Lecture Notes on

SEMIGROUPS

Short Course

2010

Tero Harju
Department of Mathematics
University of Turku
FI-20014 Turku, Finland
`http://users.utu.fi/harju/`

Contents

1

Basic Concepts

1.1 Definitions and examples

Generic notation

- We let $\mathbb{N} = \{0, 1, \dots\}$ be the set of the non-negative integers, and $\mathbb{N}_+ = \{1, 2, \dots\}$ the set of the positive integers. Also, $\mathbb{Z}$ denotes the set of the integers, $\mathbb{Q}$ the set of the rational numbers, $\mathbb{R}$ the set of the real numbers, and $\mathbb{C}$ the set of the complex numbers.
- For a function $f\colon X \to Y$, we may write $x \mapsto y$ to denote that $f(x) = y$.
- For a set X and a mapping $\alpha\colon X \to Y$, write

$$\alpha(X) = \{\alpha(x) \mid x \in X\} \subseteq Y\,,$$
$$\alpha^{-1}(X) = \{y \mid \alpha(x) = y \text{ for some } x \in X\} \subseteq Y\,.$$

- For a mapping $\alpha\colon X \to Y$, let $\alpha \restriction A$ be the **restriction** of α to the subset $A \subseteq X$, that is, $\alpha \restriction A\colon A \to Y$ is defined by

$$(\alpha \restriction A)(x) = \alpha(x) \qquad (x \in A)\,.$$

- Let $\mathcal{B}(X, Y)$ be the set of the (binary) relations in $X \times Y$:

$$\mathcal{B}(X, Y) = \{\delta \mid \delta \subseteq X \times Y\} \text{ with } \mathcal{B}(X) = \mathcal{B}(X, X)\,.$$

  We write $(x, y) \in \delta$ or $x\delta y$ to denote that the (ordered) pair (x, y) is in relation δ.
- For a relation $\delta \subseteq X \times Y$, let

$$x\delta = \{y \in Y \mid x\delta y\}.$$

  Also, for $A \subseteq X$, we adopt

$$A\delta = \bigcup_{y \in A} y\delta, \quad \operatorname{ran}(\delta) = X\delta, \quad \operatorname{dom}(\delta) = X\delta^{-1}\,,$$

  where $\delta^{-1} = \{(y, x) \mid (x, y) \in \delta\}$. In particular, $\delta \subseteq \operatorname{dom}(\delta) \times \operatorname{ran}(\delta)$.
- The **identity relation** on a set X is given by $\iota = \iota_X = \{(x, x) \mid x \in X\}$. (It is the identity function on X.) The **universal relation** in $\mathcal{B}(X)$ is the relation $\omega = \omega_X = X \times X = \{(x, y) \mid x, y \in X\}$.

- A relation $\delta \in \mathcal{B}(X)$ is an **equivalence relation**, if its is reflexive ($\iota_X \subseteq \delta$), symmetric ($\delta^{-1} = \delta$) and transitive ($\delta^2 \subseteq \delta$). The sets $x\delta$ are the **equivalence classes** of δ. They form a **partition** of the set X,

$$X = \bigcup_{x \in X} x\delta \qquad \text{and} \qquad x\delta \cap y\delta \neq \emptyset \iff x\delta = y\delta\,.$$

- The sets X and Y have the **same cardinality**, denoted $|X| = |Y|$, if there is a bijection $\alpha\colon X \to Y$. We say that a set X is **denumerable**, if it is finite or if it has the cardinality of $\mathbb{N}$. The elements of a denumerable set X can be listed: $X = \{x_1, x_2, \dots\}$. (This is not the case for $\mathbb{R}$, which is not denumerable).

Semigroups

Let S be a set and $\star\colon S \times S \to S$ a binary operation, called a **product**, and often written in infix notation $(x, y) \mapsto x \star y$. The pair $(S, \star)$ (or simply S, if there is no fear of confusion) is called a **groupoid**. We shall mostly write xy instead of $x \star y$, or if in need to emphasize the place of the operation, we write $x \cdot y$. When the groupoid is naturally given, we adopt the corresponding notation: $\cdot$, $+$, $\star$, $\circ$, $\oplus$, $\otimes$. If we do not mention the (notation for the) product, we choose: $x \cdot y$.

A groupoid S is a **semigroup**, if its defining operation is **associative**,

$$\forall x, y, z \in S\colon \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z\,.$$

Hence the placement of the brackets is irrelevant, and therefore we may write

$$x_1 x_2 \cdots x_n = x_1 \cdot (x_2 \cdot (\cdots (x_{n-1} \cdot x_n)) \cdots)\,.$$

For an element $x \in S$, let $x^1 = x$, $x^2 = x \cdot x$, and $x^{n+1} = x \cdot x^n$ for $n \geq 1$.

- A semigroup S is **commutative**, if $x \cdot y = y \cdot x$, for all $x, y \in S$.
- A semigroup S is **left cancellative** (resp., **right cancellative**) if

$$\forall x, y, z \in S\colon\ zx = zy \implies x = y \qquad (\forall x, y, z \in S\colon\ xz = yz \implies x = y).$$

If S is both left and right cancellative, then it is **cancellative**.

Example 1.1. (1) The grupoids $(\mathbb{N}, \cdot)$ and $(\mathbb{N}, +)$, with the usual operations, are commutative semigroups. Also, $(\mathbb{N}, \star)$ is a (commutative) semigroup when $n \star m = \max\{n, m\}$. Indeed,

$$\begin{aligned} n \star (m \star k) &= \max\{n, \max\{m, k\}\} = \max\{n, m, k\} \\ &= \max\{\max\{n, m\}, k\} = (n \star m) \star k\,. \end{aligned}$$

(2) Consider the upper triangular 2×2 integer matrices

$$S = \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mid\ n \geq 1 \right\}.$$

Then S is a semigroup with the usual matrix multiplication.

(3) For a nonempty set X, let

$$T_X = \{\alpha \mid \alpha\colon X \to X\}$$

be the set of all functions on X. Then $(T_X, \circ)$ is a semigroup with respect to the composition of functions: $(\beta \circ \alpha)(x) = \beta(\alpha(x))$ for all $x \in X$. The semigroup T_X is the **full transformation semigroup** on X, and it has a special place in the theory of semigroups.

(4) Let $S = \{a, b, c\}$ and define the product $\star$ as in the given table. Then S is a finite semigroup. In order to check this, we must test, for all triples, that they satisfy the associativity requirement: $x \star (y \star z) = (x \star y) \star z$. In general, this may be tedious. In this case, there are some helpful restrictions in S. For instance, if $z = c$, then the product is always c no matter what x and y are. □

$\star$	a	b	c
a	a	b	c
b	b	a	c
c	c	b	c

(5) The set $\mathcal{B}(X)$ of the relations on X forms a semigroup under composition:

$$\rho \circ \delta = \{(x, y) \in X \times X \mid \exists z\colon (x, z) \in \rho,\ (z, y) \in \delta\}\,.$$

(6) Consider the matrix semigroup $\mathbb{Z}^{2\times 2}$. Here

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

and hence $\mathbb{Z}^{2\times 2}$ is not cancellative. However, if we take all matrices A with $\det(A) = 1$, then this semigroup is cancellative. □

Example 1.2. The following table gives the number of (nonisomorphic) finite semigroups as given by the "Encyclopedia of Integer Sequences:" Let $f(n)$ denote the number of semigroups with n elements.

$f(1) =$	1	$f(5) =$	1 915
$f(2) =$	5	$f(6) =$	28 634
$f(3) =$	24	$f(7) =$	1 627 672
$f(4) =$	188		

One should compare this table to the number of groups. There are only two groups of order 6, and one for all prime numbers. □

Special elements

Let S, i.e., $(S, \cdot)$, be a semigroup.

- An element $x \in S$ is a **left identity element** (resp. **right identity element**) of S, if

  $$\forall y \in S\colon\ x \cdot y = y \qquad (\forall y \in S\colon\ y \cdot x = y)\,.$$

  If x is both a left and a right identity element of S, then x is called an **identity** (element) of S, and S is a **monoid**. The identity of a monoid S is usually denoted by 1_S, or simply by 1. (If S has an established identity, we use that notation.)

- An element $x \in S$ is a **left zero element** (resp., **right zero element**), if

$$\forall y \in S\colon\ x{\cdot}y = x \qquad (\forall y \in S\colon\ y{\cdot}x = x).$$

  A **zero element** is both a left and a right zero of S.
- An element $e \in S$ is an **idempotent**, if $e^2 = e$. Let

$$E_S = \{e \mid e^2 = e\}.$$

Lemma 1.1. *A semigroup S has at most one identity, and at most one zero element.*

Proof. Exercise. □

In fact, if S has a left identity x and a right identity y, then $x = y$. In particular, the identity of a monoid is unique.

A monoid G is a **group**, if each $x \in G$ has a (**group**) **inverse** $x^{-1} \in G$:

$$x{\cdot}x^{-1} = 1 = x^{-1}{\cdot}x\,.$$

All groups are cancellative semigroups. Although semigroups and groups seem to be rather close to each other, semigroups lack the basic symmetry properties of groups, and therefore the general theory of semigroups differs drastically from the theory of groups.

Example 1.3. (1) If G is a group, then $E_G = \{1\}$.

(2) Consider the finite semigroup S defined in Example 1.1(4). The element a is an identity of S, and hence S is a monoid. The element c is a right zero element. There are no left zeros, and hence S does not have a zero element. Both a and c are idempotents. Indeed, all left and right identities and zero elements are idempotents. The element b is not an idempotent, since $b^2 = a$.

(3) The matrix semigroup from Example 1.1(2) has no idempotents.

(4) Consider the set of all $n \times n$ matrices with integer entries. This is a monoid $\mathbb{Z}^{n\times n}$. For $n = 2$,

$$\mathbb{Z}^{2\times 2} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}.$$

The identity matrix I is the identity element of $\mathbb{Z}^{2\times 2}$ and the zero matrix 0 is its zero element. This monoid has quite many idempotents.

(5) Let $GL(n, \mathbb{R})$ denote the monoid of all nonsingular $n \times n$ matrices A with real entries – so that $\det(A) \neq 0$. Apart from the identity matrix, $GL(n, \mathbb{R})$ has no other idempotents, because it is a group, called the **generalized linear group** over $\mathbb{R}$. Also. the monoid $SL(n, \mathbb{R})$ of all n-by-n-matrices A in $\mathbb{R}^{n\times n}$ such that $\det(A) = 1$ forms a group. This is the **special linear group** over $\mathbb{R}$. □

1.2 Subsemigroups and direct products

Subsemigroups

A subset $A \subseteq S$ of a semigroup S is said to **generate** S, if

$$S = \{a_1 a_2 \cdots a_n \mid n \geq 1,\ a_i \in A\}.$$

In this case we also say that A is a **generator set** of S.

Example 1.4. (1) The additive semigroup $(\mathbb{N}_+, +)$ is generated by one element $A = \{1\}$.

(2) The multiplicative semigroup $(\mathbb{N}_+, \cdot)$ does not have finite generator sets. The smallest generator set of this semigroup consists of the prime numbers together with 1. □

We say that a subset X of a *monoid* M **generates** M (**as a monoid**), if $X \cup \{1_M\}$ generates M. Hence in a monoid the identity element is always taken into granted.

Let A be a nonempty subset of a semigroup $(S, \cdot)$. We say that $(A, \cdot)$ is a **subsemigroup** of S, denoted by $A \leq S$, if A is closed under the product of S:

$$\forall x, y \in A\colon\ x \cdot y \in A\,.$$

Example 1.5. (1) Consider the additive semigroup $S = (\mathbb{Q}, +)$ of rational numbers. Then $(\mathbb{N}, +)$ is a subsemigroup of S, but $(\mathbb{N}, \cdot)$ is not.

(2) A **numerical semigroup** S is a subsemigroup of $(\mathbb{N}, +)$ such that $0 \in S$ and the complement $\mathbb{N} \setminus S$ is finite. Each numerical semigroup S is finitely generated, that is, there exists a finite subset $A \subseteq S$ such that

$$S = \{k_1 a_1 + k_2 a_2 + \cdots + k_n a_n \mid n \geq 1,\ k_1, k_2, \ldots, k_n \geq 0 \text{ and } a_1, a_2, \ldots, a_n \in A\}.$$

If $S \neq \mathbb{N}$, then there exists the largest integer $f_S \in \mathbb{N}$ such that $f_S \notin S$. This number is called the **Frobenius number** of S. □

Theorem 1.1. *Let S be a semigroup.*

(1) Let $A_i \leq S$ be subsemigroups of for $i \in I$. If their intersection is nonempty, then

$$\bigcap_{i \in I} A_i \leq S\,.$$

(2) Let $X \subseteq S$ be nonempty. Then

$$[X]_S = \bigcup_{n=1}^{\infty} X^n = \{x_1 x_2 \cdots x_n \mid n \geq 1,\ x_i \in X\}$$

*is the smallest subsemigroup containing X, called the **subsemigroup generated by** X.*

Proof. (1) Suppose the intersection is nonempty. If $x, y \in A = \cap_{i\in I} A_i$, then $x, y \in A_i$ for all $i \in I$, and hence $xy \in A$ as required.

(2) This follows directly from part (1). □

The subsemigroup $[X]_S$ is often shortened as $[X]$. When $X = \{x_1, x_2, \ldots, x_n\}$ is finite, we write $[x_1, x_2, \ldots, x_n]_S$ for $[X]_S$.

Example 1.6. Consider the matrix semigroup $S = \mathbb{Z}^{2\times 2}$, and let

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Now,

$$M^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad M^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad M^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M^5 = M,$$

and hence $[M]_S = \{I, M, M^2, M^3\}$ is a finite subsemigroup of S. In fact, $[M]_S$ is a subgroup of S, that is, a subsemigroup which is a group. Note that $\mathbb{Z}^{2\times 2}$ is not a group. □

Index and period

If $X = \{x\}$ is a singleton subset of a semigroup S, then $[x]_S = \{x^n \mid n \geq 1\}$ is called the **cyclic semigroup** generated by x. There are two possibilities for $[x]_S$: Either it is infinite and $x^n \neq x^m$ for all $n \neq m$, or there exists an integer k such that $[x]_S = \{x, x^2, \ldots, x^{k-1}\}$ in which case $x^k = x^r$ for some $1 \leq r < k$. Here r is called the **index** and $p = k - r$ the **period** of x. We have now

$$\forall n \geq r\colon\ x^n = x^{n+p}\,.$$

Lemma 1.2. *Let r be the index and p the period of an element $x \in S$. Then*

$$K_x = \{x^r, x^{r+1}, \ldots, x^{r+p-1}\}$$

is a subgroup (i.e., a subsemigroup that is a group) of S.

Proof. Exercise. □

Direct products

The **direct product** $S \times T$ of two semigroups S and T is defined by

$$(x_1, y_1)\cdot(x_2, y_2) = (x_1x_2, y_1y_2) \qquad (x_i \in S,\ y_i \in T)\,.$$

It is easy to show that the so defined product is associative on $S \times T$, and hence the direct product is, indeed, a semigroup.

Example 1.7. Let $S = (\mathbb{N}, +)$ and $T = (\mathbb{N}, \cdot)$. Then in the direct product $S \times T$ we have $(n, r)\cdot(m, s) = (n + m, rs)$. □

The direct product is a convenient way of combining two semigroups. The new semigroup $S \times T$ inherits properties of both S and T. The mappings $\pi_1 \colon S \times T \to S$ and $\pi_2 \colon S \times T \to T$ such that $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$ are called the **projections** of the direct product $S \times T$.

1.3 Homomorphisms and transformations

Definition

Let $(S, \cdot)$ and $(P, \star)$ be two semigroups. A mapping $\alpha \colon S \to P$ is a **homomorphism**, if

$$\forall x, y \in S \colon \ \alpha(x\cdot y) = \alpha(x) \star \alpha(y)\,.$$

Thus a homomorphism respects the product of S while 'moving' elements to P (which may have a completely different operation as its product). However, a homomorphism may also identify elements: $\alpha(x) = \alpha(y)$.

Example 1.8. (1) Let $S = (\mathbb{N}, +)$ and $P = (\mathbb{N}, \cdot)$, and define $\alpha(n) = 2^n$ for all $n \in \mathbb{N}$. Now,

$$\alpha(n + m) = 2^{n+m} = 2^n\cdot 2^m = \alpha(n)\cdot\alpha(m)\,,$$

and hence $\alpha \colon S \to P$ is a homomorphism.

(2) Define S and P by the following tables:

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

$\star$	e	f	g
e	e	f	g
f	f	e	g
g	g	g	g

These are semigroups – which is never too obvious! Define a mapping $\alpha \colon S \to P$, that identifies the elements c and d of S, as follows

$$\alpha(a) = e,\ \alpha(b) = f,\ \alpha(c) = g = \alpha(d)\,.$$

Now, a is the identity of S, and its image $e = \alpha(a)$ is the identity of P. Therefore for all $x \in S$, $\alpha(a\cdot x) = \alpha(x) = e \star \alpha(x) = \alpha(a) \star \alpha(x)$, and similarly $\alpha(x\cdot a) = \alpha(x) \star \alpha(a)$. Also, the other cases can be checked easily to ensure that α is a homomorphism.

(3) Let $S = (\mathbb{Z}, \cdot)$ and $P = (\mathbb{Z}, +)$. Define $\alpha \colon S \to P$ by $\alpha(n) = n$ for all $n \in \mathbb{Z}$. Then α is *not* a homomorphism, because $6 = \alpha(6) = \alpha(2\cdot 3) \neq \alpha(2) + \alpha(3) = 5$.

(4) If $\alpha \colon S \to P$ is a homomorphism, then $\alpha(x^n) = (\alpha(x))^n$ for all $x \in S$ and $n \geq 1$. (The latter is usually written as $\alpha(x)^n$.) □

Theorem 1.2. *Let $\alpha, \beta\colon S \to P$ be homomorphisms and let $X \subseteq S$. Then $\alpha([X]_S) = [\alpha(X)]_P$. and*

$$\alpha \restriction X = \beta \restriction X \iff \alpha \restriction [X]_S = \beta \restriction [X]_S \,.$$

Proof. If $x \in [X]_S$, then, by Lemma 1.1, $x = x_1x_2\cdots x_n$ for some $x_i \in X$. Since α is a homomorphism,

$$\alpha(x) = \alpha(x_1)\alpha(x_2)\cdots\alpha(x_n) \in [\alpha(X)]_P \,,$$

and hence $\alpha([X]_S) \subseteq [\alpha(X)]_P$. On the other hand, if $y \in [\alpha(X)]_P$, then, again by Lemma 1.1, $y = \alpha(x_1)\alpha(x_2)\cdots\alpha(x_n)$ for some $\alpha(x_i) \in \alpha(X)$ with $x_i \in X$. The claim follows, since α is a homomorphism: $y = \alpha(x_1x_2\cdots x_n)$, where $x_1x_2\cdots x_n \in [X]_S$.

The second claim is an exercise. □

Corollary 1.1. *If $A \leq S$ and $\alpha\colon S \to P$ is a homomorphism, then $\alpha(A) \leq P$.*

Proof. Immediate from the preceding lemma. □

Lemma 1.3. *If $\alpha\colon S \to P$ and $\beta\colon P \to T$ are homomorphisms, so is $\beta\alpha\colon S \to T$.*

Proof. Indeed, for all x and y,

$$\beta\alpha(x\cdot y) = \beta(\alpha(x\cdot y)) = \beta(\alpha(x)\cdot\alpha(y)) = \beta(\alpha(x))\cdot\beta(\alpha(y)) = \beta\alpha(x)\cdot\beta\alpha(y).$$

□

Isomorphism and embeddings

A homomorphism $\alpha\colon S \to P$ is

- an **embedding** or a **monomorphism**, denoted by $\alpha\colon S \hookrightarrow P$, if it is injective, that is, if $\alpha(x) = \alpha(y)$ implies $x = y$;
- an **epimorphism**, denoted by $\alpha\colon S \twoheadrightarrow P$, if it is surjective, that is, if for all $y \in P$ there exists an $x \in S$ with $\alpha(x) = y$;
- an **isomorphism**, denoted by $\alpha\colon S \xrightarrow{\sim} P$, if it is both an embedding and an epimorphism;
- an **endomorphism**, if $P = S$;
- an **automorphism**, if it is both an isomorphism and an endomorphism.

Lemma 1.4. *If $\alpha\colon S \to P$ is an isomorphism, then also the inverse map $\alpha^{-1}\colon P \to S$ is an isomorphism.*

Proof. First, the inverse mapping α^{-1} exists, because α is a bijection. Furthermore, $\alpha\alpha^{-1} = \iota$, and thus because α is a homomorphism,

$$\alpha(\alpha^{-1}(x)\cdot\alpha^{-1}(y)) = \alpha(\alpha^{-1}(x))\cdot\alpha(\alpha^{-1}(y)) = xy \,,$$

and thus $\alpha^{-1}(x)\cdot\alpha^{-1}(y) = \alpha^{-1}(xy)$, which proves the claim. □

A semigroup S is **embeddable** in another semigroup P, if there exists an embedding $\alpha\colon S \hookrightarrow P$, and S is **isomorphic** to P, denoted $S \cong P$, if there exists an isomorphism $\alpha\colon S \rightarrowtail P$. Two isomorphic semigroups share their *algebraic* properties (but they may differ in their *combinatorial* properties).

Example 1.9. (1) If $S = [x]$ is an infinite cyclic semigroup, then $\alpha\colon S \to (\mathbb{N}_+, +)$, defined by $\alpha(x^n) = n$, is a homomorphism: $\alpha(x^n x^m) = \alpha(x^{n+m}) = n+m = \alpha(x^n) + \alpha(x^m)$. Moreover, α is a bijection, and thus an isomorphism. Therefore, $S \cong (\mathbb{N}_+, +)$. In conclusion, *each infinite cyclic semigroup is isomorphic to the additive semigroup of positive integers.*

(2) If $P \leq S$, then the **inclusion mapping** $\iota\colon P \to S$, $\iota(x) = x$, is a homomorphism, and it is injective (but need not be surjective). Therefore ι is an embedding (of P into S). In particular, the **identity function** $\iota\colon S \to S$, $\iota(x) = x$, is always an automorphism, the **trivial automorphism**, of S. □

Theorem 1.3. *(1) The endomorphisms of a semigroup S form a monoid.*
(2) The automorphisms of a semigroup S form a group.

Proof. Exercise. □

Representations by full transformation semigroups

Example 1.10. Let $X = \{1, 2, 3\}$. There are altogether $3^3 = 27$ functions in the full transformation semigroup T_X. A mapping $\alpha\colon X \to X$, defined by $\alpha(1) = 2$, $\alpha(2) = 3$ and $\alpha(3) = 3$, can be represented conveniently in two different ways:

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \end{pmatrix} \qquad \text{or} \qquad 1 \overset{\alpha}{\to} 2 \overset{\alpha}{\to} 3\,.$$

Let then

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

and $S = [\alpha, \beta]_{T_X}$ the subsemigroup of T_X generated by α and β. We have

$$\alpha^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix} \qquad \beta^2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \qquad \beta\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \end{pmatrix}$$
$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \end{pmatrix} \qquad \beta\alpha^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \end{pmatrix} = \alpha^2\beta\,.$$

One may check that there are no other elements in this semigroup, and hence S has seven elements, $S = \{\alpha, \beta, \alpha^2, \beta^2, \alpha\beta, \beta\alpha, \beta\alpha^2\}$. □

A homomorphism $\varphi\colon S \to T_X$ is a **representation** of the semigroup S. A representation φ is **faithful**, if it is an embedding: $\varphi\colon S \hookrightarrow T_X$.

The following theorem states that semigroups can be thought of as subsemigroups of the full transformation semigroups, that is, for each semigroup S there exists a set X such that $S \cong P \leq T_X$ for a subsemigroup P of transformations.

For a semigroup S, we define a monoid S^1 by adjoining an identity to S, if S does not have one:

$$S^1 = \begin{cases} S & \text{if } S \text{ is a monoid,} \\ S \cup \{1\} & \text{if } S \text{ is not a monoid,} \end{cases}$$

where 1 is a (new) identity element.

Theorem 1.4. *Every semigroup S has a faithful representation.*

Proof. Let $X = S^1$, that is, add the identity 1 to S if S is not a monoid. Consider the full transformation semigroup $T = T_{S^1}$. For each $x \in S$, define

$$\rho_x\colon S^1 \to S^1, \qquad \rho_x(y) = xy \quad (y \in S^1)\,.$$

Thus $\rho_x \in T$, and for all $x, y \in S$ and for all $z \in S^1$,

$$\rho_{xy}(z) = (xy)z = \rho_x(yz) = \rho_x(\rho_y(z)) = \rho_x\rho_y(z)\,,$$

and hence $\rho_{xy} = \rho_x\rho_y$. Consequently, the mapping

$$\varphi\colon S \to T \text{ given by } \varphi(x) = \rho_x$$

is a homomorphism. For injectivity we observe that

$$\varphi(x) = \varphi(y) \implies \rho_x = \rho_y \implies \rho_x(1) = \rho_y(1) \implies x = y\,.$$

□

In the previous proof the special element 1 is needed (only) to ensure injectivity of φ. It is needed, because there exists a semigroup S (without identity) that has two left identities $x \neq y$. In such a case, $xz = yz$ for all $z \in S$.

Hence, loosely speaking the theory of semigroups can be taught of as the theory of transformations. This situation is reflected in a more restricted manner in the theory of groups, where we have the following fundamental theorem of representations.

Theorem 1.5 (Cayley). *Each group can be represented faithfully as a (semi)group of permutations (bijections $X \to X$).*

Proof. Exercise. □

Example 1.11. Let S be the semigroup on the set $X = \{e, a, b, c\}$ given in the next table. We have the following faithful representation $\varphi\colon S \to T_X$ for S.

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

Note that the mappings ρ_x are just the rows of the multiplication table.

$$\varphi(e) = \begin{pmatrix} e\,a\,b\,c \\ e\,a\,b\,c \end{pmatrix}, \qquad \varphi(a) = \begin{pmatrix} e\,a\,b\,c \\ a\,e\,c\,b \end{pmatrix},$$
$$\varphi(b) = \begin{pmatrix} e\,a\,b\,c \\ b\,b\,b\,b \end{pmatrix}, \qquad \varphi(c) = \begin{pmatrix} e\,a\,b\,c \\ c\,c\,c\,c \end{pmatrix}.$$

The representation φ is by no means unique. Consider $\varphi_1\colon S \to T_{\{1,2\}}$ defined by

$$\varphi_1(e) = \begin{pmatrix} 1\,2 \\ 1\,2 \end{pmatrix}, \ \varphi_1(a) = \begin{pmatrix} 1\,2 \\ 2\,1 \end{pmatrix}, \ \varphi_1(b) = \begin{pmatrix} 1\,2 \\ 1\,1 \end{pmatrix}, \ \varphi_1(c) = \begin{pmatrix} 1\,2 \\ 2\,2 \end{pmatrix}.$$

One may check that this simpler φ_1 also represents S by checking that the transformations $\varphi_1(x)$ satisfy the multiplication table of S. □

Semilattice of idempotents

A relation $\delta \in \mathcal{B}(X)$ is a **partial order**, if it is reflexive, antisymmetric ($\delta \cap \delta^{-1} \subseteq \iota$) and transitive. A partial order is often denoted by the symbols $\leq$ or $\subseteq$.

The idempotents E_S of a semigroup S are partially ordered as follows, if $E_S \neq \emptyset$. Define for $e, f \in E_S$:

$$e \leq f \iff ef = e = fe\,.$$

Lemma 1.5. *The relation $\leq$ is a partial order on E_S.*

Proof. First of all, for all $e \in E_S$, $e^2 = e$, and hence $e \leq e$ (reflexive). If $e \leq f$ and $f \leq e$, then $e = ef = f$ (antisymmetric). If $e \leq f$ and $f \leq h$, then

$$e = ef = efh = eh \qquad \text{and} \qquad e = fe = hfe = he\,,$$

and so also $e \leq h$ (transitive). □

If S is a commutative semigroup and all its elements are idempotents ($S = E_S$), then S is called a **semilattice**. Hence, in a semilattice, $x^2 = x$ and $xy = yx$ for all $x, y \in S$.

The relation $\leq$ on idempotents is defined on the entire semilattice S. An element g is a **lower bound** of elements e and f, if $g \leq e$ and $g \leq f$.

Lemma 1.6. *Let S be a semilattice. Then $ef \in S$ is the greatest lower bound of the elements e and f of S.*

Proof. Let $e, f \in S$. Then $ef = fe = ffe = eff$, and thus $(ef)f = ef = f(ef)$, which means that $ef \leq f$. Similarly, $ef \leq e$ holds, and hence ef is a lower bound of e and f. If g is also a lower bound of e and f, then $g = gf = ge$, and therefore

$$g(ef) = (ge)f = gf = g\,,$$

and hence $g \leq ef$. Therefore $ef \in S$ is the greatest lower bound of the elements e and f in the semilattice S. □

2

General Structure Results

2.1 Quotient semigroups

Congruences

An equivalence relation ρ on a semigroup S is a **left congruence** (resp., **right congruence**), if

$$\forall x, y, z \in S\colon\ x\rho y \implies (zx)\rho(zy) \qquad (\forall x, y, z \in S\colon\ x\rho y \implies (xz)\rho(yz))$$

and ρ is a **congruence**, if it is both a left and a right congruence. An equivalence class of a congruence is called a **congruence class**. We can also use the symbols $\sim$, $\simeq$, $\approx$, $\equiv$ for congruences.

Let $\mathrm{Con}(S)$ be the set of the congruences of the semigroup S.

If ρ is a congruence, then it respects the product of S as in the following lemma.

x_1		$x_1 \cdot x_2$		x_2
	$\vdots$		$\vdots$	
y_1		$y_1 \cdot y_2$		y_2

Lemma 2.1. *An equivalence relation ρ on S is a congruence if and only if for all $x_1, x_2, y_1, y_2 \in S$:*

$$\left.\begin{matrix} x_1\rho y_1 \\ x_2\rho y_2 \end{matrix}\right\} \implies (x_1x_2)\rho(y_1y_2)\,.$$

Proof. Suppose first that $\rho \in \mathrm{Con}(S)$. If $x_1\rho y_1$ and $x_2\rho y_2$, then, by the definition, $(x_1x_2)\rho(x_1y_2)$ and $(x_1y_2)\rho(y_1y_2)$ and hence also $(x_1x_2)\rho(y_1y_2)$ by transitivity. In the converse, the claim is trivial. □

Quotients

Let $\rho \in \mathrm{Con}(S)$ be a congruence of a semigroup S, and let

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

be the set of the congruence classes of ρ. We define a new semigroup, the **quotient semigroup** (of S **modulo** ρ), on S/ρ by

$$x\rho \cdot y\rho = (xy)\rho\,.$$

This is a well defined binary operation by Lemma 2.1. Associativity is easily checked:

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

The quotient S/ρ is thus obtained from S by contracting each congruence class $x\rho$ to a single element.

Example 2.1. (1) Consider the finite semigroup S defined in the leftmost table. Here e and f are idempotents. Indeed, e is an identity of S.

$\cdot$	e	a	f	b
e	e	a	f	b
a	a	e	b	f
f	f	b	f	b
b	b	f	b	f

$\cdot$	c	d	g
c	c	d	g
d	d	c	g
g	g	g	g

Let $\rho = \iota \cup \{(f,b),(b,f)\}$. Then ρ is a congruence of S, and its congruence classes are: $c = \{e\}$, $d = \{a\}$ and $g = \{f,b\}$. The multiplication table for the quotient semigroup S/ρ is given in the rightmost table.

Also, $\rho' = \iota \cup \{(e,a),(a,e),(f,b),(b,f)\}$ is a congruence. It has only two congruence classes, $\{e,a\}$ and $\{f,b\}$, and hence the quotient S/ρ' is a semigroup of two elements.

The relation $\rho_{\mathrm{no}} = \iota \cup \{(a,b),(b,a)\}$ is *not* a congruence, because $a\cdot a = e$ and $a\cdot b = f$ in S, but $e\rho_{\mathrm{no}}f$ does not hold.

(2) Consider the additive semigroup $S = (\mathbb{Z},+)$ with a congruence $\sim$. Then $n \sim m$ implies $(n+k) \sim (m+k)$ for all $k \in \mathbb{Z}$. Suppose that k is the smallest non-negative integer such that $n \sim (n+k)$ for some $n \in \mathbb{Z}$. In particular, $(n-n) \sim (n+k-n)$, i.e., $0 \sim k$. Let $m \equiv \overline{m} \pmod{k}$, where $0 \leq \overline{m} < m$. By the above $m \sim \overline{m}$. Also the converse holds, and thus the congruences of $(\mathbb{Z},+)$ are exactly the ordinary number theoretic congruences, $\sim$ equals $\bmod k$ ($k \geq 0$). □

We prove now that the congruences of a semigroup S are closed under intersections.

Lemma 2.2. *(1) Let $\rho_i \in \mathrm{Con}(S)$ be a congruence of S for each $i \in I$. Then also the intersection $\bigcap_{i\in I}\rho_i$ is a congruence.*
(2) Let $\delta \subseteq S \times S$ be a relation. Then

$$\delta^c = \bigcap\{\rho \mid \delta \subseteq \rho,\ \rho \in \mathrm{Con}(S)\}$$

is the smallest congruence of S containing δ.

Proof. Denote $\rho = \cap_{i\in I}\rho_i$. Assume $x\rho y$ and $z \in S$, then also $x\rho_i y$ for all $i \in I$, and hence $(zx)\rho_i(zy)$ and $(xz)\rho_i(yz)$, which implies that $(zx)\rho(zy)$ and $(xz)\rho(yz)$. Therefore ρ is a congruence of S.

For the second claim, observe first that δ^c is a congruence, since it is an intersection of congruences. If ρ is any congruence with $\delta \subseteq \rho$, then ρ takes part in the intersection, and thus $\delta^c \subseteq \rho$. The claim follows from this. □

2.2 Homomorphism theorem

Natural homomorphisms

For a congruence $\rho \in \mathrm{Con}(S)$, define a mapping $\rho^{\mathrm{nat}}\colon S \to S/\rho$ as follows

$$\rho^{\mathrm{nat}}(x) = x\rho\,.$$

Theorem 2.1. *Let $\rho \in \mathrm{Con}(S)$. The mapping ρ^{nat} is an epimorphism, called the **natural epimorphism**.*

Proof. Let $x, y \in S$. Now, $\rho^{\mathrm{nat}}(xy) = (xy)\rho = x\rho\cdot y\rho = \rho^{\mathrm{nat}}(x)\rho^{\mathrm{nat}}(y)$, and hence ρ^{nat} is a homomorphism. On the other hand, if $u = x\rho \in S/\rho$, then clearly $\rho^{\mathrm{nat}}(x) = u$, and therefore ρ^{nat} is surjective. □

For a homomorphism $\alpha\colon S \to P$, we define its **kernel** as the relation

$$\ker(\alpha) = \{(x, y) \mid \alpha(x) = \alpha(y)\} = \alpha^{-1}\alpha\,.$$

Lemma 2.3. *For each homomorphism $\alpha\colon S \to P$, $\ker(\alpha) \in \mathrm{Con}(S)$.*

Proof. We leave it as an exercise to show that $\ker(\alpha)$ is an equivalence relation. Let $(x, y) \in \ker(\alpha)$, that is, $\alpha(x) = \alpha(y)$. Now, for all $z \in S$,

$$\alpha(zx) = \alpha(z)\alpha(x) = \alpha(z)\alpha(y) = \alpha(xy)\,,$$

and similarly, $\alpha(xz) = \alpha(yz)$. This proves the claim. □

Lemma 2.4. *Let $\rho \in \mathrm{Con}(S)$. Then $\rho = \ker(\rho^{\mathrm{nat}})$.*

Proof. Indeed,

$$x\rho y \iff x\rho = y\rho \iff \rho^{\mathrm{nat}}(x) = \rho^{\mathrm{nat}}(y) \iff (x, y) \in \ker(\rho^{\mathrm{nat}})\,.$$

□

Corollary 2.1. *Every congruence is a kernel of some homomorphism.*

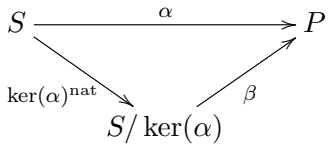
Example 2.2. Often homomorphisms are easier to handle that congruences. Consider $S = (\mathbb{N}, +)$. If $\alpha\colon S \to P$ is a homomorphism, then

$$\alpha(n) = \alpha(1 + 1 + \cdots + 1) = \alpha(1) + \alpha(1) + \cdots + \alpha(1) = n\alpha(1)\,.$$

In particular, $\alpha(0) = 0$. This implies that α depends only on 1. Of course, this follows already from the fact that 1 generates $(\mathbb{N}, +)$. It is now easy to show that $\ker(\alpha)$ equals $\bmod k$ for some k. □

Homomorphism theorem

Theorem 2.2. *Let $\alpha\colon S \to P$ be any homomorphism. There exists a unique embedding $\beta\colon S/\ker(\alpha) \hookrightarrow P$ such that $\alpha = \beta \circ \ker(\alpha)^{\mathrm{nat}}$.*



Proof. Denote $\rho = \ker(\alpha)$, and let $\rho^{\mathrm{nat}}\colon S \to S/\rho$ be the corresponding natural homomorphism. Define $\beta\colon S/\rho \to P$ by

$$\forall x \in S\colon\ \beta(x\rho) = \alpha(x)\,.$$

(1) β is a well defined function, since

$$x\rho = y\rho \iff (x,y) \in \ker(\alpha) \iff \alpha(x) = \alpha(y) \iff \beta(x\rho) = \beta(y\rho) \qquad (2.1)$$

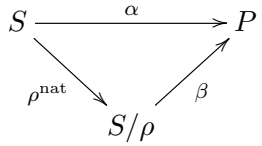
and hence the value $\beta(x\rho)$ is independent of the choice of the representative of the congruence class $x\rho$.
(2) β is a homomorphism, since

$$\beta(x\rho \cdot y\rho) = \beta((xy)\rho) = \alpha(xy) = \alpha(x)\alpha(y) = \beta(x\rho)\beta(y\rho)\,.$$

(3) β is injective by (2.1).
Finally, β is unique, since if $\gamma\colon S/\rho \to P$ is another embedding, then $\alpha = \gamma\rho^{\mathrm{nat}}$, and hence $\alpha(x) = \gamma(x\rho)$ for all $x \in S$. But this means that $\gamma = \beta$. □

The previous theorem has the following improvement.

Theorem 2.3 (Homomorphism theorem). *Let $\alpha\colon S \to P$ homomorphism and let $\rho \in \mathrm{Con}(S)$ be such that $\rho \subseteq \ker(\alpha)$. Then there exists a unique homomorphism $\beta\colon S/\rho \to P$ such that the following diagram commutes.*



Proof. The proof is similar to the previous one. Here we observe that the mapping β defined by $\beta(x\rho) = \alpha(x)$ is well defined, since $\rho \subseteq \ker(\alpha)$, and hence

$$x\rho = y\rho \implies x\rho y \implies (x,y) \in \ker(\alpha) \implies \alpha(x) = \alpha(y)\,.$$

□

The homomorphism theorem, as well as the next isomorphism theorem, are standard (universal) algebraic results, that is, they hold in all algebras (groups, rings, Boolean algebras, and so forth).

Theorem 2.4 (Isomorphism theorem). *Let $\alpha\colon S \to P$ be a homomorphism. Then*

$$\alpha(S) \cong S/\ker(\alpha)\,.$$

Proof. The homomorphism α is an epimorphism $S \to \alpha(S) \leq P$. When Theorem 2.2 is applied to $\alpha\colon S \to \alpha(S)$ we obtain a unique embedding $\beta\colon S/\ker(\alpha) \to \alpha(S)$. Moreover, β is surjective, since α maps S onto $\alpha(S)$ and $\alpha = \beta\gamma$ for the homomorphism $\gamma = \ker(\alpha)^{\mathrm{nat}}$. Consequently, β is a bijective homomorphism, i.e., an isomorphism. □

2.3 Ideals

Ideals in semigroups

A nonempty subset I of a semigroup S is a **left ideal** (resp., **right ideal**), if $SI \subseteq I$; (resp., $IS \subseteq I$). Also, I is an **ideal**, if it is both a left and a right ideal.

Lemma 2.5. *A nonempty subset $I \subseteq S$ is*

(1) a left ideal of S, if for all $a \in I$ and $x \in S$, $sa \in I$;
(2) a right ideal of S, if for all $a \in I$ and $x \in S$, $as \in I$;
(3) an ideal of S, if for all $a \in I$ and $x \in S$, $as, sa \in I$.

Proof. Exercises. □

Lemma 2.6. *If I is a (left or a right) ideal of S, then I is a subsemigroup of S.*

Proof. If $SI \subseteq I$ or $IS \subseteq I$, then certainly $II \subseteq I$, since $I \subseteq S$. □

Lemma 2.7. *If I and J are left (right) ideals of a semigroup S with $I \cap J \neq \emptyset$, then $I \cap J$ is a left (right) ideal.*

Proof. Let $a \in I \cap J$ and $x \in S$. Then $sa \in I$ and $sa \in J$, since I and J are ideals, and so $sa \in I \cap J$. Hence $(I \cap J)S \subseteq (I \cap J)$. □

Example 2.3. (1) A group G has only the trivial ideals $\{1\}$ and G. Indeed, if I is an ideal and $g \in G$, then $g = ga^{-1}a \in I$ for each $a \in I$.

(2) Let I be an ideal of S. Define the **Rees congruence** (with respect to I) as follows:

$$\rho_I = (I \times I) \cup \iota\,.$$

Hence $x\rho_I y$ holds if and only if either $x, y \in I$ or $x = y$. This means that ρ_I contracts the ideal I and leaves the rest of S as it was. If I is an ideal of S, then $\rho_I \in \mathrm{Con}(S)$.

The quotient semigroup S/ρ_I will be denoted simply S/I, and it is called the **Rees quotient** (of the ideal I). S/I has the elements I and $\{x\}$ for $x \in S \setminus I$. In order to simplify the notation, we identify the elements $\{x\}$ $(= x\rho_I)$ with the corresponding elements $x \in S \setminus I$. The product of elements in S/I is as follows: $x \cdot y = xy$ for $x, y \notin I$, and $Ix = I = xI$ for all x. Therefore I is a zero element of the semigroup S/I. □

An ideal I of S is **minimal**, if for all ideals J of S, $J \subseteq I$ implies that $J = I$.

Lemma 2.8. *If I is a minimal ideal, and J is any ideal of S, then $I \subseteq J$.*

Proof. First of all, $I \cap J \neq \emptyset$. Indeed, by the definition of an ideal $IS \subseteq I$, and hence also $IJ \subseteq I$. Similarly, $IJ \subseteq J$, and so $IJ \subseteq I \cap J$. Here, of course, $IJ \neq \emptyset$. Now, $I \cap J \subseteq I$ implies that $I \cap J = I$, since I is a minimal ideal, and $I \cap J$ is an ideal. It follows that $I \subseteq J$ as required. □

By the above,

Theorem 2.5. *If a semigroup S has a minimal ideal, then it is unique.*

Every *finite* semigroup S does have a minimal ideal. Indeed, consider an ideal I, which has the least number of elements. Such an ideal exists because S is finite and S is its own ideal. By Theorem 2.5, it is unique.

Example 2.4. The ideals of the semigroup $(\mathbb{N}, +)$ are the sets $n + \mathbb{N} = \{n + k \mid k \geq 0\}$ (Exercise). Further, $m + \mathbb{N} \subseteq n + \mathbb{N}$ if and only if $m \geq n$. Therefore $(\mathbb{N}, +)$ has no minimal ideals. □

Simple semigroups

A semigroup S is said to be **simple**, if it has no (proper) ideals $I \neq S$.

Lemma 2.9. *S is simple if and only if $S = SxS$ for all $x \in S$.*

Proof. Clearly, for any $x \in S$, SxS is an ideal of S, and hence if S is simple, then $S = SxS$ for all $x \in S$.

In converse, assume $S = SxS$ for all $x \in S$. Let I is an ideal of S and let $x \in I$. Hence $S = SxS \subseteq I$, and so $I = S$, from which it follows that S is simple. □

Example 2.5. By the above, a semigroup S is simple if and only if for all $s, r \in S$ the equation $xsy = r$ has a solution in S. Using this, one can show that the semigroup of all 2×2 matrices

$$\begin{pmatrix} x & 0 \\ y & 1 \end{pmatrix} \qquad (x, y \in \mathbb{Q} \text{ with } x, y > 0)$$

is a simple semigroup. This is an exercise. □

3

Free Semigroups and Presentations

3.1 Free semigroups

Word semigroups

Let A be a set, called an **alphabet**, of symbols or **letters**. Any finite sequence of letters is a **word** (or a string) **over** A. The set of all words over A, with at least one letter, is denoted by A^+. The set A^+ is a semigroup, the **word semigroup** over A with respect to **catenation** of words, that is, the product of the words $w_1 = a_1a_2\cdots a_n$, $w_2 = b_1b_2\cdots b_m$ ($a_i, b_i \in A$) is the word $w_1 \cdot w_2 = w_1w_2 = a_1a_2\cdots a_nb_1b_2\cdots b_m$. When we adjoin the **empty word** ε (with no letters) to A^+, we obtain the **word monoid** $A^* = A^+ \cup \{\varepsilon\}$. Clearly, $\varepsilon\cdot w = w = w\cdot\varepsilon$ for all words $w \in A^*$.

Example 3.1. Let $A = \{a, b\}$ be a **binary** alphabet. Then

$$A^* = \{\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, \dots\}.$$

As usual, w^k means the catenation of w with itself k times. We have $w^0 = \varepsilon$, and, for example, $v = ab^3(ba)^2 = abbbbaba = ab^4aba$. □

Free semigroups

Let S be a semigroup. A subset $X \subseteq S$ **generates** S **freely**, if

- $S = [X]_S$ and,
- every mapping $\alpha_0\colon X \to P$ (with P is any semigroup) extends to a homomorphism $\alpha\colon S \to P$ such that $\alpha \restriction X = \alpha_0$.

Here we say that α is a **morphic extension** of α_0. If S is freely generated by some subset, then S is a **free semigroup**.

Example 3.2. (1) The additive semigroup $(\mathbb{N}_+, +)$ is free, for $X = \{1\}$ generates it freely: Let $\alpha_0\colon X \to P$ be a homomorphism, and define $\alpha\colon \mathbb{N}_+ \to P$ by $\alpha(n) = \alpha_0(1)^n$. Now, $\alpha \restriction X = \alpha_0$, and α is a homomorphism: $\alpha(n+m) = \alpha_0(1)^{n+m} = \alpha_0(1)^n\cdot\alpha_0(1)^m = \alpha(n)\cdot\alpha(m)$.

(2) On the other hand, the multiplicative semigroup $(\mathbb{N}_+, \cdot)$ is *not* free. For, suppose $X \subseteq \mathbb{N}_+$, choose $P = (\mathbb{N}_+, +)$, and let $\alpha_0(n) = n$ for all $n \in X$. If $\alpha\colon (\mathbb{N}_+, \cdot) \to P$ is any homomorphism, then $\alpha(n) = \alpha(1\cdot n) = \alpha(1) + \alpha(n)$, and thus $\alpha(1) = 0 \notin P$. So certainly no α can be an extension of α_0.

(3) Let $S = \{a, a^2\}$ be a cyclic semigroup with $a^3 = a^2$. If X generates S, then $a \in X$, since a is not a product of two elements of S. Let $\alpha_0 \colon S \to P = (\mathbb{N}_+, +)$ be such that $\alpha_0(a) = 1$. If $\alpha \colon S \to P$ is an extension of α_0, then $\alpha(a^2) = \alpha(a) + \alpha(a) = \alpha_0(a) + \alpha_0(a) = 2$ and similarly $\alpha(a^3) = 3$. However, $a^2 = a^3$ in S, and this is a contradiction. Thus S is not free. □

Theorem 3.1. *For any alphabet A, A^+ is a free semigroup freely generated by A.*

Proof. Clearly, A generates A^+. Let then S be any semigroup, and $\alpha_0 \colon A \to S$ a mapping. Define $\alpha \colon A^+ \to S$ by

$$\alpha(a_1 a_2 \cdots a_n) = \alpha_0(a_1)\alpha_0(a_2) \cdots \alpha_0(a_n) \qquad (n \geq 1, a_i \in A)\,.$$

Clearly, $\alpha \restriction A = \alpha_0$, and α is a homomorphism by its definition. □

Theorem 3.2. *If S is freely generated by X and $\alpha_0 \colon X \to P$ is a mapping to a semigroup P, then α_0 has a unique morphic extension $\alpha \colon S \to P$.*

Proof. By the definition, each α_0 has an extension. For uniqueness, suppose both $\alpha \colon S \to P$ and $\beta \colon S \to P$ are morphic extensions of α_0. Now, for all $x \in S$, $x = a_1 a_2 \cdots a_n$ for some $a_i \in X$, since X generates S. Therefore,

$$\begin{aligned}\alpha(x) =& \alpha(a_1)\alpha(a_2) \cdots \alpha(a_n) = \alpha_0(a_1)\alpha_0(a_2) \cdots \alpha_0(a_n) \\ &= \beta(a_1)\beta(a_2) \cdots \beta(a_n) = \beta(x)\,,\end{aligned}$$

and hence $\alpha = \beta$. □

The next result states that every semigroup S is a homomorphic image of free semigroup, and therefore the free semigroups are the 'basic' semigroups wherefrom all other semigroups can be derived.

Theorem 3.3. *For each semigroup S, there exists an alphabet A, possibly infinite, and an epimorphism $\psi \colon A^+ \twoheadrightarrow S$.*

Proof. Let X be any generating set of S; you may even choose $X = S$. Let A be an alphabet with $|A| = |X|$, and let $\psi_0 \colon A \to X$ be a bijection. Since A^+ is free, ψ_0 has a morphic extension $\psi \colon A^+ \to S$. This extension is surjective, since X generates S and hence $[\psi(X)]_S = \psi([X]_S) = \psi(S)$. □

Corollary 3.1. *Every semigroup is a quotient of a free semigroup. Indeed,*

$$S \cong A^+ / \ker(\psi)$$

for a suitable epimorphism $\psi \colon A^+ \twoheadrightarrow S$.

There are other free semigroups than the word semigroups, but all of them are isomorphic to word semigroups.

Theorem 3.4. *A semigroup S is free if and only if $S \cong A^+$ for some alphabet A.*

Proof. Suppose S is freely generated by a subset $X \subseteq S$, and let A be an alphabet with $|A| = |X|$. Let $\psi_0 \colon A \to X$ be a bijection. Since A generates A^+ freely there is an epimorphic extension $\psi \colon A^+ \twoheadrightarrow S$ as in the previous theorem. Also, the mapping $\psi_0^{-1} \colon X \to A$ is a bijection, which has an epimorphic extension $\beta \colon S \twoheadrightarrow A^+$, since S is freely generated by X. The composition $\beta\psi \colon A^+ \to A^+$ is an epimorphism, for which

$$\beta\psi \restriction A = \beta\psi_0 = (\beta \restriction X)\psi_0 = \psi_0^{-1}\psi_0 = \iota_A \,.$$

Now, the identity mapping $\iota_A \colon A \to A$ extends uniquely to the identity homomorphism $\iota_{A^+} \colon A^+ \to A^+$, and hence $\beta\psi \restriction A$ extends also uniquely to ι_{A^+}, that is, $\beta\psi = \iota_{A^+}$. It follows that $\beta = \psi^{-1}$ and thus ψ is a bijection, and hence an isomorphism.

Conversely, suppose that there exists an isomorphism $\psi \colon A^+ \to S$. In this case, $S = [\psi(A)]_S$, and ψ has an inverse mapping $\psi^{-1} \colon S \to A^+$, which is a homomorphism (an isomorphism, by Lemma 1.4). Denote $\psi_0 = \psi \restriction A$ and $X = \psi(A)$. Let P be any semigroup, and $\alpha_0 \colon X \to P$ any mapping. Now, the mapping $\alpha_0\psi_0 \colon A \to P$ extends uniquely to a homomorphism $\gamma \colon A^+ \to P$. Consider the mapping $\beta = \gamma\psi^{-1} \colon S \to P$. This is a homomorphism, since both ψ^{-1} and γ are. Further, for each $x \in X$,

$$\beta(x) = \gamma(\psi^{-1}(x)) = \alpha_0\psi_0\psi_0^{-1}(x) = \alpha_0(x)$$

and hence $\beta \restriction X = \alpha_0$, that is, β is a morphic extension of α_0. By the definition, S is freely generated by X. □

The above proof reveals also

Corollary 3.2. *If S is freely generated by a set X, then $S \cong A^+$, where $|A| = |X|$.*

Corollary 3.3. *If S and R are free semigroups generated freely by X and Y, respectively, such that $|X| = |Y|$, then $S \cong R$.*

Corollary 3.4. *Every free semigroup is cancellative.*

Proof. This is immediate, since A^+ is cancellative. □

A criterion for freeness

Let $X \subseteq S$ for a semigroup S. We say that $x = x_1x_2\cdots x_n$ is a **factorization** of x over X, if $x_i \in X$ for each i. Now, if X generates S, then every element $x \in X$ has a factorization over X. Usually, this factorization is not unique, that is, we may have $x_1x_2\cdots x_n = x = y_1y_2\cdots y_m$ for some $x_i \in X$ $(i = 1, 2, \ldots, n)$ and for some different $y_i \in X$ $(i = 1, 2, \ldots, m)$.

Theorem 3.5. *A semigroup S is freely generated by X if and only if every $x \in S$ has a unique factorization over X.*

Proof. We observe first that the claim holds for the word semigroups A^+, for which A is the only minimal generating set. Let A be an alphabet such that $|A| = |X|$ and let $\alpha_0 \colon X \to A$ be a bijection.

Suppose X generates S freely, and

$$x_1 x_2 \cdots x_n = y_1 y_2 \cdots y_m \text{ for some } x_i, y_i \in X\,.$$

For the morphic extension α of α_0, we have

$$\alpha_0(x_1)\alpha_0(x_2)\cdots\alpha_0(x_n) = \alpha_0(y_1)\alpha_0(y_2)\cdots\alpha_0(y_m)$$

in A^+. Since the claim holds for A^+, and $\alpha_0(x_i)$, $\alpha(y_i) \in A$ each i, we must have that $\alpha_0(x_i) = \alpha_0(y_i)$ for all $i = 1, 2, \ldots, n$ and $m = n$. Moreover, α_0 is injective, and hence $x_i = y_i$ for all i, and thus also S satisfies the claim.

Suppose then that S satisfies the uniqueness condition. Denote $\beta_0 = \alpha_0^{-1}$ and let $\beta \colon A^+ \to S$ be the morphic extension of β_0. Now, β is surjective, since X generates S. It is also injective, because if $\beta(u) = \beta(v)$ for some words $u \neq v \in A^+$, then $\beta(u)$ would have two different factorizations over X. Hence β is an isomorphism, and the claim follows from this. □

Let the **base** of a semigroup S be defined by

$$\mathrm{Base}(S) = S \setminus S^2 = \{x \in S \mid \forall y, z \in S \colon\ x \neq yz\}\,,$$

consisting of those elements $x \in S$ which are not products of any two or more elements of S. The following result is an exercise.

Theorem 3.6. *A semigroup is free S if and only if* $\mathrm{Base}(S)$ *generates it freely.*

Example 3.3. (1) Let $A = \{a, b, c\}$ be an alphabet. The set $X = \{ab, bab, ba\}$ generates a subsemigroup S of A^+ that is *not free*, since the element $w = babab$ has two different factorizations over X: $w = ba{\cdot}bab = bab{\cdot}ab$.

(2) Let A be as above, but let S be generated by $X = \{aab, ab, aa\}$. Then S is a free. Indeed, if there were two different factorizations $u_1 u_2 \cdots u_n = v_1 v_2 \cdots v_m$ over X of a word, then either $u_1 = v_1$ (and, by cancellation, there would be a shorter word $u_2 \cdots u_n = v_2 \cdots v_m$ with two different factorizations) or $u_1 = aa$ and $v_1 = aab$ (or symmetrically, $u_1 = aab$ and $v_1 = aa$). But in this case u_2 cannot exist, because it should start with the letter b.

(3) Let then $A = \{a, b\}$ be a binary alphabet, and consider $S = [ab, ab^2, ab^3, \ldots]$. Clearly, $\mathrm{Base}(S) = S \setminus S^2 = \{ab^n \mid n \geq 1\}$, and S is freely generated by $\mathrm{Base}(S)$.

(4) For $S = (\mathbb{R}, \cdot)$, we have $\mathrm{Base}(S) = \emptyset$. In particular, this semigroup cannot be free. □

Free monoids

No monoid M can be a free semigroup, because $1_M = 1_M \cdot 1_M$. We say that M is a **free monoid** freely generated by a subset X with $1_M \notin X$, if $X \cup \{1_M\}$ generates M, and every mapping $\alpha_0 \colon X \to P$ (where P is a monoid) extends to a monoid homomorphism $\alpha \colon M \to P$ such that $\alpha \restriction X = \alpha_0$ and $\alpha(1_M) = 1_P$.

Again, if M is free and $\alpha_0 \colon X \to P$ a mapping to a monoid P, then its extension $\alpha \colon M \to P$ is unique.

We have immediately:

Theorem 3.7. *If S is a free semigroup, then S^1 is a free monoid.*

Corollary 3.5. *The word monoid A^* is a free monoid for all alphabets A.*

The converse of Theorem 3.7 holds, too.

Theorem 3.8. *A monoid M is a free monoid if and only if $M \setminus \{1_M\}$ is a free semigroup.*

Proof. Suppose first that M is a free monoid, freely generated by X. Then $M \setminus \{1_M\}$ is a subsemigroup of M, for otherwise, $1_M = x_1 x_2 \dots x_n$ for some $x_i \in X$, and a mapping $\alpha_0 \colon X \to A^*$ with $\alpha_0(x_i) \in A$ does not extend. The rest of the claim follows from the definition of a free semigroup.

The converse claim is an exercise. □

The other results for free semigroups can also be modified for free monoids:

Theorem 3.9. *(1) A monoid M is freely generated by X if and only if each nonidentity $x \in M \setminus \{1_M\}$ has a unique factorization over X.*

(2) Every monoid is an epimorphic image of a word monoid A^*.

(3) A monoid M is free if and only if it is isomorphic to A^* for some alphabet A.

A free matrix monoid

Let $SL(2, \mathbb{N})$ be the set of all matrices $M \in \mathbb{N}^{2\times 2}$ with $\det(M) = 1$, that is,

$$M = \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix}, \qquad n_{11}n_{22} - n_{12}n_{21} = 1 \qquad (n_{ij} \in \mathbb{N}). \tag{3.1}$$

This is a monoid, called the **special linear monoid**. The identity element is the identity matrix I. Note that if $\det(M_1) = 1 = \det(M_2)$, then also $\det(M_1 M_2) = 1$.

Theorem 3.10 (Nielsen). *$SL(2, \mathbb{N})$ is a free monoid, freely generated by the matrices*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Proof. The inverses of the matrices A and B are

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

They have integer entries, but are not in $SL(2,\mathbb{N})$.

Let M be as in (3.1) with $M \neq I$. Then

$$A^{-1}M = \begin{pmatrix} n_{11}-n_{21} & n_{12}-n_{22} \\ n_{21} & n_{22} \end{pmatrix} \text{ and } B^{-1}M = \begin{pmatrix} n_{11} & n_{12} \\ n_{21}-n_{11} & n_{22}-n_{12} \end{pmatrix}.$$

Either one of these is in $SL(2,\mathbb{N})$: $A^{-1}M \in SL(2,\mathbb{N})$, if $n_{11} \geq n_{21}$, and $B^{-1}M \in SL(2,\mathbb{N})$, if $n_{11} < n_{21}$.

Moreover, the trace of M, $\operatorname{tr}(M) = n_{11} + n_{22}$, is greater than the trace of $A^{-1}M$ or $B^{-1}M$ whichever is in $SL(2,\mathbb{N})$:

$$\operatorname{tr}(A^{-1}M) = \operatorname{tr}(M) - n_{21} \text{ and } \operatorname{tr}(B^{-1}M) = \operatorname{tr}(M) - n_{12}\,.$$

Therefore, for each $M \in SL(2,\mathbb{N})$ with $M \neq I$, there exists a unique sequence $C_1, C_2, \ldots, C_k$ ($C_i = A$ or B) such that $C_k^{-1}C_{k-1}^{-1}\cdots C_1^{-1}M = I$, that is, $M = C_1C_2\cdots C_k$. By Theorem 3.5 and Theorem 3.8, $SL(2,\mathbb{N})$ is a freely generated by $\{A, B\}$. □

3.2 Presentations of semigroups

Generators and relations

Let S be a semigroup. By Theorem 3.3, there exists an epimorphism $\psi\colon A^+ \twoheadrightarrow S$, where A^+ is a suitable word semigroup. By Theorem 2.4,

$$S \cong A^+/\ker(\psi)\,.$$

We say that

$$S = \langle a_1, a_2, \cdots \mid u_i = v_i\ (i \in I)\rangle \quad \text{or} \quad S = \langle A \mid R\rangle\,,$$

is a **presentation** of S **in generators** $A = \{a_1, a_2, \ldots\}$ **and relations** $R = \{(u_i, v_i) \mid i \in I\}$, if $\ker(\psi)$ is the smallest congruence of A^+ that contains the relation R. In particular,

$$\psi(u) = \psi(v) \quad \text{for all} \quad (u,v) \text{ in } R\,. \tag{3.2}$$

The set R of relations is supposed to be symmetric, that is, if $(u,v) \in R$, then also $(v,u) \in R$, and we often write $u = v$ (in S) instead of $(u,v) \in R$.

One should remember that the words $w \in A^+$ are not elements of S but of the word semigroup A^+, which is mapped onto S. We say that a word $w \in A^+$ **presents** the

element $\psi(w)$ of S. The same element can be presented by many different words. In fact, if $\psi(u) = \psi(v)$, then both u and v present the same element of S.

Let $S = \langle A \mid R\rangle$ be a presentation. We say that a word v is **directly derivable** from the word u, if

$$u = w_1u'w_2 \quad \text{and} \quad v = w_1v'w_2 \quad \text{for some } (u', v') \in R\,. \tag{3.3}$$

Clearly, if v is derivable from u, then u is derivable from v, since R is supposed to be symmetric, and, in the notation of (3.3),

$$\psi(u) = \psi(w_1u'w_2) = \psi(w_1)\psi(u')\psi(w_2) = \psi(w_1)\psi(v')\psi(w_2) = \psi(w_1v'w_2) = \psi(v).$$

The word v is **derivable** from u, denoted by $u =_S v$ or simply $u = v$ (in S), if there exists a finite sequence $u = u_1, u_2, \ldots, u_k = v$ such that for all $j = 1, 2, \ldots, k-1$, u_{j+1} is directly derivable from u_j. Now, if v is derivable from u, then also $\psi(u) = \psi(v)$, since $\psi(u) = \psi(u_1) = \cdots = \psi(u_k) = \psi(v)$. This can be written as

$$u = u_1 = u_2 = \cdots = u_k = v\,.$$

We allow the case that an element u is derivable from itself.

Theorem 3.11. *Let $S = \langle A \mid R\rangle$ be a presentation (with R symmetric). Then*

$$R^c = \{(u, v) \mid \ v \ \textit{is derivable from} \ u\}\,.$$

Hence $u = v$ in S if and only if v is derivable from u.

Proof. Let the relation ρ be defined by

$$u\rho v \iff u = v \text{ as words or } v \text{ is derivable from } u\,.$$

Clearly $\iota_S \subseteq \rho$, and hence ρ is reflexive. It is also symmetric, since R is symmetric. The transitivity of ρ is easily verified, and hence ρ is an equivalence relation.

If $w \in A^+$ and v is derivable from u, then clearly also wv is derivable from wu and vw is derivable from uw. This proves that ρ is a congruence.

Let then θ be any congruence such that $R \subseteq \theta$. Suppose v is directly derivable from u, say $u = w_1u'w_2$ and $v = w_1v'w_2$ with $(u', v') \in R$. Since $R \subseteq \theta$, also $(u', v') \in \theta$, and since θ is a congruence, also $(w_1u'w_2, w_1v'w_2) \in \theta$, that is, $u\theta v$. Therefore, by transitivity of ρ, we have θ, $\rho \subseteq \theta$. In conclusion, ρ is the smallest congruence that contains R, that is, $\rho = R^c$. □

Theorem 3.12. *Let A be an alphabet and $R \subseteq A^+ \times A^+$ a symmetric relation. The semigroup $S = A^+/R^c$, where R^c is the smallest congruence containing R, has the presentation*

$$S = \langle A \mid u = v \ \textit{for all} \ (u, v) \in R\rangle\,.$$

Moreover, all semigroups having a common presentation are isomorphic.

Proof. The claim is immediate from the above considerations. □

Example 3.4. (1) Consider the following presentation

$$S = \langle a, b \mid aa = ab,\ ba = aab,\ bbb = aba \rangle .$$

In this presentation there are two generators and three relations. For instance, S satisfies the relation $baabbaa = bbaaaba$, since $u_1 = baabbaa = b{\cdot}aab{\cdot}baa = b{\cdot}ba{\cdot}baa = u_2$ and $u_2 = bbabaa = bba{\cdot}ba{\cdot}a = bba{\cdot}aab{\cdot}a = bbaaaba$. Also, $aaab = aabb = abbb = aaba = baa = bab$ in S and hence $aaab = bab$ in S.

(2) The free word semigroup A^+ does not need any relations: $A^+ = \langle A \mid \emptyset \rangle$.

(3) The presentation

$$S = \langle a, b, c, d, e \mid ac = ca,\ ad = da,\ bc = cb,\ bd = db,\ eca = ce,\ edb = de,\ cca = ccae \rangle$$

is called **Tzeitin's semigroup**. This surprisingly simple semigroup has *an undecidable word problem*, that is, there exists no algorithm that determines whether $u = v$ is a relation in this semigroup, where $u, v \in \{a, b, c, d, e\}^+$ are given (input) words. Hence in Tzeitin's semigroup one cannot effectively decide whether a word is derivable from another given word. □

All semigroups, and also monoids and groups, have presentations. Indeed, $S = \langle A \mid \ker(\psi) \rangle$ is one such presentation, when $\psi\colon A^+ \to S$ is the presenting epimorphisms. Usually this presentation is complicated. We are mostly interested in semigroups that have a **finite presentation**, that is, a presentation $S = \langle A \mid R \rangle$, where A is a finite alphabet and R is finite set of relations. However, not all semigroups have such presentations.

Monoid presentations

All monoids have a semigroup presentation, but it is more convenient to use monoid presentations for these in order to take advantage of the identity element:

$$M = \langle a_1, a_2, \ldots \mid u_i = v_i\ (i \in I) \rangle$$

is a **monoid presentation**, if $u_i, v_i \in A^*$ where $A = \{a_1, a_2, \ldots\}$ is an alphabet. In a monoid presentation we may thus have relations of the form $u = 1$, which means that the word u can be erased from another word or added somewhere in between two letters.

Example 3.5. (1) Let $M = \langle a, b \mid ab = ba \rangle$ be a monoid presentation. Hence $M \cong A^*/R^c$, where $A = \{a, b\}$ and R has the single relation $ab = ba$. There is an epimorphism $\psi\colon A^* \twoheadrightarrow M$ and M is generated by the elements $x = \psi(a)$ and $y = \psi(b)$. The monoid M is commutative, because of the relation $ab = ba$. *If the generators of M*

commute with each other, then M is commutative. Furthermore, each element $z \in M$ has a *normal form*: Suppose $z = z_1 z_2 \dots z_n$ with $z_i = \psi(a_i)$ ($a_i = a$ or b), and so

$$z = \psi(a_1)\psi(a_2)\dots\psi(a_n) = \psi(a_1 a_2 \dots a_n) = \psi(a^k b^m) = \psi(a)^k \psi(b)^m$$

for some $k, m \geq 0$. The monoid M is a **free commutative monoid**, and it can be shown that every commutative monoid generated by two elements is an epimorphic image of M.

(2) The (monoid) presentation

$$\langle a, b \mid aba = 1 \rangle$$

defines a group is isomorphic to $(\mathbb{Z}, +)$. Indeed, let M be a monoid with the above presentation, i.e., $M \cong A^*/R^c$, where $A = \{a, b\}$ and R consists of $aba = 1$, and let $\psi\colon A^* \twoheadrightarrow M$ be the corresponding epimorphism. Then M is generated by the elements $x = \psi(a)$ and $y = \psi(b)$. Furthermore, $ab = ab{\cdot}aba = aba{\cdot}ba = ba$, and hence $xy = yx$. It follows that M is a commutative monoid. Consequently, every element $z \in M$ has the form $z = \psi(a^n b^m)$ for some $n, m \geq 0$. Also, $a{\cdot}ba = 1$ and $ba{\cdot}a = ab{\cdot}a = 1$ means that ba is a group inverse of a. Similarly, a^2 is a group inverse of b, $b = (aa)^{-1}$. This means that all elements of M have the form $z = \psi(a^k)$ for $k \in \mathbb{Z}$. It is now easy to show that $\alpha\colon M \to (\mathbb{Z}, +)$ defined by $\alpha(a) = -1$ and $\alpha(b) = 2$ is an isomorphism.

(3) Define two mappings $\alpha, \beta\colon \mathbb{N} \to \mathbb{N}$ as follows:

$$\alpha(n) = \begin{cases} 0 & \text{if } n = 0 \\ n-1 & \text{if } n \geq 1 \end{cases} \qquad \text{and} \qquad \beta(n) = n + 1 \quad (n \geq 0)\,.$$

Consider the **bicyclic monoid** $\mathbb{B} = [\alpha, \beta]$ generated by these two transformations. We observe that $\alpha\beta = \iota$, but

$$\beta\alpha(n) = \begin{cases} 1 & \text{if } n = 0\,, \\ n & \text{if } n \geq 1\,, \end{cases}$$

and, more generally,

$$\beta^k\alpha^k(n) = \begin{cases} k & \text{if } n < k\,, \\ n & \text{if } n \geq k\,. \end{cases}$$

Let $A = \{a, b\}$ be an alphabet, and define a homomorphism $\psi\colon A^* \to \mathbb{B}$ by $\psi(a) = \alpha$ and $\psi(b) = \beta$. By the extension property, ψ becomes uniquely defined by the images of the generators a and b. Hence ψ is an epimorphism and $\mathbb{B} \cong A^+/\ker(\psi)$.

By above $ab = 1$ is a relation in $\mathbb{B}$. Let then $\gamma \in \mathbb{B}$ be any element of the bicyclic monoid, $\gamma = \gamma_n\gamma_{n-1}\dots\gamma_1$, where $\gamma_i = \alpha$ or $\gamma_i = \beta$. Since $\alpha\beta = \iota$, we may assume that if $\gamma_j = \beta$ for some index j, then $\gamma_t = \beta$ for all t with $j \leq t \leq n$. This shows that $\gamma = \beta^k\alpha^m$ for some $k, m \geq 0$, and hence

$$\mathbb{B} = \{\beta^k\alpha^m \mid k, m \geq 0\}\,.$$

Furthermore, these elements are all different from each other: if $\gamma = \beta^k\alpha^m$ and $\delta = \beta^r\alpha^s$, then

$$\gamma(0) = \beta^k\alpha^m(0) = \beta^k(0) = k \quad \text{and} \quad \delta(0) = \beta^r\alpha^s(0) = \beta^r(0) = r\,,$$

and for $n \geq \max\{m, s\}$,

$$\begin{aligned}\gamma(n) &= \beta^k\alpha^m(n) = \beta^k(n-m) = n+k-m\,,\\ \delta(n) &= \beta^r\alpha^s(n) = \beta^r(n-s) = n+r-s\,.\end{aligned}$$

Therefore, $\gamma = \delta$ just in case $k = r$ and $m = s$. This means that

$$\mathbb{B} = \langle a, b \mid ab = 1\rangle$$

is a (monoid) presentation of $\mathbb{B}$.

The bicyclic monoid has many (semigroup) presentations, of which we mention

$$\mathbb{B} = \langle a, b \mid aba = aab, a = aab, bab = abb, b = abb\rangle\,.$$

□

3.3 Embeddings into 2-generator semigroups

Evans' embedding result

The following embedding result was proved by Evans in 1952. The proof given here is due to Subbiah (1973).

Theorem 3.13 (Evans). *Let S be a semigroup generated by denumerably many elements. Then S can be embedded into a semigroup generated by two elements.*

We shall use the fact that each semigroup is isomorphic to a subsemigroup of the full transformation semigroup, and we actually prove

Theorem 3.14 (Sierpinski). *Let $\alpha_1, \alpha_2, \ldots : \mathbb{N}_+ \to \mathbb{N}_+$ be any transformations. There exist two transformations $\beta_1, \beta_2 \colon \mathbb{N}_+ \to \mathbb{N}_+$ such that each α_i is a composition of β_1 and β_2.*

Proof. Define

$$X_n = \{2^n(2m-1) - 1 \mid m = 1, 2, \ldots\}\,.$$

Now, $X_n \cap X_m = \emptyset$ for all $n \neq m$, because $k \in X_n$ if and only if n is the highest power for which 2^n divides $k+1$. Also,

$$\bigcup_{n\geq 1} X_n = 2\mathbb{N} + 1\,,$$

and hence the sets X_n form a partition of the odd positive integers. Define

$$\beta_1(n) = 2n \ (n \geq 1)\,, \qquad \beta_2(n) = \begin{cases} n-1 & \text{if } n \text{ is even,} \\ \alpha_k(2^{-(k+1)} \cdot (n+1) + 2^{-1}) & \text{if } n \in X_k. \end{cases}$$

Now, for all $k \geq 1$,

$$\alpha_k = \beta_2^2 \beta_1^k \beta_2 \beta_1\,,$$

since

$$\begin{aligned} \beta_2^2\beta_1^k\beta_2\beta_1(n) = \ & \beta_2^2\beta_1^k\beta_2(2n) = \beta_2^2\beta_1^k(2n-1) = \beta_2^2(2^k(2n-1)) = \\ & \beta_2(2^k(2n-1)-1) = \alpha_k(2^{-(k+1)}(2^k(2n-1)+2^{-1}) = \\ & \alpha_k(2^{-1}(2n-1+1) = \alpha_k(n)\,. \end{aligned}$$

This proves the claim. □

Theorem 3.13 follows from Sierpinski's result by using suitable isomorphisms. We omit this straightforward reduction.

Embedding word monoids

Let $A = \{a_1, a_2, \ldots, a_n\}$ be a finite alphabet, and let $B = \{a, b\}$ be a binary alphabet. The homomorphism $\alpha\colon A^* \to B^*$ defined by

$$\alpha(a_i) = ab^i \quad (i = 1, 2, \ldots, n)$$

is injective as can be easily shown. Therefore

Theorem 3.15. *Each finitely generated word monoid A^* can be embedded in a word monoid B^* generated by two letters.*

4

Green's Relations

4.1 Definitions

Introducing the relations

Let S be a semigroup, and define the following relations on S,

$$\begin{aligned} x\mathcal{L}y &\iff S^1x = S^1y \\ x\mathcal{R}y &\iff xS^1 = yS^1 \\ x\mathcal{J}y &\iff S^1xS^1 = S^1yS^1\,. \end{aligned}$$

Here $S^1x,\ xS^1$ and S^1xS^1 are the **principal left ideal**, the **principal right ideal** and the **principal ideal** generated by $x \in S$. By the definitions,

$$\begin{aligned} x\mathcal{L}y &\iff \exists s, s' \in S^1 \colon x = sy \text{ and } y = s'x\,, \\ x\mathcal{R}y &\iff \exists r, r' \in S^1 \colon x = yr \text{ and } y = xr'\,. \end{aligned}$$

As an exercise, we state

Lemma 4.1. *The relations $\mathcal{L}$, $\mathcal{R}$ and $\mathcal{J}$ are equivalence relations on S. In fact, $\mathcal{L}$ is a right congruence and $\mathcal{R}$ is a left congruence of S.*

We denote the corresponding equivalence classes containing x by

$$L_x = \{y \mid x\mathcal{L}y\}, \quad R_x = \{y \mid x\mathcal{R}y\}, \quad J_x = \{y \mid x\mathcal{J}y\}\,.$$

Example 4.1. (1) Consider the semigroup S from the table. Then

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

$S^1a = \{a, d\},\ S^1b = \{a, b, c, d\},\ S^1c = \{a, b, c, d\},\ S^1d = \{a, d\}$
$aS^1 = \{a\},\ bS^1 = \{a, b\},\ cS^1 = \{a, c\},\ dS^1 = \{d\}\,.$

The equivalence classes with respect to $\mathcal{L}$ are $L_a = \{a, d\} = L_d$ and $L_b = \{b, c\} = L_c$ and those with respect to $\mathcal{R}$, $R_a = \{a\}$, $R_b = \{b\}$, $R_c = \{c\}$ and $R_d = \{d\}$.

(2) Let T_X be the full transformation semigroup on X. Now, for $\alpha, \beta \in T_X$,

$$\alpha\mathcal{R}\beta \iff \exists \gamma, \gamma' \in T_X :\ \alpha = \beta\gamma \text{ and } \beta = \alpha\gamma'\,.$$

Therefore, $\alpha\mathcal{R}\beta$ implies that $\alpha(X) = \beta(X)$. On the other hand, if $\alpha(X) = \beta(X)$, then define $\gamma \in T_X$ as a function, for which $\gamma(x) =$ some $\overline{x}$ with $\beta(\overline{x}) = \alpha(x)$. Then, in the above notation, $\beta\gamma(x) = \beta(\overline{x}) = \alpha(x)$, and hence $\alpha \in \beta T^1_X$. Similarly, $\beta \in \alpha T^1_X$, and hence, $\alpha\mathcal{R}\beta$ if and only if $\alpha(X) = \beta(X)$. □

Theorem 4.1. *The relations $\mathcal{L}$ and $\mathcal{R}$ commute: $\mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L}$.*

Proof. Suppose that $(x, y) \in \mathcal{L} \circ \mathcal{R}$. This means that there exists an element $z \in S$ such that $x\mathcal{L}z$ and $z\mathcal{R}y$. Therefore there are elements $s, s', r, r' \in S$ such that

$$x = sz,\ z = s'x,\ z = yr,\ y = zr'\,.$$

Denote $t = szr'$. Then

$$\begin{aligned} t &= sz{\cdot}r' = xr', \\ x &= sz = syr = szr'r = tr\,, \end{aligned}$$

which means that $x\mathcal{R}t$. On the other hand,

$$\begin{aligned} t &= s{\cdot}zr' = sy, \\ y &= zr' = s'xr' = s'szr' = s't\,, \end{aligned}$$

which implies that $y\mathcal{L}t$. Since $x\mathcal{R}t$ and $t\mathcal{L}y$, also $(x, y) \in \mathcal{R} \circ \mathcal{L}$. We have shown that $\mathcal{L} \circ \mathcal{R} \subseteq \mathcal{R} \circ \mathcal{L}$. The inclusion in the other direction is proved in the same way. □

D- and H-relations

The previous result is important because it implies that the product

$$\mathcal{D} = \mathcal{L} \circ \mathcal{R}$$

is an equivalence relation.

Lemma 4.2. *$\mathcal{D}$ is the smallest equivalence relation that contains both $\mathcal{L}$ and $\mathcal{R}$.*

Proof. Clearly, $\mathcal{L} \subseteq \mathcal{D}$ and $\mathcal{R} \subseteq \mathcal{D}$, since $x\mathcal{L}x$ and $x\mathcal{R}x$ hold for all $x \in S^1$.

In order to show that $\mathcal{D}$ is an equivalence relation one needs to show that it is transitive: if $x\mathcal{D}z$ and $z\mathcal{D}y$, then also $x\mathcal{D}y$. This is an exercise.

In order to show that $\mathcal{D}$ is the smallest equivalence relation containing $\mathcal{L} \cup \mathcal{R}$ one assumes an equivalence relation $\mathcal{C}$ with $\mathcal{L} \cup \mathcal{R} \subseteq \mathcal{C}$, and shows that $\mathcal{D} \subseteq \mathcal{C}$. This also is an exercise. □

Let

$$\mathcal{H} = \mathcal{L} \cap \mathcal{R}\,.$$

Hence $\mathcal{H}$ is an equivalence relation, and it is the largest equivalence relation of S that is contained in both $\mathcal{L}$ and $\mathcal{R}$.

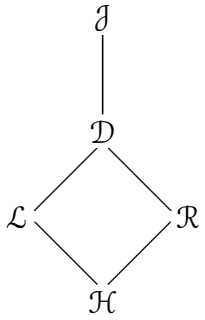


Fig. 4.1. The inclusion diagram of the Green's relations

The inclusion diagram of the **Green's relations** is given in Fig. 4.1, where one needs to observe that $\mathcal{D} \subseteq \mathcal{J}$ by Lemma 4.2.

We denote by D_x and H_x the corresponding equivalence classes that contain the element $x \in S$. Clearly, for all $x \in S$

$$H_x = L_x \cap R_x \,.$$

On the D-classes

Lemma 4.3. *For each semigroup S we have*

$$x \mathcal{D} y \iff L_x \cap R_y \neq \emptyset \iff L_y \cap R_x \neq \emptyset \,.$$

Moreover,

$$D_x = \bigcup_{y \in D_x} L_y = \bigcup_{y \in D_x} R_y \,.$$

Proof. By the definition of $\mathcal{D}$,

$$x \mathcal{D} y \iff \exists z \in S \, (x \mathcal{L} z \text{ and } z \mathcal{R} y) \iff \exists s \in S \, (x \mathcal{R} s \text{ and } s \mathcal{L} y) \,.$$

The first claim follows from this. For the second claim is trivial, since $\mathcal{L} \subseteq \mathcal{D}$ and $\mathcal{R} \subseteq \mathcal{D}$. □

The situation can be visualized by the **eggbox** picture of Fig. 4.2, where the rows are $\mathcal{R}$-classes and the columns are $\mathcal{L}$-classes. Their intersections are $\mathcal{H}$-classes, if nonempty (the intersection $L_y \cap R_z$ is nonempty, if $y \mathcal{D} z$). Indeed, if $u \in L_y \cap R_z$, then $L_y = L_u$ and $R_z = R_u$, and so $L_y \cap R_z = H_u$.

Example 4.2. Let $X = \{1, 2, \ldots, 7\}$, and consider the subsemigroup S of the full transformation semigroup T_X generated by α and β as defined in the table.

	1	2	3	4	5	6	7
α	2	2	5	6	5	6	5
β	3	4	3	4	7	4	7

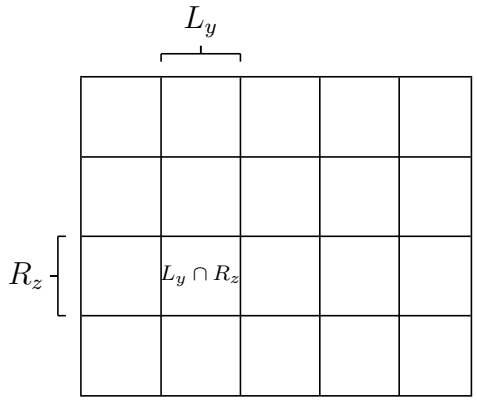


Fig. 4.2. The eggbox visualization of a $\mathcal{D}$-class D_x

This semigroup consists of six elements α, β, $\alpha\beta$, $\beta\alpha$, $\alpha\beta\alpha$ and $\beta\alpha\beta$. Computation shows that $\beta\alpha\beta\alpha = \beta\alpha$, and hence $\beta\alpha\mathcal{R}\beta\alpha\beta$. Also, $\alpha\beta\mathcal{R}\alpha\beta\alpha$ holds, as do $\beta\alpha\mathcal{L}\alpha\beta\alpha$ and $\alpha\beta\mathcal{L}\beta\alpha\beta$.

We conclude (after some calculations) that the eggbox for the class $D_{\beta\alpha}$ is as in Fig. 4.3. □

$\alpha\beta$	$\alpha\beta\alpha$
$\beta\alpha\beta$	$\beta\alpha$

Fig. 4.3. Eggbox for $D_{\beta\alpha}$

4.2 Green's Lemma and corollaries

Green's lemma

Let S be a semigroup. For each $s \in S$, define the mapping $\rho_s \colon S^1 \to S^1$ by

$$\forall z \in S^1 : \ \rho_s(z) = sz\,.$$

The usefulness of the Green's relations are due to the following version of **Green's lemma**.

Lemma 4.4. *Let S be a semigroup, $x, y \in S$ such that $x\mathcal{L}y$, and let $s, s' \in S^1$ be such that $sx = y$ and $s'y = x$. Then*

(1) $\rho_s \colon R_x \to R_y$ *is a bijection, and* $\rho_{s'} \colon R_y \to R_x$ *is a bijection,*
(2) $\rho_{s'} = \rho_s^{-1}$ *is the inverse function of* ρ_s *restricted to* R_x *,*
(3) ρ_s *fixes the* $\mathcal{L}$*-classes, that is,* $z\mathcal{L}\rho_s(z)$ *for all* $z \in R_x$ *,*
(4) ρ_s *preserves* $\mathcal{H}$*-classes, that is, for all* $u, v \in R_x$*:* $u\mathcal{H}v \iff \rho_s(u)\mathcal{H}\rho_s(v)$ *.*

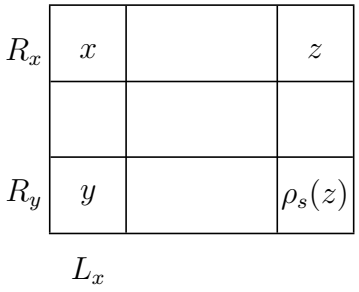


Fig. 4.4. Green's lemma

Proof. First we have to prove that ρ_s maps R_x into R_y. For this, let $z \in R_x$, that is, $zS^1 = xS^1$. We have now

$$szS^1 = sxS^1 = yS^1 \,,$$

which shows that $\rho_s(z) = sz \in R_y$ as required. A symmetric argument shows that $\rho_{s'}$ maps R_y into R_x.

If $z \in R_x$, then $z\mathcal{R}x$ and therefore there are elements $u, u' \in S^1$ such that $z = xu$ and $x = zu'$. Now,

$$s'sz = s's{\cdot}xu = s'{\cdot}sx{\cdot}u = s'yu = xu = z \,,$$

and hence

$$\forall z \in R_x : \; s'sz = z \,. \tag{4.1}$$

Hence

$$\rho_{s'}\rho_s(z) = \rho_{s'}(sz) = s'sz = z \,,$$

and so $\rho_{s'}\rho_s$ is the identity mapping of R_x. Similarly, $\rho_s\rho_{s'}$ is the identity mapping of R_y, from which we conclude Claims (1) and (2).

For Case (3) assume $z \in R_x$. By (4.1), the elements z and $\rho_s(z)$ $(= sz)$ are in the same $\mathcal{L}$-class.

For Case (4) we notice that $u\mathcal{H}v$ if and only if $u\mathcal{L}v$ and $u\mathcal{R}v$, and hence

$$u\mathcal{H}v \implies \rho_s(u)\mathcal{L}\rho_s(v) \text{ and } \rho_s(u)\mathcal{R}\rho_s(v) \implies \rho_s(u)\mathcal{H}\rho_s(v)$$

by Case (3) and since $\mathcal{R}$ is a left congruence (and $\rho_s(u) = su$, $\rho_s(v) = sv$). On the other hand, if $\rho_s(u)\mathcal{H}\rho_s(v)$, that is, $su\mathcal{H}sv$, then $u\mathcal{H}v$, since, by equation (4.1), $u = s'su$ and $v = s'sv$ (and since $\mathcal{R}$ is a left congruence and $\rho_{s'}$ fixes the $\mathcal{L}$-classes). □

Thus, ρ_s maps each $\mathcal{H}$-class H_z (with $z \in R_x$) bijectively onto the $\mathcal{H}$-class $H_{\rho_s(z)}$.

The next dual version of Lemma 4.4 is proved similarly. Here $\lambda_r \colon S^1 \to S^1$ is the opposite version of ρ_r:

$$\forall z \in S : \ \lambda_r(z) = zr \, .$$

Lemma 4.5. *Let S be a semigroup, $y\mathcal{R}z$, and let $r, r' \in S^1$ be such that $yr = z$ and $zr' = y$. Then*

(1) $\lambda_r \colon L_y \to L_z$ is a bijection, and $\lambda_{r'} \colon L_z \to L_y$ is a bijection,
(2) $\lambda_{r'} = \lambda_r^{-1}$ is the inverse function of λ_r restricted to L_y, and
(3) λ_r preserves the $\mathcal{R}$-classes, that is, $w\mathcal{R}\lambda_r(w)$ for all $w \in L_y$.

Lemma 4.6. *Let $e \in E_S$ be an idempotent. If $x\mathcal{L}e$, then $xe = x$. If $x\mathcal{R}e$, then $ex = x$.*

Proof. If $x\mathcal{L}e$, then $x \in L_e$, that is, $x = se$ for some $s \in S^1$. Therefore, since $e = ee$, $x = se \cdot e = xe$. The other case is similar. □

Corollary 4.1. *Each $\mathcal{H}$-class contains at most one idempotent.*

The sizes of an H-class

In the above notation, ρ_s is a bijection $R_x \to R_y$ and λ_r is a bijection $L_y \to L_z$ and therefore

Corollary 4.2. *The $\mathcal{H}$-classes inside a $\mathcal{D}$-class have the same cardinality, that is, if $x\mathcal{D}y$, then there exists a bijection between H_x and H_y.*

Proof. Indeed, if $x, z \in S$ are in the same $\mathcal{D}$-class such that $x\mathcal{L}y$, $y\mathcal{R}z$ with

$$sx = y, \ s'y = x, \ yr = z, \ zr' = y \, ,$$

then by the above lemmas $\rho_s \colon H_x \to H_y$ and $\lambda_r \colon H_y \to H_z$ are bijections. Therefore $\lambda_r \rho_s \colon H_x \to H_z$ is a bijection. □

Green's theorem for H-classes

The following crucial result is the **location theorem** of Miller and Clifford (1956).

Theorem 4.2. *Let $x, y \in S$. Then*

$$xy \in R_x \cap L_y \iff R_y \cap L_x \ \textit{contains a unique idempotent}.$$

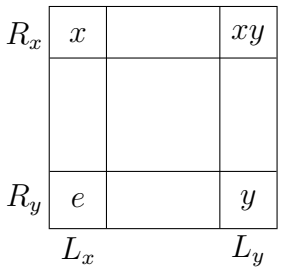


Fig. 4.5. Location theorem

Proof. Suppose first that $xy \in R_x \cap L_y$. Since $y\mathcal{L}xy$, we may choose $s = x$ in Lemma 4.4, and hence $\rho_x \colon R_y \to R_{xy}$ is a bijection. Also $xy\mathcal{R}x$, and hence $R_{xy} = R_x$, which means that $\rho_x \colon R_y \to R_x$ is a bijection. The mapping ρ_x preserves $\mathcal{L}$-classes and so ρ_x maps $R_y \cap L_x$ onto $R_x \cap L_x = H_x$. Therefore there exists a $z \in R_y \cap L_x$ such that $\rho_x(z) = x$, that is, $xz = x$.

Because $z\mathcal{L}x$, there exists a $u \in S^1$ with $z = ux$. We have thus obtained $xux = xz = x$, and thus $zz = uxux = ux = z$, which means that $z \in E_S$.

In the other direction, suppose there exists an idempotent $e \in R_y \cap L_x$. Now, by Lemma 4.6, $ey = y$ and $xe = x$. From $e\mathcal{R}y$ we obtain that $xe\mathcal{R}xy$, and thus $x\mathcal{R}xy$. From $e\mathcal{L}x$ we obtain that $ey\mathcal{L}xy$, and thus $y\mathcal{L}xy$. So $xy \in R_x \cap L_y$ as required.

Finally, if $R_y \cap L_x$ contains an idempotent, it must be unique by Corollary 4.1, since $R_y \cap L_x$ is an $\mathcal{H}$-class. □

In the following **Green's theorem**, G is a **subgroup** of a semigroup S, if G is a subsemigroup, which is a group.

We need the following lemma; see Fig. 4.6.

Lemma 4.7. *Let $e, f \in E_S$. Then for all $x \in R_e \cap L_f$, there exists a unique $y \in R_f \cap L_e$ such that $xy = e$ and $yx = f$.*

Proof. Let $x \in R_e \cap L_f$. By Lemma 4.6, $x = ex = xf$, and there are $u, v \in S^1$ for which $e = xu$ and $f = vx$. Now, $y = fu$ is the required element of the claim. Indeed,

$$f = vx = vex = vxux = fux = yx \tag{4.2}$$

and

$$e = xu = xfu = xy\,. \tag{4.3}$$

Hence, $y \in R_f$ by (4.2) and the fact that $y = f \cdot u$; and $y \in L_e$, by (4.3) and since $y = fu = vxu = ve$. Therefore $y \in R_f \cap L_e$.

For the uniqueness, assume also y' satisfies the conclusion. Then

$$y' = y'e = y'xy = fy = y$$

using Lemma 4.6 twice. □

$R_x = R_e$	x		$e = xy$
R_f	$f = yx$		y
	$L_x = L_f$		L_e

Fig. 4.6. The eggbox of D_x for a regular x

Theorem 4.3. *Let H be an $\mathcal{H}$-class of S. Then the following are equivalent:*

(1) H contains an idempotent.
(2) There exist $x, y \in H$ such that $xy \in H$.
(3) H is a subgroup of S.

Proof. (1) implies (2), since we can choose $x = e = y$ for the idempotent $e \in H$.

Assume (2). Then (1) follows by Theorem 4.2, since now $R_x \cap L_y = H = R_y \cap L_x$. Therefore there exists an $e \in E_H$. Hence, by the converse claim of Theorem 4.2, we know that H is a subsemigroup of S. By Lemma 4.6, H is a monoid with identity e. We apply then Lemma 4.7 for $e = f$. This shows that H is a group, and (2) implies (3).

(3) implies (1), since the identity element of the group H is an idempotent of S. □

Example: periodic semigroups

We say that a semigroup S is **periodic**, if each of its elements has finite order, that is, the cyclic subsemigroup $[x]$ is finite for all $x \in S$.

As an exercise we have

Lemma 4.8. *If S is periodic, then each $[x]$ contains an idempotent.*

Theorem 4.4. *For periodic semigroups $\mathcal{J} = \mathcal{D}$.*

Proof. First, $\mathcal{D} \subseteq \mathcal{J}$ is valid for all semigroups. Thus we need to show that $\mathcal{J} \subseteq \mathcal{D}$.

Suppose that $x \mathcal{J} y$. Therefore there exist $u, v, r, s \in S^1$ such that

$$x = uyv \quad \text{and} \quad y = rxs\,.$$

Now,

$$\forall i \geq 0 : \; x = (ur)^i x (sv)^i \quad \text{and } y = (ru)^i y (vs)^i\,.$$

By Lemma 4.8, there exists a nonnegative n such that $(ur)^n \in E_S$. We may suppose that $ur \neq 1$, for otherwise already $x \mathcal{R} y$ and thus $x \mathcal{D} y$, since $\mathcal{R} \subseteq \mathcal{D}$.

Denote $z = rx$, for short. We have then

$$x = (ur)^n x (sv)^n = (ur)^n (ur)^n x (sv)^n = (ur)^n x = (ur)^{n-1} uz\,,$$

and hence $x\mathcal{L}z$, since $z = r\cdot x$. Further, $y = rxs = zs$ and if $(vs)^k \in E_S$, then

$$z = rx = r(ur)^{k+1}x(sv)^{k+1} = (ru)^{k+1}rxs(vs)^k v = (ru)^{k+1}y(vs)^k v = (ru)^{k+1}y(vs)^k(vs)^k v = (ru)^{k+1}y(vs)^{k+1}(vs)^{k-1}v = y(vs)^{k-1}v\,,$$

which shows that $z\mathcal{R}y$, and so $x\mathcal{D}y$. This proves that $\mathcal{J} = \mathcal{D}$. □

Corollary 4.3. *For all finite semigroups S, $\mathcal{J} = \mathcal{D}$.*

5

Inverse Semigroups

5.1 Regular semigroups

Basic properties of regular elements

An element x of a semigroup S is **regular**, if there exists an element $y \in S$ such that

$$x = xyx\,.$$

Also, S is **regular**, if each of its elements is regular.

Example 5.1. (1) Groups are regular: $x = xx^{-1}x$, where x^{-1} is the group inverse of x. A regular semigroup is a group if and only if it has exactly one idempotent (Exercise).

(2) Idempotents $e \in E_S$ are regular, since $e = eee$. Let $x \in S$ be regular with $x = xyx$. Then $xy \in E_S$ and $yx \in E_S$.

(3) The full transformation semigroup T_X is regular (Exercise). □

Lemma 5.1. *Let $x \in S$. Then*

$$x \text{ is regular} \iff \exists e \in E_S\colon\ x\mathcal{R}e \iff \exists f \in E_S\colon\ x\mathcal{L}f\,.$$

Proof. If x is regular, say $x = xyx$, then $e = xy \in E_S$ and $f = yx$ satisfy $x\mathcal{R}e$ and $x\mathcal{L}f$. On the other hand, if $x\mathcal{R}e$, then $x = ex$ by Lemma 4.6, and there exists a $y \in S^1$ such that $e = xy$. Therefore $x = ex = xyx$, and x is regular. The case for $\mathcal{L}$ is similar. □

Regular D-classes

We say that a class D_x is **regular**, if it contains only regular elements.

Lemma 5.2. *Let $x \in S$ be a regular element. Then the D-class D_x is regular.*

Proof. By Lemma 5.1, the class D_x contains an idempotent e and hence $D_x = D_e$. Let then $z \in D_x$, i.e., $z\mathcal{D}e$. Thus there exists $u \in D_x$ with $e\mathcal{R}u$ and $u\mathcal{L}z$. Hence there are elements $r, s, s' \in S$ such that

$$e = ur,\ u = eu,\ u = sz,\ z = s'u\,.$$

Here, $u = eu$ by Lemma 4.6. Now,

$$z = s'u = s'eu = s'esz = s'ursz = z{\cdot}rs{\cdot}z\,,$$

and hence z is regular. □

In particular, D_e is regular for each idempotent $e \in E_S$.

Lemma 5.3. *If a $\mathcal{D}$-class D is regular, then each $L_x \subseteq D$ and each $R_x \subseteq D$ contains an idempotent.*

Proof. When $x \in D$, then $x = xyx$ for some y, and hence $x\mathcal{L}yx$ and $x\mathcal{R}xy$, where xy and yx are idempotents. □

From these we obtain

Theorem 5.1. *A $\mathcal{D}$-class is regular if and only if it contains an idempotent.*

Inverse elements in semigroups

We say that $y \in S$ is an **inverse element** of $x \in S$, if

$$x = xyx \quad \text{and} \quad y = yxy\,.$$

Note that an inverse element of x, if such an element exists, need not be unique.

Lemma 5.4. *Each regular element $x \in S$ has an inverse element.*

Proof. If $x \in S$ is regular, then for some $y \in S$, $x = xyx$. Now, $yxy = yxy{\cdot}x{\cdot}yxy$, and yxy is also regular. Also, $x = x{\cdot}yxy{\cdot}x$ and consequently yxy is an inverse of x. □

Lallement's lemma

Lemma 5.5. *Let S be a regular semigroup, and let $\alpha\colon S \twoheadrightarrow P$ be an epimorphism onto a semigroup P. If $e \in E_P$, then there exists an idempotent $f \in E_S$ such that $\alpha(f) = e$.*

Proof. Let $x \in S$ be such that $\alpha(x) = e$, and let y be an inverse element of x^2 in S: $x^2 = x^2yx^2$ and $y = yx^2y$. Then for $f = xyx$,

$$\begin{aligned}\alpha(f) &= \alpha(x)\alpha(y)\alpha(x) = \alpha(x)^2\alpha(y)\alpha(x)^2\\ &= \alpha(x^2yx^2) = \alpha(x^2) = \alpha(x)^2 = e^2 = e\,,\end{aligned}$$

that is, $\alpha(f) = e$. Here f is an idempotent: $xyx{\cdot}xyx = x{\cdot}yx^2y{\cdot}x = xyx$. □

As a consequence we have

Theorem 5.2. *Let ρ be a congruence of a regular S, then*

$$x\rho \in E_{S/\rho} \implies \exists e \in E_S\colon\ x\rho = e\rho\,.$$

As an exercise

Theorem 5.3. *If $\alpha\colon S \to P$ is a homomorphism from a regular semigroup S, then $\alpha(S)$ is regular. In particular, if α is an epimorphism, then P is regular.*

5.2 Inverse semigroups

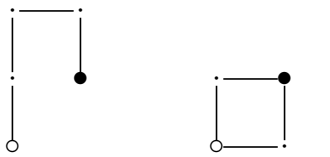
Example

A semigroup S is called an **inverse semigroup**, if each $x \in S$ has a *unique* inverse element x^{-1}:

$$x = xx^{-1}x \qquad \text{and} \qquad x^{-1} = x^{-1}xx^{-1}\,.$$

Example 5.2. (1) If S is a group, then it is an inverse semigroup, and x^{-1} is the group inverse of the element x.

(2) Consider $\mathbb{Z} \times \mathbb{Z}$ as a drawing board so that you can draw a unit line *up*, *down*, *right* and *left*, using the alphabet $A = \{u, d, r, \ell\}$. A word $w \in A^*$ produces a figure in $\mathbb{Z} \times \mathbb{Z}$ starting from $(0, 0)$ and following the instructions given by the word w; see Fig. 5.1.



$w = uurd$ $\qquad$ $v = ru\ell dur$

Fig. 5.1. The figures drawn by the words w and v

We identify the words w_1 and w_2, and denote this by $w_1 \bowtie w_2$, if they draw the same figure with the same endpoint. This means that

$$u \bowtie udu, \quad d \bowtie dud, \quad r \bowtie r\ell r \;\text{ and }\; \ell \bowtie \ell r \ell\,. \tag{5.1}$$

It should be clear that if $w_1 \bowtie w_2$ and $v \in A^*$, then also $vw_1 \bowtie vw_2$ and $w_1 v \bowtie w_2 v$, and hence $\bowtie$ is a congruence on A^+. The quotient $\mathrm{T}^2 = A^+/\bowtie$ might be called the **turtle semigroup**. It is an inverse semigroup with $u^{-1} = d$, $d^{-1} = u$, $r^{-1} = \ell$ and $\ell^{-1} = r$, and the inverse of an element $w \bowtie$ for a word $w = x_1 x_2 \dots x_n \in A^+$ is obtained by reversing the order of the instructions: $w^{-1} = x_n^{-1} x_{n-1}^{-1} \dots x_1^{-1}$. □

The semilattice of idempotents

If $e \in E_S$ for an inverse semigroup, then $eee = e$, and hence $e^{-1} = e$.

Theorem 5.4. *Let S be an inverse semigroup. Then the idempotents E_S form a subsemigroup of S. Moreover, E_S is a semilattice, that is, the idempotents of an inverse semigroup commute.*

Proof. Let $e, f \in E_S$ and consider the (unique) inverse element $x = (ef)^{-1}$ of ef. Now,

$$ef = ef \cdot x \cdot ef = \begin{cases} ef \cdot xe \cdot ef \\ ef \cdot fx \cdot ef \end{cases}$$

and

$$\begin{aligned} xe \cdot ef \cdot xe &= xefx \cdot e = xe\,, \\ fx \cdot ef \cdot fx &= f \cdot xefx = fx\,. \end{aligned}$$

This means that $x = (ef)^{-1} = xe = fx$. Here $x \in E_S$, since

$$x^2 = xe \cdot fx = x \cdot ef \cdot x = x\,,$$

and hence $ef \in E_S$ for all $e, f \in E_S$, that is, E_S is a subsemigroup of S.

Furthermore, E_S is commutative: For $e, f \in E_S$, also $ef, fe \in E_S$, and

$$ef \cdot fe \cdot ef = efef = (ef)^2 = ef \quad \text{and} \quad fe \cdot ef \cdot fe = fefe = (fe)^2 = fe\,,$$

meaning that $fe = (ef)^{-1} = ef$. □

Corollary 5.1. *Assume S is an inverse semigroup. Then*

$$(x_1 x_2 \dots x_n)^{-1} = x_n^{-1} x_{n-1}^{-1} \dots x_1^{-1}$$

for all $x_i \in S$.

Proof. We have

$$xy \cdot y^{-1}x^{-1} \cdot xy = x \cdot yy^{-1} \cdot x^{-1}x \cdot y = x \cdot x^{-1}x \cdot yy^{-1} \cdot y = xy\,.$$

The claim follows by induction. □

Corollary 5.2. *In an inverse semigroup S, we have $x = (x^{-1})^{-1}$ for all $x \in S$.*

A characterization

Theorem 5.5. *Let S be a semigroup. The following are equivalent:*

(1) S is an inverse semigroup.
(2) S is regular and its idempotents commute.
(3) Each $\mathcal{L}$-class and $\mathcal{R}$-class contains a unique idempotent.

Proof. (1) implies (2) by Theorem 5.4.

Suppose (2). By Lemma 5.3, each $\mathcal{L}$-class and $\mathcal{R}$-class contains an idempotent. For the uniqueness, let $f \in L_e$, where $e, f \in E_S$. Hence $e \mathcal{L} f$, and therefore there are $x, y \in S^1$ such that $e = xf$ and $f = ye$. We obtain

$$e = xf = xff = ef = fe = yee = ye = f\,.$$

Similarly, $e\mathcal{R}f$ implies that $e = f$. Hence (2) implies (3).

Suppose (3). Now each $\mathcal{D}$-class contains an idempotent, and, by Theorem 5.1, each $x \in S$ has an inverse element. Suppose an element x has two inverse elements y and z. Now, $yx, zx \in E_S$ with $yx\mathcal{L}x$ and $zx\mathcal{L}x$. Then, by assumption, $yx = zx$. A similar reasoning using $\mathcal{R}$ shows that $xy = xz$. Therefore $y = yxy = zxz = z$, and (1) follows. □

As exercises we have

Corollary 5.3. *Let S be an inverse semigroup. Then*

$$\forall x \in S\colon\ x^{-1}E_S x \subseteq E_S\,.$$

Theorem 5.6. *Let S be an inverse semigroup, and let $x, y \in S$ and $e, f \in E_S$. Then*

(1) $x\mathcal{L}y \iff x^{-1}x = y^{-1}y$.
(2) $x\mathcal{R}y \iff xx^{-1} = yy^{-1}$.
(3) $e\mathcal{D}f \iff \exists z \in S\colon\ e = zz^{-1}$ *and* $f = z^{-1}z$.

Partial ordering inverse semigroups

Recall that in any semigroup S the idempotents can be partially ordered by the relation:

$$e \leq f \iff ef = e = fe\,.$$

This partial order generalizes in an inverse semigroup S to all elements of S as follows,

$$x \leq y \iff \exists e \in E_S\colon\ x = ey\,.$$

Indeed, here $\leq$ is

- reflexive, since $x = (xx^{-1})\cdot x$, where $xx^{-1} \in E_S$;
- antisymmetric, since if $x = ey$ and $y = fx$, then $x = ey = eey = ex$, and $x = ey = efx = fex = fx = y$;
- transitive, since if $x = ey$ and $y = fz$, then also $x = ey = efz$, where $ef \in E_S$.

If you restrict $\leq$ to E_S, you get the above partial order of idempotents. Indeed, if $e \leq f$, then there exists $g \in E_S$ such that $e = gf$, and here $e = gff = ef = fe$ as required.

As an exercise we state

Lemma 5.6. *In an inverse semigroup S we have*

$$\begin{aligned} x \leq y &\iff \exists e \in E_S :\ x = ye \iff xx^{-1} = yx^{-1} \\ &\iff x = xy^{-1}x \iff xx^{-1} = xy^{-1} \\ &\iff x^{-1}x = y^{-1}x \iff x^{-1}x = x^{-1}y \\ &\iff x = xx^{-1}y\,. \end{aligned}$$

Example 5.3. All groups are inverse semigroups. Let S be a semigroup. As an exercise we state that the following conditions are equivalent.

(1) S is a group.
(2) $\forall x \in S\ \exists! x' \in S\colon\ x = xx'x$.
(3) $\forall x \in S\ \exists! x' \in S\colon\ xx' \in E_S$.
(4) S is an inverse semigroup that satisfies: $x = xyx \implies y = yxy$.

□

Example 5.4. If an inverse semigroup S is right cancellative, then it is a group. □

5.3 Representations by injective partial mappings

Partial mappings

Let $X \neq \emptyset$ be a set. A **partial mapping** $\alpha\colon X \to X$ is a function from a subset $Y = \mathrm{dom}(\alpha)$ of X onto $\mathrm{ran}(\alpha) = \alpha(Y) \subseteq X$. A partial mapping $\alpha\colon X \to X$ is **undefined** on all $x \notin \mathrm{dom}(\alpha)$.

Example 5.5. Let $X = \{1, 2, \ldots, 5\}$ and let $\mathrm{dom}(\alpha) = \{2, 4, 5\}$ with $\alpha(2) = 1$, $\alpha(4) = 5$ and $\alpha(5) = 1$. We can represent α as follows

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

Here $\mathrm{ran}(\alpha) = \{1, 5\}$. □

We say that a partial mapping $\alpha\colon X \to X$ is **injective**, if $\alpha(x) \neq \alpha(y)$ for all $x \neq y$ with $x, y \in \mathrm{dom}(\alpha)$. The injective partial mappings form a semigroup, denoted $\mathcal{I}_X$, under the usual composition:

$$(\beta\alpha)(x) = \beta(\alpha(x)) \qquad \text{if} \quad x \in \mathrm{dom}(\alpha) \ \text{ and } \ \alpha(x) \in \mathrm{dom}(\beta)\,.$$

We observe that

$$\mathrm{dom}(\beta\alpha) = \alpha^{-1}(\mathrm{ran}(\alpha) \cap \mathrm{dom}(\beta)) \quad \text{ and } \quad \mathrm{ran}(\beta\alpha) = \beta(\mathrm{ran}(\alpha) \cap \mathrm{dom}(\beta))\,. \tag{5.2}$$

We denote by $\iota_Y\colon X \to X$ the partial function such that $\mathrm{dom}(\iota_Y) = Y = \mathrm{ran}(\iota)$ and $\iota_Y(y) = y$ for all $y \in Y$.

Theorem 5.7. *$\mathcal{I}_X$ is an inverse semigroup.*

Proof. Let $\alpha \in \mathcal{I}_X$, and let $\alpha^{-1}\colon X \to X$ be defined by

$$\alpha^{-1}(y) = x \quad \text{ if } \ \alpha(x) = y \ \text{ for } x \in \mathrm{dom}(\alpha)\,,$$

that is, $\mathrm{dom}(\alpha^{-1}) = \mathrm{ran}(\alpha)$ and $\mathrm{ran}(\alpha^{-1}) = \mathrm{dom}(\alpha)$. Moreover, $\alpha^{-1}(\alpha(x)) = x$ for all $x \in \mathrm{dom}(\alpha)$. Clearly, $\alpha^{-1} \in \mathcal{I}_X$. Furthermore,

$$\alpha^{-1}\alpha = \iota_{\mathrm{dom}(\alpha)} \qquad \text{and} \qquad \alpha\alpha^{-1} = \iota_{\mathrm{ran}(\alpha)}\,.$$

Now, $\alpha\alpha^{-1}\alpha = \alpha$ and $\alpha^{-1}\alpha\alpha^{-1} = \alpha^{-1}$, and so α^{-1} is an inverse element of α, and $\mathcal{I}_X$ is a regular semigroup.

By Theorem 5.4, we need to show that the idempotents of $\mathcal{I}_X$ commute. For this, we prove that

$$\varepsilon \in E_{\mathcal{I}_X} \iff \varepsilon = \iota_Y \quad \text{for some } Y \subseteq X\,.$$

Indeed, suppose ε is an idempotent of $\mathcal{I}_X$. Then by (5.2),

$$\mathrm{dom}(\varepsilon^2) = \varepsilon^{-1}(\mathrm{ran}(\varepsilon) \cap \mathrm{dom}(\varepsilon)) = \mathrm{dom}(\varepsilon)\,,$$

and, by injectivity,

$$\mathrm{ran}(\varepsilon) \cap \mathrm{dom}(\varepsilon) = \varepsilon(\mathrm{dom}(\varepsilon)) = \mathrm{ran}(\varepsilon)\,,$$

which implies that $\mathrm{dom}(\varepsilon) \subseteq \mathrm{ran}(\varepsilon)$. Similarly, $\mathrm{ran}(\varepsilon^2) = \mathrm{ran}(\varepsilon)$ implies $\mathrm{ran}(\varepsilon) \subseteq \mathrm{dom}(\varepsilon)$, and hence $\mathrm{ran}(\varepsilon) = \mathrm{dom}(\varepsilon)$. Now, $\varepsilon^2(x) = \varepsilon(x)$ for all $x \in \mathrm{dom}(\varepsilon)$, and thus, by injectivity, $\varepsilon(x) = x$ for all $x \in \mathrm{dom}(\varepsilon)$.

Finally, the idempotents commute, since for all $Y, Z \subseteq X$,

$$\iota_Z\iota_Y = \iota_{Y\cap Z} = \iota_Y\iota_Z\,,$$

and the theorem is proved. □

The Vagner-Preston representation

Theorem 5.8. *Each inverse semigroup S has a faithful representation as a semigroup of injective partial mappings, i.e., there exists an embedding $\varphi\colon S \hookrightarrow \mathcal{I}_X$ for some set X.*

Proof. We choose $X = S$, and define for each $x \in S$ the mapping

$$\tau_x\colon x^{-1}S \to S \quad \text{by} \quad \tau_x(y) = xy \quad (y \in x^{-1}S)\,.$$

First of all,

$$x^{-1}S = x^{-1}xS, \tag{5.3}$$

because $x^{-1}xS \subseteq x^{-1}S$, and $x^{-1}S = x^{-1}xx^{-1}S \subseteq x^{-1}xS$. In particular, since $x^{-1}x \in E_S$,

$$y \in \mathrm{dom}(\tau_x) = x^{-1}S \iff y = x^{-1}xy\,. \tag{5.4}$$

For $\tau_x \in \mathcal{I}_X$ we need to show that it is injective. Let $y, z \in x^{-1}S$ be such that $\tau_x(y) = \tau_x(z)$. Thus $xy = xz$, and, by (5.4), $y = x^{-1}xy = x^{-1}xz = z$.

We need also

$$\mathrm{dom}(\tau_{yx}) = \mathrm{dom}(\tau_y\tau_x)\,. \tag{5.5}$$

To prove this, notice that

$$z \in \mathrm{dom}(\tau_y\tau_x) \iff z \in \mathrm{dom}(\tau_x) \text{ and } \tau_x(z) \in \mathrm{dom}(\tau_y)\,,$$

that is, by (5.4),

$$z \in \mathrm{dom}(\tau_y\tau_x) \iff z = x^{-1}xz \text{ and } xz = y^{-1}yxz\,.$$

Therefore if $z \in \mathrm{dom}(\tau_y\tau_x)$, then

$$z = x^{-1}xz = x^{-1}y^{-1}yxz = (yx)^{-1}\cdot yx\cdot z$$

and hence $z \in \mathrm{dom}(\tau_{yx})$ by (5.4). On the other hand, if $z \in \mathrm{dom}(\tau_{yx})$, then $z = (yx)^{-1}\cdot yx\cdot z$, and so $z = x^{-1}x\cdot x^{-1}y^{-1}yxz = x^{-1}xz$, from which it follows that $z \in \mathrm{dom}(\tau_x)$. Also, in this case, by commuting the idempotents,

$$xz = xx^{-1}y^{-1}yxz = y^{-1}yxx^{-1}xz = y^{-1}yxz\,,$$

and hence $\tau_x(z) = xz \in \mathrm{dom}(\tau_y)$. This proves (5.5).

Define then $\varphi\colon S \to \mathcal{I}_X$ by

$$\varphi(x) = \tau_x \qquad (x \in S)\,.$$

Now, φ is a homomorphism: When $z \in \mathrm{dom}(\tau_y\tau_x) = \mathrm{dom}(\tau_{yx})$, then

$$\tau_y\tau_x(z) = \tau_y(xz) = yxz = \tau_{yx}(z)\,,$$

and so $\tau_y\tau_x = \tau_{yx}$, that is,

$$\varphi(yx) = \tau_{yx} = \tau_y\tau_x = \varphi(y)\varphi(x)\,.$$

Finally, φ is injective: If $\varphi(x) = \varphi(y)$, then $\tau_x = \tau_y$. By (5.3), $x^{-1}xS = y^{-1}yS$, from which it follows that $x^{-1}x\mathcal{R}y^{-1}y$ for the idempotents $x^{-1}x$ and $y^{-1}y$. By Theorem 5.6, $x^{-1}x = y^{-1}y$. Now, since $x^{-1}x \in x^{-1}S = \mathrm{dom}(\tau_x) = \mathrm{dom}(\tau_y)$, it follows that $\tau_x(x^{-1}x) = \tau_y(x^{-1}x)$. Here $\tau_x(x^{-1}x) = xx^{-1}x = x$ and $\tau_y(x^{-1}x) = yx^{-1}x = yy^{-1}y = y$. Therefore $x = y$ as required. □

5.4 Congruences of inverse semigroups

Heritage of images

Lemma 5.7. *Let S be an inverse semigroup and $\alpha\colon S \to P$ a homomorphism. Then $\alpha(S)$ is an inverse subsemigroup of P.*

Proof. We have for all $x \in S$, $\alpha(x) = \alpha(xx^{-1}x) = \alpha(x)\cdot\alpha(x^{-1})\cdot\alpha(x)$, and so $\alpha(S)$ is a regular subsemigroup of P. Again, it sufficies to show that the idempotents of $\alpha(S)$ commute.

Let $g, h \in E_{\alpha(S)}$. By Lemma 5.5, there exist $e, f \in E_S$ so that $g = \alpha(e)$ and $h = \alpha(f)$. Now, $gf = \alpha(ef) = \alpha(fe) = fg$, from which the claim follows. □

In particular,

Corollary 5.4. *If ρ is a congruence of an inverse semigroup S, then S/ρ is an inverse semigroup.*

Therefore,

Lemma 5.8. *Let S be an inverse semigroup, and ρ its congruence. Then*

$$x\rho y \iff x^{-1}\rho y^{-1}\,.$$

We obtain also that for each homomorphism $\alpha\colon S \to P$ for an inverse semigroup S,

$$\forall x \in S:\ \alpha(x^{-1}) = \alpha(x)^{-1}\,.$$

A subsemigroup T of an inverse semigroup S is called a **inverse subsemigroup**, if for all $x \in T$ also $x^{-1} \in T$, where x^{-1} is the inverse element of x in S. Notice that not all subsemigroups of an inverse semigroup are inverse subsemigroups.

The following lemma is an exercise.

Lemma 5.9. *Let S be an inverse semigroup, and let A be a subsemigroup of S. Then A is an inverse subsemigroup of S if and only if $x^{-1} \in A$ for all $x \in A$.*

Lemma 5.10. *Let S be an inverse semigroup, $\alpha\colon S \to P$ an epimorphism, and let $e \in E_P$. Then $\alpha^{-1}(e)$ is an inverse subsemigroup of S.*

Proof. First of all, if $\alpha(x) = e = \alpha(y)$, then $\alpha(xy) = e^2 = e$, and so $xy \in \alpha^{-1}(e)$. This means that $\alpha^{-1}(e)$ is a sunsemigroup of S.

The claim follows when we show that $x \in \alpha^{-1}(e)$ implies $x^{-1} \in \alpha^{-1}(e)$. For this we just notice that if $\alpha(x) = e$, then $\alpha(x^{-1}) = \alpha(x)^{-1} = e^{-1} = e$. □

For ideals we have a strong closure property (which will be an exercise). In general, if I is an ideal of a semigroup S, then S is called an **ideal extension of** I by S/I, where S/I is the Rees quotient.

Theorem 5.9. *Let I be an ideal of a semigroup S. Then S is an inverse semigroup if and only if I and S/I are inverse semigroups*

Kernels and traces

Let ρ be a congruence of a semigroup S. We define its **kernel** $\ker(\rho)$ and **trace** $\mathrm{tr}(\rho)$ as follows:

$$\ker(\rho) = \{x \in S \mid x\rho e \text{ for some } e \in E_S\} = \bigcup_{e \in E_S} e\rho\,,$$

$$\mathrm{tr}(\rho) = \rho \restriction E_S \times E_S\,.$$

Theorem 5.10. *Let S be an inverse semigroup. Then for all congruences ρ and δ,*

$$\rho \subseteq \delta \iff \forall e \in E_S : e\rho \subseteq e\delta \,.$$

Proof. In one direction the claim is trivial. Suppose then that $e\rho \subseteq e\delta$ for all $e \in E_S$. Now,

$$\begin{aligned} x\rho y &\implies xx^{-1}\rho yx^{-1} \implies xx^{-1}\delta yx^{-1} \implies x\delta yx^{-1}x \,, \\ x\rho y &\implies x^{-1}x\rho x^{-1}y \implies x^{-1}x\delta x^{-1}y \implies yx^{-1}x\delta yx^{-1}y \,, \end{aligned}$$

and thus

$$x\rho y \implies x\delta yx^{-1}y \,. \tag{5.6}$$

On the other hand,

$$x\rho y \implies x^{-1}\rho y^{-1} \implies x^{-1}y\rho y^{-1}y \implies x^{-1}y\delta y^{-1}y \implies yx^{-1}y\delta y \,,$$

and, by combining this with (5.6), we obtain that $x\delta y$ holds. This shows that $\rho \subseteq \delta$. □

Corollary 5.5. *For an inverse semigroup S,*

$$\rho = \delta \iff \forall e \in E_S : e\rho = e\delta$$

for all congruences ρ and δ.

We have then **Vagner's theorem**:

Theorem 5.11. *Let S be an inverse semigroup, and let ρ and δ be its congruences. Then*

$$\rho = \delta \iff \ker(\rho) = \ker(\delta) \text{ and } \operatorname{tr}(\rho) = \operatorname{tr}(\delta) \,.$$

In other words, If $\alpha\colon S \twoheadrightarrow P$ and $\beta\colon S \twoheadrightarrow T$ are epimorphisms from an inverse semigroup S, then $\ker(\alpha) = \ker(\beta)$ if and only if for all $x \in S$ and $e, f \in E_S$,

$$\begin{aligned} \alpha(x) \in E_P &\iff \beta(x) \in E_T \,, \\ \alpha(e) = \alpha(f) &\iff \beta(e) = \beta(f) \,. \end{aligned}$$

Proof. In one direction the claim is again trivial. Suppose then that $\ker(\rho) = \ker(\delta)$ and $\operatorname{tr}(\rho) = \operatorname{tr}(\delta)$. If $e \in E_S$, then

$$\begin{aligned} e\rho x &\implies \exists f \in E_S : f\delta x \implies ff^{-1} = f\delta xx^{-1} \,, \\ e\rho x &\implies ee^{-1} = e\rho xx^{-1} \implies e\delta xx^{-1} \,, \end{aligned}$$

and thus $e\delta f$ and $f\delta x$ hold, from which we obtain $e\delta x$, that is, $e\rho \subseteq e\delta$. The case for $e\delta \subseteq e\rho$ is similar, and hence we can conclude that $e\rho = e\delta$, and finally $\rho = \delta$ by Corollary 5.5. □

Classifications according to traces

Congruences of an inverse semigroup are classified according to their traces.

Lemma 5.11. *For all $x \in S$ and $e \in E_S$, $x^{-1}ex \in E_S$.*

Proof. Indeed, by commuting idempotents, $x^{-1}ex \cdot x^{-1}ex = x^{-1}xx^{-1}eex = x^{-1}ex$. □

For a congruence ρ of an inverse semigroup S, we obtain a congruence $\rho_{\min}$ by defining

$$x\rho_{\min}y \iff \exists e \in E_S : \ xe = ye, \ x^{-1}x\rho e \ \text{ and } y^{-1}y\rho e\,. \tag{5.7}$$

The next theorem states that $\rho_{\min}$ identifies as few elements as possible under the restriction that it should identify exactly the same idempotents as the original ρ. In this way the quotient $S/\rho_{\min}$ is as large as possible.

Theorem 5.12. *For a congruence ρ of an inverse semigroup S, $\rho_{\min}$ is the smallest congruence whose trace equals $\mathrm{tr}(\rho)$.*

Proof. To show that $\rho_{\min}$ is an equivalence relation demands some calculations, which we leave as an exercise.

We show that $\rho_{\min}$ is a congruence. For this suppose $x\rho_{\min}y$ and let $e \in E_S$ be such that (5.7) holds. Let $z \in S$. Now, $f = z^{-1}ez \in E_S$, and

$$xz\cdot f = xzz^{-1}\cdot ez = xe\cdot zz^{-1}z = ye\cdot zz^{-1}z = yzz^{-1}\cdot ez = yz\cdot f\,,$$

and so there exists an idempotent $f \in E_S$ such that $xz\cdot f = yz\cdot f$. Further,

$$(xz)^{-1}\cdot xz = z^{-1}x^{-1}xz \ \rho \ z^{-1}ez = f\,. \tag{5.8}$$

Similarly, $(yz)^{-1}yz \,\rho\, f$, and therefore $xz\rho_{\min}yz$. On the other hand, $xe = ye$ implies that $ex^{-1} = ey^{-1}$ by taking inverses of both sides, and using this we obtain

$$\begin{aligned}
&(zx)^{-1}\cdot zx = x^{-1}\cdot z^{-1}z\cdot x = x^{-1}x\cdot x^{-1}z^{-1}z\cdot x \ \rho \ ex^{-1}z^{-1}zx\,,\\
&(zy)^{-1}\cdot zy = y^{-1}\cdot z^{-1}z\cdot y = y^{-1}y\cdot y^{-1}z^{-1}zy\cdot y^{-1}y \ \rho \ ey^{-1}z^{-1}zye\\
&= ex^{-1}\cdot z^{-1}z\cdot xe \ \rho \ ex^{-1}z^{-1}zx\cdot x^{-1}x \ = ex^{-1}z^{-1}zx\,,
\end{aligned}$$

which shows that $zx\rho_{\min}zy$, since $ex^{-1}z^{-1}zx$ is an idempotent. Thus $\rho_{\min}$ is a congruence.

Next we show that $\mathrm{tr}(\rho_{\min}) = \mathrm{tr}(\rho)$. For $f, g \in E_S$,

$$f\rho_{\min}g \implies \exists e \in E : \ fe = ge, \ f\rho e \ \text{ and } \ g\rho e \implies f\rho g\,.$$

On the other hand, suppose $f\rho g$. Now, for $e = fg$ we have $fe = ffg = fg = fgg = gfg = ge$, and, clearly, $f\rho e$ and $g\rho e$. Hence also $f\rho_{\min}g$ and the traces are the same.

Finally, $\rho_{\min}$ is the smallest of such congruences: For any congruence δ with $\mathrm{tr}(\delta) = \mathrm{tr}(\rho)$,

$$x\rho_{\min}y \implies \exists e \in E_S : xe = ye,\ x^{-1}x\rho e,\ y^{-1}y\rho e$$
$$\implies x^{-1}x\delta e \text{ and } y^{-1}y\delta e \implies x\delta xe \text{ and } y\delta ye \implies x\delta y\,,$$

and therefore $\rho_{\min} \subseteq \delta$ as required. □

In particular, we have that $\rho_{\min} \subseteq \rho$ for all congruences ρ of an inverse semigroup S.

For a congruence ρ of an inverse semigroup S define $\rho_{\max}$ by

$$x\rho_{\max}y \iff \forall e \in E_S : x^{-1}ex\rho y^{-1}ey\,.$$

Theorem 5.13. *Let S be an inverse semigroup and ρ its congruence. Then $\rho_{\max}$ is the largest congruence of S whose trace equals $\mathrm{tr}(\rho)$.*

Proof. First of all we show that $\rho_{\max}$ is a congruence. That it is an equivalence relation is an easy exercise. Let then $x\rho_{\max}y$ and let $z \in S$. We have

$$zx\rho_{\max}zy \iff \forall e \in E_S : (zx)^{-1}ezx\rho(zy)^{-1}ezy \iff x^{-1}z^{-1}ezx\rho y^{-1}z^{-1}ezy\,.$$

Since $z^{-1}ez \in E_S$, we have

$$x\rho_{\max}y \implies x^{-1}\cdot z^{-1}ez\cdot x\rho y^{-1}\cdot z^{-1}ez\cdot y\,,$$

and thus $zx\rho_{\max}zy$ holds if $x\rho_{\max}y$ holds. Similarly, for $xz\rho_{\max}yz$. We conclude that $\rho_{\max}$ is a congruence.

Next we show that $\mathrm{tr}(\rho_{\max}) = \mathrm{tr}(\rho)$. For this let $f, g \in E_S$. Then

$$f\rho_{\max}g \implies f^{-1}ff\rho g^{-1}fg \implies f\rho fg\,,$$

and in the same way, $f\rho_{\max}g \implies g\rho fg$, from which we get that $f\rho g$, and so $\mathrm{tr}(\rho_{\max}) \subseteq \mathrm{tr}(\rho)$. In the other direction, for all $e \in E_S$, if $f\rho g$, then $f^{-1}ef\rho g^{-1}eg$ and so we deduce that $f\rho_{\max}g$, that is, the traces are the same.

Finally $\rho_{\max}$ is the largest of such congruences. For this assume δ is a congruence such that $\mathrm{tr}(\delta) = \mathrm{tr}(\rho)$. Now, for all $e \in E_S$,

$$x\delta y \implies x^{-1}ex\delta y^{-1}ey \implies x^{-1}ex\rho y^{-1}ey \implies x\rho_{\max}y\,,$$

since $x^{-1}ex, y^{-1}ey \in E_S$; and so $\delta \subseteq \rho_{\max}$ as required. □

The above theorem states that $\rho_{\max}$ identifies as many elements of S as possible with the restriction that it does not identify any idempotents unless ρ does so. Certainly, $\rho \subseteq \rho_{\max}$, and so the quotient $S/\rho_{\max}$ is an epimorphic image of S/ρ.

Group congruences

We say that a congruence ρ of a semigroup S is a **group congruence**, if S/ρ is a group.

The following lemma holds already for regular semigroups.

Lemma 5.12. *An inverse semigroup is a group if and only if it has a unique idempotent.*

For a congruence ρ of an inverse semigroup, we have by Theorem 5.2 that

$$x\rho \in E_{S/\rho} \implies \exists e \in E_S : \ e\rho = x\rho \,,$$

and hence

Theorem 5.14. *A congruence ρ of an inverse semigroup S is a group congruence if and only if* $\mathrm{tr}(\rho) = E_S \times E_S$.

Proof. If $\mathrm{tr}(\rho) = E_S \times E_S$, then S/ρ has exactly one idempotent by Theorem 5.2. In the other direction the claim is equally clear. □

If ρ is a group congruence of an inverse semigroup S, then so is $\rho_{\min}$, because now $\mathrm{tr}(\rho) = \mathrm{tr}(\rho_{\min}) = E_S \times E_S$. In particular, $\rho_{\min}$ is the smallest group congruence of S, and for all group congruences δ of S, $\delta_{\min} = \rho_{\min}$.

The **smallest group congruence** of an inverse semigroup S is denoted by σ_S.

Let then ρ be a group congruence. Then $\sigma_S = \rho_{\min} \subseteq \rho$, and hence, by the homomorphism theorem, there exists a unique homomorphism β such that

$$\rho^{\mathrm{nat}} = \beta\sigma_S^{\mathrm{nat}} \,.$$

This means that every group G, which is a homomorphic image of S, is a homomorphic image of the group S/σ_S, and in this sense S/σ_S is a maximal homomorphic image of S.

Remark 5.1. If S is *not* an inverse semigroup, it need not have the smallest group congruence. As an example consider $(\mathbb{N}_+, +)$. The group congruences of this semigroup are exactly $\rho_n = \{(p, q) \mid p \equiv p(\mod n)\}$. □

Theorem 5.15 (Munn). *In an inverse semigroup S,*

$$x\sigma_S y \iff \exists e \in E_S : \ xe = ye \,.$$

Proof. Now, $\sigma_S = \sigma_{\min}$, and so

$$x\sigma_S y \iff \exists e \in E_S : \ xe = ye, \ x^{-1}x\sigma_S e, \ y^{-1}y\sigma_S e \,,$$

where $x^{-1}x, y^{-1}y \in E_S$, and so the condition reduces to the claim. □

As an exercise we have

Theorem 5.16. *In an inverse semigroup S,*

(1) $x\sigma_S y \iff \exists e \in E_S : \ ex = ey.$
(2) $\ker(\sigma_S) = \{x \in S \mid \exists y \in S : \ xy = x\}.$

Idempotent separating congruences

A group congruence puts all idempotents in the same congruence class. An idempotent separating congruence does the opposite, it puts different idempotents to different classes.

A congruence ρ of a semigroup S is **idempotent separating**, if

$$\forall e, f \in E_S : \ e\rho f \implies e = f \, .$$

From this definition we have immediately

Lemma 5.13. *If ρ is an idempotent separating congruence of an inverse semigroup S, then $\mathrm{tr}(\rho) = \iota_E = \{(e, e) \mid e \in E_S\}$.*

By Theorem 5.13, for each inverse semigroup S there exists the **greatest idempotent separating congruence**, which will be denoted by μ_S.

Theorem 5.17. *For all inverse semigroups S,*

$$x\mu_S y \iff \forall e \in E_S : \ x^{-1}ex = y^{-1}ey \, .$$

Proof. This is clear, since from Theorem 5.13, since $\mathrm{tr}(\mu_S) = \iota_E$, and so the claim follows. □

As an exercise, using Theorem 5.6, we have

Theorem 5.18. *For an inverse semigroup S,*

(1) $\mu_S \subseteq \mathcal{H}$;
(2) if $\rho \subseteq \mathcal{L}$, then ρ is idempotent separating.

Books on semigroups:

A.H. Clifford and G.B. Preston. *The algebraic theory of semigroups, Vol. I & II*, Mathematical Surveys of the Amer. Math. Soc. **7** (1961 & 1967).

P.A. Grillet, *Semigroups*, Marcel Dekker, 1995.

P.M. Higgins. *Techniques of semigroup theory*. Oxford University Press 1992

J.M. Howie. *An introduction to semigroup theory*. Academic Press, 1976.

J.M. Howie. *Fundamentals of semigroup theory*. Clarendon Press 1995.

G. Lallement. *Semigroups and combinatorial applications*. Wiley, 1979.

M.L. Lawson. *Inverse Semigroups*. World Scientific 1998.

E.S. Lyapin. *Semigroups*. Transl. of Math. Monographs, Vol 3. Amer. Math. Soc. 1974.

M. Petrich. *Introduction to semigroups*. Merrill 1973.

M. Petrich. *Lectures in semigroups*. Wiley 1977.

M. Petrich. *Inverse semigroups*. Wiley 1984.

Related books:

J. Berstel and D. Perrin. *Theory of Codes*. Academic Press 1985.

S. Burris and H.P. Sankappanavar. *A course in universal algebra*. Springer 1981. Also freely available on the web.

P.M. Cohn. *Universal algebra*. Reidel 1981.

M. Lothaire. *Combinatorics on Words*, Addison-Wesley 1983.

M. Lothaire. *Algebraic Combinatorics on Words*, Cambridge Univ. Press 2002.

R.C. Lyndon and P. Schupp. *Combinatorial group theory*. Springer 1977.

W. Magnus, A. Karrass and D. Solitar. *Combinatorial group theory*. Dover 1976