated by the mapping which only send 1 into 2 is the zero semigroup with two elements; this semigroup is not regular, let alone inverse.

Today there is a rich theory of inverse semigroups. For example, there are descriptions of congruences on these semigroups in terms of subsemigroups, in a way similar to, but more complicated than, groups. Also, there is a separate theory of free inverse semigroups, and inverse semigroup presentations.