Varieties of Formal Languages

by

J. E. Pin

Varieties of Formal Languages

by

J. E. Pin

Translated by

A. Howie

THE LANGUAGE OF SCIENCE
Plenum
PUBLISHING CORPORATION

English edition first published 1986
by North Oxford Academic Publishers Ltd,
a subsidiary of Kogan Page Ltd, 120 Pentonville Road,
London N1 9JN

Published in the United States of America by Plenum Press,
a Division of Plenum Publishing Corporation.
233 Spring Street
New York, NY 10013

Library of Congress Catalog Card Number

86-060593

ISBN 0-306-42294-8

Printed and bound in Great Britain.

Preface

The theory of finite automata and of rational languages could be likened to the ground floor of a huge building under construction which is theoretical computer science. The metaphor would indicate first that it can be entered on the ground level and secondly that it is more convenient to go through it in order to reach the higher levels. It is also the first purely mathematical theory to emerge from the needs and intuitions of computer science in the wider sense. In fact, at the end of the 1950s Kleene, who was intrigued by electronic models of the nervous system which were then very fashionable, proposed characterizing feasible calculations by means of a system making use of a single bounded memory.

This led him to discover what are now called rational languages which are the subject of the present book by J. E. Pin. Subsequent work has revealed that this class is a particularly fundamental mathematical entity in the study of finite systems, for they appear quite naturally starting from considerations as diverse as those of restricted logical systems or the standard rational functions of analysis.

From the start, one of the principal problems was found to be a problem of classification, or rather of hierarchization. J. Rhodes showed that the composition of automata preserved the associated groups and McNaughton discovered that the existence of non-trivial groups of this kind was intimately related to the presence of loops within the system of calculation. The development by S. Eilenberg of the notion of *variety of language* gave a new impetus to research by coordinating these results with others such as the excellent theorem of Imre Simon. J. E. Pin has been one of the most active investigators in this area and we are indebted to him for numerous original contributions to the subject of varieties.

However, this rapid growth has necessitated a new synthesis incorporating the techniques discovered since the treatise of S. Eilenberg. J. E. Pin has undertaken this task and has been successful in presenting the subject with the care of an inspiring teacher, beginning with the most elementary aspects. Although it can easily be included in the more general framework of the theory of automata, it is an independent work, both of introduction and of preparation for research, which the author presents to the public interested in mathematics and in data processing.

M. P. SCHÜTZENBERGER
Professor in the University of Paris VII
Corresponding Member of the Academy of Sciences

Foreword

The aim of this book is to present the fundamental results of the theory of finite automata and of recognizable languages, or regular languages.

The concept of a finite automaton occurs frequently in computer science processing, whether it is used to model a particular machine, to formalize records of communication or to describe logic circuits. The notion of a recognizable language is equally fundamental: it enables us to take account of the linking of the calculations in a program, to express the behaviour of a process or to describe certain operations of a text editor and more generally to describe any iterative algorithm. Moreover, recognizable languages constitute the first link in a hierarchy of progressively more complex languages. Under this heading they were studied quite early: Kleene's theorem, on which the entire theory of finite automata is based, dates from 1956. Subsequently, the work of numerous investigators, in the first rank of whom appear M. P. Schützenberger, R. McNaughton, J. A. Brzozowski and I. Simon, has made clear the profound connections which exist between finite automata, recognizable languages and finite semigroups. The concept of variety of languages, which was introduced by Eilenberg in 1976, has enabled us to formalize this triple approach—automata, language and semigroups—and has provided a coherent and unified framework for the theory.

Since then the modern theory of automata has been constructed around this fundamental idea.

Thus the aim of this book is to present the theory of automata from the point of view of variety of languages. This approach has at least two advantages: it enables us to handle classical results in a concise and rigorous manner, and it facilitates access to the most recent results and problems. The first four chapters of the book are devoted to fundamental statements of the theory of automata. These theorems are proved at the same time (with the exception of Kleene's theorem, which is stated and admitted without proof) and illustrated by numerous examples. The final chapter presents a succinct review of the most recent results of the theory but contains no proofs.

The content of this book is approximately that of a course on DEA given at the University of Paris VI in the period 1981–1983. The level of presentation is thus that of an advanced graduate course. However, certain parts of the book—for example Chapter 1, part of Chapter 2 or the results in Chapter 4—can be used for a master's course on the theory of languages. The whole book should enable the reader to arrive quickly at the research level; thus the most important references

are accompanied by a brief commentary aimed at facilitating the orientation of the reader. The reader of this book does not require any previous knowledge of formal languages or automata. However, it is necessary to have some familiarity with the formalism of algebra although, again, the previous theoretical knowledge required is limited to the notion of a group.

Chapter 0 defines the notation used in the book. Chapter 1 presents the basic material: semigroups, finite automata and recognizable languages. All the details of the algorithms for calculating the syntactic monoid of a language are also given in this chapter. Moreover, these algorithms have been implemented on a computer (APL programs of d'Autebert, Cousineau, Perrot and Rifflet). Chapter 2 is devoted to varieties. First the varieties of semigroups and of finite monoids are introduced, then their interpretation in terms of equations is given and finally Eilenberg's variety theorem is proved. Elementary examples of varieties of languages are presented at the end of the chapter. Chapter 3 is an introduction to the theory of finite semigroups. The subjects dealt with are Green's relations, simple and 0-simple semigroups, and the structure of regular $\mathscr{D}$-classes and of the minimal ideal of a finite semigroup. The algorithm for calculation of a regular $\mathscr{D}$-class is presented in detail and is illustrated by numerous examples. This algorithm has also been implemented on a computer (the APL programs of the authors cited above). We return to varieties of semigroups in the last two sections: the first presents the varieties of semigroups defined by Green's relations, and the second introduces relational morphisms and V-morphisms. The theorem of I. Simon on piecewise-testable languages and that of Schützenberger on star-free languages are proved in Chapter 4. Some applications of Simon's theorem to finite semigroups and the characterization of $\mathscr{R}$-trivial and $\mathscr{L}$-trivial languages are also included in this chapter. Chapter 5 presents various aspects of the theory of automata. The first section is devoted to operations on the languages: concatenation, mixing, star, morphisms, sequential functions etc. The second gives a résumé of recent work on hierarchies of languages—including the connection with symbolic logic—and the third presents the relations with the theory of variable-length codes. Finally the last section recalls briefly some other lines of research. The problems appearing at the end of the chapters are often the subjects of research. However, we have not given an indication of the difficulty of these problems except for problems which are open or have been recently solved—to determine the difficulty of a problem is itself a difficult problem.

I wish to thank in particular my friends S. W. Margolis, H. Straubing and D. Thérien for their numerous remarks and suggestions during the preparation of this book. I am also grateful to all the people who have read through or commented on various parts of the manuscript: J. Berstel, J. P. Pécuchet, D. Perrin, Ch. Reutenauer, G. Rindone, S. Schwer and W. Thomas as well as all the students on my course in Paris. I should like also to thank G. Lallement, J. F. Perrot and M. P. Schützenberger to whom I owe my interest in the theory of automata. Finally I thank Madame A. Dupont for her excellent work on the typing of the manuscript.

J. E. PIN

Contents

Introduction

Relations

This brief preliminary chapter is aimed at making precise certain definitions and properties referring to binary relations.

Let E and F be two sets. A **relation** between E and F is a subset R of $E \times F$. If $E = F$, we say that R is a relation on E. If $(u, v) \in E \times F$ we often write $u R v$ in place of $(u, v) \in R$. For example it is more convenient to write $2 \leq 3$ than $(2, 3) \in \leq$.

A relation R on a set E is **reflexive** if, for every $u \in E$, uRu. It is **symmetric** if, for every $(u, v) \in E \times E$, uRv implies vRu. It is **transitive** if, for every $(u, v, w) \in E \times E \times E$, uRv and vRw imply uRw. It is **antisymmetric** if, for every $(u, v) \in E \times E$, uRv and vRu imply $u = v$. An **equivalence relation** is a relation which is simultaneously reflexive, symmetric and transitive. A **quasi-order relation** is a relation which is reflexive and transitive. A **partial order relation** is a relation which is reflexive, transitive and antisymmetric. A **total order relation** is an order relation such that, for every $(u, v) \in E \times E$, we have $u R v$ or $v R u$.

A (partial) function $\varphi : E \to F$ is a relation over $E \times F$ such that for every $x \in E$ there exists one and only one (in the case of a partial function, at most one) element $y \in F$ such that $(x, y) \in \varphi$. When this y exists, we denote it by $\varphi(x)$ or $x\varphi$. In this book we shall employ the notation $x\varphi$, which is more convenient in the theory of automata.

We can also consider each relation $R \subset E \times F$ in a dynamic way and associate with it the function τ from E into the set of subsets of F defined by

$$u\tau = \{v \in F \mid (u, v) \in R\}$$

Conversely, if τ is a function from E into the set of subsets of F, the **graph** R of τ, which is defined by

$$R = \{(u, v) \in E \times F \mid v \in u\tau\}$$

is a relation between E and F. By abuse of language, we say that $\tau : E \to F$ is a relation from E *into* F.

A relation $\tau : E \to F$ is called **injective** if, for every $u, v \in E$, $u\tau \cap v\tau \neq \varnothing$ implies $u = v$. In particular, if τ is a function, we find again the standard notion of an injective function. Moreover, τ is a **surjective** relation if, for every $v \in F$, there exists $u \in E$ such that $v \in u\tau$.

The relations over a set E are ordered by inclusion: if $R, S \subset E \times E$ are two

relations, we say that R is **finer** than S (or that S is **coarser** than R) if $R \subset S$. In Chapter 4 we shall have to use the following elementary result.

Proposition 0.1

For every partial order relation R on a finite set E, there exists a total order relation on E which is coarser than R.

Proof

Put $S(R) = \{(a,b) \in E \times E | (a,b) \notin R \text{ and } (b,a) \notin R\}$. If $S(R) = \varnothing$, R is a total order. Otherwise fix $(a,b) \in S(R)$ and put $R' = R \cup \{(x,y) \in E \times E | (x,a) \in R \text{ and } (b,y) \in R\}$. Then R' is a partial ordering on E which is coarser than R. Since $(a,b) \in R'$, $S(R')$ is strictly included in $S(R)$ and we reach the conclusion by induction over the cardinality of $S(R)$.

Given a relation $\tau : E \to F$ which is a graph $R \subset E \times F$, we denote by $\tau^{-1} : F \to E$ the graph relation $R^{-1} = \{(v,u) \in F \times E | (u,v) \in R\}$. We can then see easily that, for every $v \in F$, $v\tau^{-1} = \{u \in E | v \in u\tau\}$.

More generally, if X is a subset of E, we put

$$X\tau = \bigcup_{x \in X} x\tau$$

If Y is a subset of F, we then have

$$Y\tau^{-1} = \bigcup_{y \in Y} y\tau^{-1} = \{u \in E | \text{there exists } y \in Y \text{ such that } y \in u\tau\}$$

i.e.

$$Y\tau^{-1} = \{u \in E | u\tau \cap Y \neq \varnothing\}$$

Given two relations $\tau_1 : E \to F$ and $\tau_2 : F \to G$, we denote by $\tau_1\tau_2$ the relation $E \to G$ defined, for every $u \in E$, by $u(\tau_1\tau_2) = \{w \in G | \text{there exists } v \in F \text{ such that } v \in u\tau_1 \text{ and } w \in v\tau_2\}$. In the case in which τ_1 and τ_2 are partial functions we find again the standard notion of the composition of two partial functions — up to the order of the factors — since we are using a 'post-fixed' notation. With a prefixed notation, we could write $\tau_2 \circ \tau_1$ in place of $\tau_1\tau_2$.

We now define some elementary properties which will frequently be used without reference in subsequent chapters. The proofs are immediate and are left to the reader.

Proposition 0.2

Let $\varphi : E \to F$ be a partial function. Then

(1) the relation φ^{-1} is injective;

(2) if φ is an injective function, φ^{-1} is an injective partial function and $\varphi\varphi^{-1}$ is the identity on E;

(3) if $\varphi : E \to F$ is a surjective partial function, then $\varphi^{-1}\varphi$ is the identity on F.

Proposition 0.3
Let $\tau: E \to F$ be a relation. Then for every $X, Y \subset E$, we have $(X \cup Y)\tau = X\tau \cup Y\tau$.

In the case in which τ is injective, we can be more precise.

Proposition 0.4
Let $\tau: E \to F$ be an injective relation (in particular $\tau = \varphi^{-1}$ where $\varphi: F \to E$ is a partial function). Then for every $X, Y \subset E$ we have

(1) $$(X \cup Y)\tau = X\tau \cup Y\tau$$
(2) $$(X \cap Y)\tau = X\tau \cap Y\tau$$
(3) $$(X \setminus Y)\tau = X\tau \setminus Y\tau$$

Proposition 0.5
Let $\varphi: E \to F$ be a surjective partial function. Then for every $X \subset E$ and $Y \subset F$, we have $X\varphi \cap Y = X\varphi \cap Y\varphi^{-1}\varphi = (X \cap Y\varphi^{-1})\varphi$.

Proposition 0.6
Let E, F, G be three sets and $\alpha: G \to E$, $\beta: G \to F$ be two functions. Suppose that α is surjective and that, for every $s, t \in G$, $s\alpha = t\alpha$ implies $s\beta = t\beta$. Then the relation $\alpha^{-1}\beta: E \to F$ is a function.

Apart from the 'post-fixed' notation used for functions and relations, we have followed the terminology and notation of M. Lothaire for everything that concerns free monoids and we have retained most of the notation of Eilenberg elsewhere. This notation is consistent with the notation regularly used in mathematics with one exception. The notation $\mathbb{Z}_p$ designates not the p-adic numbers but the group of integers modulo p which is regularly denoted by $\mathbb{Z}/p\mathbb{Z}$. Finally, following an abuse of notation which is quite widely accepted, we shall sometimes identify the singleton $\{s\}$ with the element s.

Chapter 1

Semigroups, Languages and Automata

The aim of this chapter is to give most of the definitions relating to the semigroups and languages which will be used in subsequent chapters. Some general statements on semigroups, almost all elementary, will also be found; the only difficult statement of this chapter is Theorem 1.10, which is a consequence of Ramsey's theorem. The second part of the chapter gives a brief résumé of the relations between automata, semigroups and languages. Finally, the last section is devoted to the explicit calculation of two syntactic semigroups.

1. Semigroups

1.1. Semigroups, monoids, morphisms

A **semigroup** is a couple formed from a set S and an internal associative law of composition defined on S. This law is generally denoted in a multiplicative way. Given two semigroups S and T, a semigroup morphism $\varphi: S \to T$ is a function from S into T such that, for all $x, y \in S$, $(xy)\varphi = (x\varphi)(y\varphi)$.

A **monoid** is a triplet formed from a set M, an internal associative law of composition defined over M and a distinct element of M, denoted by 1, such that, for every $x \in M$, $1x = x1 = x$. In practice, we usually denote the monoid (or semigroup) and the underlying set by the same letter. Given two monoids M and N, a monoid morphism $\varphi: M \to N$ is a function from M into N such that $1\varphi = 1$ and such that, for every $x, y \in M$, $(xy)\varphi = (x\varphi)(y\varphi)$. In the remainder of this book the word 'morphism' denotes, according to context, a semigroup morphism or a monoid morphism.

Given a semigroup S, we denote by S^1 the following monoid: if S is a monoid, $S^1 = S$; if S is not a monoid, $S^1 = S \cup \{1\}$ together with the law $*$ defined by $x * y = xy$ if $x, y \in S$ and $1 * x = x * 1 = x$ for every $x \in S^1$.

The semigroups (or monoids), together with the morphism which we have just defined, form a category. We shall see later that there exists another interesting category whose objects are semigroups and whose morphisms will be called 'relational morphisms'.

In agreement with the general definition, we say that a morphism $\varphi: S \to T$ is an isomorphism if there exists a morphism $\psi: T \to S$ such that $\varphi\psi = Id_S$ and $\psi\varphi = Id_T$. In fact a morphism is an isomorphism if and only if it is bijective. As a general rule we shall identify two isomorphic semigroups. This rule applies in particular to the definition of subsemigroups: we say that S is a **subsemigroup** of T if there exists an injective morphism $\varphi: S \to T$. S is then identified with $S\varphi$ together with the law induced by that of T. We shall say that T is a **quotient** of S if there exists a surjective morphism $\varphi: S \to T$.

A **submonoid** of a monoid M is a subsemigroup of M containing 1. If this submonoid is a group, we say that it is a **subgroup** of M. In particular the set U of invertible elements of M is the maximal subgroup of M, which is also called the group of **units** of M. The notion of a subgroup of a monoid must not be confused with that of a group within a **semigroup** S; a group *within* S is a subsemigroup of S which is a group.

We say that a semigroup S divides a semigroup T (notation $S < T$) if S is a quotient of a subsemigroup of T.

Proposition 1.1
The division relation is transitive.

Proof
Suppose $S_1 < S_2 < S_3$. Then there exists a subsemigroup T_1 of S_2, a subsemigroup T_2 of S_3 and surjective morphisms $\pi_1: T_1 \to S_1$ and $\pi_2: T_2 \to S_2$. Put $T = T_1\pi_2^{-1}$. Then T is a subsemigroup of S_3 and S_1 is a quotient of T since $T\pi_2\pi_1 = T_1\pi_1 = S_1$. Then S_1 divides S_3.

Given a family $(S_i)_{i\in I}$ of semigroups, the product

$$\prod_{i\in I} S_i$$

is the semigroup defined on the set

$$\prod_{i\in I} S_i$$

by the law $(s_i)_{i\in I} \cdot (s'_i)_{i\in I} = (s_i s'_i)_{i\in I}$. Since the semigroup 1 consisting of a single element is the identity with respect to the product operation, following the usual practice we put

$$\prod_{i\in\emptyset} S_i = 1$$

We note that the product of a family of monoids is a monoid.

A semigroup S is **generated** by a subset P of S if every element of S can be written in the form $p_1 \dots p_n$ with $n > 0$ and $p_1 \dots p_n \in P$.

1.2. Idempotents, zero, ideal

An element e of a semigroup S is **idempotent** if $e = e^2$. We shall denote by $E(S)$ the set of idempotents of S. As we shall see, the idempotents play a fundamental role in the study of finite semigroups.

We call a **zero** of S an element, denoted by 0, such that $0s = s0 = 0$ for every $s \in S$. If S is a semigroup, we denote by S^0 the semigroup obtained from S by the addition of a zero: the support of S^0 is the disjoint union of S and the singleton $\{0\}$ and the law (here denoted $*$) is defined by $s * s' = ss'$ if $s, s' \in S$ and $s * 0 = 0 * s = 0$ for every $s \in S^0$.

A subset I of S is an **ideal**, a **right ideal** or a **left ideal** if $S^1IS^1 \subset I$, $IS^1 \subset I$ or $S^1I \subset I$ respectively.

A non-empty ideal I of a semigroup S is called **minimal** if, for every non-empty ideal J of S, $J \subset I$ implies $J = I$. We note that if such an ideal exists it is necessarily unique. The existence of a minimal ideal is assured in at least two important cases, namely if S is finite or if S possesses a zero. In this last case $\{0\}$ is the minimal ideal. A non-empty ideal $I \neq \{0\}$ such that, for every non-empty ideal J of S, $J \subset I$ implies $J = \{0\}$ or $J = I$ is called a **0-minimal ideal**. It should be noted that a semigroup can have several 0-minimal ideals.

1.3. Congruences

A **congruence** on a semigroup S is an equivalence relation $\sim$ on S compatible on the left and on the right with multiplication, i.e. such that, for every $a, b, c \in S$, $a \sim b$ implies $ac \sim bc$ and $ca \sim cb$. Classically the quotient set $S/\sim$ is then naturally provided with a semigroup structure. Three particular cases of congruences will be extensively used in the remainder of this book.

Rees congruence

Let I be an ideal of S and let $\equiv_I$ be the equivalence relation identifying all the elements of I and separating the other elements. Formally $s \equiv_I s'$ if and only if $s = s'$ or $s, s' \in I$. $\equiv_I$ is then a congruence called the Rees congruence. Traditionally we write S/I for the quotient of S by $\equiv_I$.

Syntactic congruence

Let P be a subset of S and $\equiv$ be an equivalence relation over S. We say that $\equiv$ saturates P if P is the union of classes modulo $\equiv$, which amounts to saying that, for every $u, v \in S$, $u \equiv v$ and $u \in P$ imply $v \in P$. The **syntactic congruence** of P is the congruence $\sim_P$ over S defined by $u \sim_P v$ if and only if, for every $s, t \in S^1$, $(sut \in P \Leftrightarrow svt \in P)$. We can show (this is left as an exercise) that $\sim_P$ is the coarsest congruence saturating P. This congruence is particularly important in the theory of languages as we shall see a little later.

Nuclear congruence

Let $\varphi: S \to T$ be a semigroup morphism. We denote by $\sim_\varphi$ the (nuclear) congruence associated with φ and defined by

$$u \sim_\varphi v \text{ if and only if } u\varphi = v\varphi$$

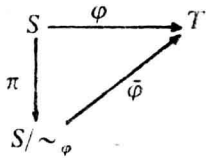
We then have the classical result.

Proposition 1.2

Let $\varphi: S \to T$ be a semigroup morphism and $\pi: S \to S/\sim_\varphi$ be the natural projection. Then there exists a unique morphism $\bar{\varphi}: S/\sim_\varphi \to T$ such that $\varphi = \pi\bar{\varphi}$. Moreover, $\bar{\varphi}$ is an isomorphism of $S/\sim_\varphi$ over $S\varphi$.

Proof

The situation is summed up in the following diagram:



Necessarily $\bar{\varphi} = \pi^{-1}\varphi$. It is now necessary to verify that $\bar{\varphi}$ is indeed a morphism. Now if $u, v \in s\pi^{-1}\varphi$, there exists $x, y \in s\pi^{-1}$ such that $x\varphi = u$ and $y\varphi = v$. Since $x, y \in s\pi^{-1}$, $x \sim_\varphi y$, i.e. $x\varphi = y\varphi$. Thus $\bar{\varphi}$ is a function. Moreover, if $x_1 \in s_1\pi^{-1}$ and $x_2 \in s_2\pi^{-1}$, it follows that $x_1x_2 \in (s_1s_2)\pi^{-1}$ whence

$$(s_1\bar{\varphi})(s_2\bar{\varphi}) = (x_1\varphi)(x_2\varphi) = (x_1x_2)\varphi = (s_1s_2)\bar{\varphi}$$

$\bar{\varphi}$ is injective; if $s_1\bar{\varphi} = s_2\bar{\varphi}$, there exists $x_1 \in s_1\pi^{-1}$ and $x_2 \in s_2\pi^{-1}$ such that $x_1\varphi = x_2\varphi$. Hence we can deduce $x_1 \sim_\varphi x_2$, i.e. $x_1\pi = x_2\pi$, whence $s_1 = s_2$. Then $\bar{\varphi}$ induces an isomorphism of $S/\sim_\varphi$ over its image $(S/\sim_\varphi)\bar{\varphi} = S\varphi$.

Let $(\sim_i)_{i\in I}$ be a family of congruences over a semigroup S. We denote by $\sim$ the intersection of the family $(\sim_i)_{i\in I}$; by definition $u \sim v$ if and only if, for every $i \in I$, $u \sim_i v$.

Proposition 1.3

With the preceding notation $S/\sim$ is a subsemigroup of

$$\prod_{i\in I} S/\sim_i$$

Proof

We denote by $\pi_i: S \to S/\sim_i$ the projections and by

$$\pi: S \to \prod_{i\in I} S/\sim_i$$

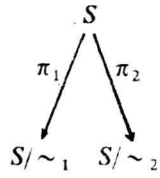
the morphism defined by $s\pi = (s\pi_i)_{i\in I}$ for every $s \in S$. The nuclear congruence of π is none other than $\sim$, and thus $S/\sim$ is isomorphic to $S\pi$ in accordance with Proposition 1.2.

Proposition 1.4

Let $\sim_1$ and $\sim_2$ be two congruences defined over a semigroup S. We suppose that, for every $s, t \in S$, $s \sim_1 t$ implies $s \sim_2 t$. Then $S/\sim_2$ is a quotient of $S/\sim_1$.

Proof

We denote by $\pi_1 : S \to S/\sim_1$ and $\pi_2 : S \to S/\sim_2$ the canonical morphisms. The condition of the statement implies that $\pi = \pi_1^{-1}\pi_2$ is a surjective morphism $S/\sim_1 \to S/\sim_2$.



1.4. Semigroups of transformations

If E is a set, we denote by $\mathscr{T}(E)$ the monoid of functions from E into E together with the composition of functions. If $E = \{1, \ldots, n\}$, we generally write $\mathscr{T}_n$ for the monoid $\mathscr{T}(E)$.

A transformation semigroup over E is a subsemigroup of $\mathscr{T}(E)$. The importance of transformation semigroups is emphasized by the following proposition.

Proposition 1.5

Every semigroup is isomorphic to a transformation semigroup. In particular every finite semigroup S is isomorphic to a subsemigroup of $\mathscr{T}_n$ for a certain integer n.

Proof

We associate with each element s of S the right translation $\rho_s : S^1 \to S^1$ defined by $a\rho_s = as$ for every $a \in S^1$. We can easily verify that the function $s \to \rho_s$ thus defined is an injective morphism from S into $\mathscr{T}(S^1)$.

There follows another elementary theorem; it is concerned with the structure of semigroups generated by a single element (sometimes called **monogenic** semigroups).

Proposition 1.6

Let S be a semigroup generated by an element a. Then either $S = (\mathbb{N}\setminus\{0\}, +)$ or S is finite. In the latter case there exist integers $n \geq 0$ and $p > 0$ such that $a^n = a^{n+p}$ and $S = \{a, a^2, \ldots, a^{n+p-1}\}$. Then S contains a single idempotent, which is the identity of the group $G = \{a^n, a^{n+1}, \ldots, a^{n+p-1}\}$.

Proof

If all the powers of a are distinct we clearly find ourselves in the first case. Otherwise let n be the smallest positive integer such that there exists k satisfying $a^n = a^{n+k}$ and let us write p for the smallest k satisfying this last relation. Then all