Analysis on Semigroups

Function Spaces, Compactifications, Representations

JOHN F. BERGLUND

Virginia Commonwealth University,
Richmond

HUGO D. JUNGHENN

George Washington University,
Washington, D.C.

PAUL MILNES

The University of Western Ontario,
London, Canada

WILEY

A Wiley-Interscience Publication

JOHN WILEY & SONS

New York · Chichester · Brisbane · Toronto · Singapore

Library of Congress Cataloging-in-Publication Data

Berglund, John F.

Analysis on semigroups: function spaces, compactifications, representations/by John F. Berglund, Hugo D. Junghenn, Paul Milnes.

p. cm.—(Canadian Mathematical Society series of monographs and advanced texts = Monographies et études de la Société mathématique du Canada)

"A Wiley-Interscience publication."

Bibliography: p.

Includes indexes.

ISBN 0-471-61208-1

1. Topological semigroups. 2. Representations of semigroups. 3. Bohr compactification. I. Junghenn, Hugo D. II. Milnes, Paul. III. Title. IV. Series: Canadian Mathematical Society series of monographs and advanced texts.

QA387.B47 1988
512′.55—dc19 88-14292
CIP

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To the memory of
Vicki Ostrolenk,
loving wife, devoted friend

Preface

This book presents a unified treatment of certain topics in analysis on semigroups, in particular, those topics that pertain to the functional analytic and dynamical theory of continuous representations of semitopological semigroups.

It is well known that the study of such representations is facilitated by the use of semigroup compactifications. The importance of compactifications in this respect derives from the fact that the dynamical and structural properties of a given representation frequently appear as algebraic and/or topological properties of an associated semigroup compactification. Thus the introduction of a suitable compactification makes available powerful results from the theory of compact semigroups. The interplay between the dynamics of semigroup representations and the algebraic and topological properties of semigroup compactifications is the main theme of this book.

A representation of some importance to us is the so-called right regular representation, which is the representation of a semigroup S by right translation operators on the C^*-algebra $\mathcal{B}(S)$ of bounded, complex-valued functions on S. The study of this representation reduces essentially to the study of translation invariant subspaces of $\mathcal{B}(S)$. A significant portion of the monograph is devoted to the investigation of the structure of these function spaces and the dynamical properties of their members.

The subject of analysis on semigroups can trace its origins back to the work of H. Bohr (1925, 1926) on almost periodic functions on the real line. Bohr's definition of almost periodic function (which is found in Chapter 4 under Example 1.2(c)) is a natural generalization of that of periodic function, and his original methods involve reduction to the periodic case. In 1927, S. Bochner gave a functional analytic characterization of almost periodicity, and this led J. von Neumann (1934a) and Bochner and von Neumann (1935) to develop a theory of almost periodic functions on an arbitrary group. Subsequently, A. Weil (1935, 1940) and E.R. van Kampen (1936) used group compactifications to show that the theory of almost periodic functions on a discrete group may be reduced to the theory of continuous functions on a compact topological group.

As a specific illustration of the utility of semigroup compactifications, consider the almost periodic compactification of the additive group $\mathbb{R}$ of real numbers. By definition, this compactification consists of a pair (ψ, G), where G is a compact, Hausdorff, topological group and ψ is a continuous homomorphism from $\mathbb{R}$ into G such that $\{ f \circ \psi : f \in \mathcal{C}(G) \}$ is the space of almost periodic functions on $\mathbb{R}$. Here, $\mathcal{C}(G)$ denotes the subalgebra of $\mathcal{B}(G)$ consisting of the continuous functions. Now, a classical result in the theory of almost periodic functions on $\mathbb{R}$ asserts that each such function may be uniformly approximated by trigonometric polynomials. Bohr's proof of this result relies on his theory of periodic functions of infinitely many variables. The defining property of the compactification (ψ, G), however, allows one to infer this result immediately from the Peter–Weyl theorem.

Generalizations of the classical theory of almost periodicity have taken several directions. We mention a few of those that are of particular interest to us. First, Bochner's definition of almost periodic function on a group does not make use of the existence of inverses in a group and hence applies equally well to semigroups. Furthermore, the definition of continuous almost periodic function on a topological group does not involve the joint continuity property of multiplication. Thus the natural domain of a continuous almost periodic function (and the setting of the modern theory of almost periodicity) is a semitopological semigroup, that is, a semigroup with a topology relative to which multiplication is separately continuous. Apart from their suitability in this context, such semigroups have become important in applications, as they arise naturally in the study of semigroups of operators on Banach spaces.

A second direction the theory of almost periodicity has taken is the broader study of functions of "almost periodic type." An early example of such a function is the weakly almost periodic function, which was first defined and investigated by W. F. Eberlein (1949). Although these functions have many of the characteristics of almost periodic functions (e.g., the space of weakly almost periodic functions on a group admits an invariant mean), there are essential differences between the two kinds of functions. These differences show up clearly in the structure of the associated compactifications: the almost periodic compactification of a semigroup is always a topological semigroup (i.e., multiplication is jointly continuous), whereas the weakly almost periodic compactification is, in general, only a semitopological semigroup.

Another direction the theory has taken is the investigation of almost periodic properties of representations of a semitopological semigroup S by operators on an arbitrary Banach space. Here, the notion of almost periodic function is replaced by the more general concept of almost periodic vector. This generalization of almost periodicity was initiated by K. Jacobs (1956) and was further developed by K. de Leeuw and I. Glicksberg (1961a). The basic idea is this. If $s \to U_s$ is a continuous representation of S by operators on a Banach space X, one defines the space X_a of almost periodic vectors in X as the set of all vectors x such that $U_S x$ is norm relatively compact. Then X_a is a closed subspace of X, which reduces to the space of almost periodic functions if $X = \mathcal{C}(S)$ and $U_s = R_s$. By replacing the norm topology in the previous definition by the weak topology one obtains the space X_w of weakly almost periodic vectors in X. Such vectors occur in great profusion in representation theory. For example, if U_S is uniformly bounded and X is a reflexive Banach space, then every vector in X is weakly almost periodic. The point here is that dynamical properties of the representation U on the space X_w may be deduced from the algebraic structure of the weakly almost periodic compactification of S.

An effective tool in the study of semigroup representations is the invariant mean. For example, it is the existence of such a mean on the space of weakly almost periodic functions on a group that guarantees that a representation U on X_w possesses certain desirable dynamical properties. Here again, compactifications are useful, since a mean may be represented as a probability measure on an associated compactification and hence may be studied by measure theoretic methods.

The book falls roughly into four parts. The first part, Chapter 1, is the study of semigroups with topology. In the first two sections of the chapter we develop the requisite algebraic theory of semigroups. Section 1 presents the elementary aspects of the theory, whereas Section 2 gives a detailed description of the structure of the minimal ideal of a semigroup with minimal idempotents. In Section 3 we introduce the notions of right topological, semitopological, and topological semigroup and prove the fundamental structure theorems for compact right topological semigroups. Section 4 takes up the problem of generating points of joint continuity for separately continuous actions. The results of this section are used in Section 5 to refine the structure theorems of Section 3. In Section 6 we give a brief introduction to the general theory of flows and conclude the section with I. Namioka's flow theoretic proof of Ryll–Nardzewski's celebrated fixed point theorem.

The second part of the book, Chapter 2, develops the general theory of means on function spaces. The basic properties of means are assembled in Section 1. In Section 2 we introduce the notion of introversion, which is the essential ingredient in the theory of semigroup compactifications as developed in this monograph. Sections 3, 4, and 5 discuss some of the more important features of the theory of invariant means on semigroups. (Sections 4 and 5 are somewhat more specialized than the other sections and may be omitted on first reading.)

In Chapter 3, which comprises the third part of the book, we construct the machinery of semigroup compactifications. The general theory of semigroup compactifications is presented in Section 1. In Section 2 we introduce the basic device used in the construction of universal compactifications, namely subdirect products. The fundamental theorem on the existence of universal compactifications is given in Section 3, along with many examples. In Section 4 we develop the theory of affine compactifications.

The final part of the book consists of Chapters 4, 5, and 6. In these chapters, we use compactifications to determine the dynamical behavior of semigroup representations. The right translation representation, which is the subject of Chapters 4 and 5, is studied in terms of the structural properties of various spaces of functions of almost periodic type. In all, 10 distinct types of functions are investigated, starting with the space of (Bochner) almost periodic functions and ending with the related class of Bohr almost periodic functions. In Chapter 6 we take up the study of arbitrary weakly almost periodic representations of semigroups. The general theory is developed in the first two sections, and applications to ergodic theory and Markov operators are given in Sections 3 and 4.

Appendices on weak compactness, joint continuity, and invariant measures are included at the end of the book.

The book contains more than 200 exercises. These range from simple applications and examples to significant complements to the theory. Many of the more difficult exercises are supplied with hints.

The prerequisite for reading the book is a working knowledge of the basic principles of functional analysis, general topology, and measure theory as found, say, in the core curriculum of a traditional U.S. or Canadian Master's Degree program.

The book is organized as follows. Each of the six chapters is divided into sections. In each section, theorems, corollaries, propositions, definitions, remarks, examples, and exercises are numbered m.n, where m is the section number and n the number of the item within the section. Cross references to items inside the current chapter are written (m.n), whereas references to items outside the chapter are written (k.m.n), where k is the number of the chapter containing the item. Section m of Chapter k is referred to as Section k.m outside Chapter k and simply as Section m inside the chapter. The three appendices are labelled A, B, and C. The *n*th item of Appendix A, say, is marked and cross-referenced as A.n.

Bibliographical references are given in the Notes section at the end of each chapter. Although we do not claim completeness for the bibliography, the listing is sufficiently detailed to allow further investigation of the topics presented in this monograph. Failure to cite a reference for a particular result should not be taken as a claim of originality on our part.

Finally, we would like to acknowledge our indebtedness, spiritual and otherwise, to the many mathematicians who have influenced us before and during the preparation of this monograph. We mention in particular J.W. Baker, M.M. Day, K. de Leeuw, I. Glicksberg, K.H. Hofmann, T. Mitchell, I. Namioka, and J.S. Pym.

JOHN F. BERGLUND
HUGO D. JUNGHENN
PAUL MILNES

Virginia Commonwealth University, Richmond
George Washington University, Washington, D.C.
The University of Western Ontario, London

Summary of Notation

We indicate here the notational conventions and basic terminology that will be used throughout the book.

The symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ denote the sets of natural numbers, integers, rational numbers, real numbers, and complex numbers, respectively. We also define $\mathbb{R}^+ = [0, \infty)$, $\mathbb{Z}^+ = \mathbb{Z} \cap \mathbb{R}^+$, $\mathbb{Q}^+ = \mathbb{Q} \cap \mathbb{R}^+$, $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, and $\mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$.

Unless otherwise stipulated, we shall take the scalar field of a vector space to be the field of complex numbers. If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces or, more generally, locally convex topological vector spaces, then $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ denotes the vector space of all continuous linear mappings from $\mathcal{X}$ into $\mathcal{Y}$. $\mathcal{L}(\mathcal{X}, \mathcal{X})$ is denoted by $\mathcal{L}(\mathcal{X})$, and $\mathcal{L}(\mathcal{X}, \mathbb{C})$, the *dual space* of $\mathcal{X}$, is denoted by $\mathcal{X}^*$.

If $A \subset \mathcal{X}$ and $B \subset \mathcal{X}^*$, then $\sigma(A, B)$ denotes the weakest topology on A relative to which the restriction to A of each member of B is continuous. A net $\{x_\alpha\}$ in A $\sigma(A, B)$-converges to $x \in A$ if and only if $x^*(x_\alpha) \to x^*(x)$ for all $x^* \in B$. With the topology $\sigma(\mathcal{X}, B)$, $\mathcal{X}$ is a locally convex topological vector space. A typical basic convex neighborhood of zero in this topology is the set $\{x \in \mathcal{X} : |x_i^*(x)| < \epsilon, i = 1, 2, \ldots, n\}$, where $x_1^*, x_2^*, \ldots, x_n^* \in B$ and $\epsilon > 0$. $\sigma(\mathcal{X}, \mathcal{X}^*)$ is called the *weak topology* of $\mathcal{X}$. Dually, $\sigma(B, A)$ is the weakest topology on B relative to which the mapping $x^* \to x^*(x) : B \mapsto \mathbb{C}$ is continuous for each $x \in A$. $\sigma(\mathcal{X}^*, \mathcal{X})$ is called the *weak* topology* of $\mathcal{X}^*$.

If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, then $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a Banach space under the *uniform operator norm*

$$\|U\| = \sup\left\{\|Ux\| : \|x\| \leq 1\right\} \qquad (U \in \mathcal{L}(\mathcal{X}, \mathcal{Y})).$$

There are two additional locally convex topologies on $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ that are of interest to us: the *strong operator topology*, which is the weakest topology of $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ relative to which the mapping $U \to Ux : \mathcal{L}(\mathcal{X}, \mathcal{Y}) \to \mathcal{Y}$ is continuous for each $x \in \mathcal{X}$, and the *weak operator topology*, which is the weakest topology of $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ relative to which the mapping $U \to y^*(Ux) : \mathcal{L}(\mathcal{X}, \mathcal{Y}) \to \mathbb{C}$ is continuous for each $x \in \mathcal{X}$ and $y^* \in \mathcal{Y}^*$.

We denote the closure of a set A in a topological space by A^- or $\overline{A}$. If $\overline{A}$ is a subset of a locally convex topological vector space, then sp A and co A denote, respectively, the linear span of A and the convex hull of A. The closures of these sets are denoted by $\overline{\text{sp}}\, A$ and $\overline{\text{co}}\, A$, respectively. The convex circled (or convex balanced) hull of A is denoted by cco A. If A is convex, then ex A denotes the set of extreme points of A.

For a nonempty set S, $\mathcal{B}(S)$ denotes the set of all bounded complex-valued functions on S. $\mathcal{B}(S)$ is a C^*-algebra with respect to the usual operations of (pointwise) addition, multiplication, scalar multiplication, and complex conjugation, and with respect to the *uniform* (or *supremum*) *norm* given by $\| f \| = \sup \{ | f(s) | : s \in S \}$. Unless otherwise stipulated, any mention of norm on a subspace of $\mathcal{B}(S)$ will refer to the uniform norm.

For $f \in \mathcal{B}(S)$, the functions $\mathcal{R}e\, f$, $\mathcal{I}m\, f$, $\bar{f}$, and $| f |$ are defined as

$$(\mathcal{R}e\, f)(s) = \mathcal{R}e\,(f(s)),\ (\mathcal{I}m\, f)(s) = \mathcal{I}m\,(f(s)),$$

$$\bar{f}(s) = \overline{f(s)},\ | f |(s) = | f(s) | \qquad (s \in S).$$

Also, if f and g are real-valued members of $\mathcal{B}(S)$, then $f \vee g$ and $f \wedge g$ are defined by

$$(f \vee g)(s) = f(s) \vee g(s) \quad \text{and} \quad (f \wedge g)(s) = f(s) \wedge g(s) \qquad (s \in S)$$

where, for real numbers x and y, $x \vee y$ is the maximum of x and y, and $x \wedge y$ is the minimum. If c is a complex number, we shall use the same symbol to denote the function whose constant value is c. If $A \subset S$, then 1_A denotes the *indicator function* of A, that is, the function on S whose value is 1 on A and 0 on $S \setminus A$. For subsets T of S and F of $\mathcal{B}(S)$, the set of functions $\{ f|_T : f \in F \}$ is denoted by $F|_T$. If $S = S_1 \times S_2$ and $f_i \in \mathcal{B}(S_i)$, $i = 1, 2$, we shall write $f_1 \otimes f_2$ for the function on S whose value at (s_1, s_2) is $f(s_1) f(s_2)$.

The space of bounded, continuous, complex-valued functions on a topological space S is denoted by $\mathcal{C}(S)$. Clearly, $\mathcal{C}(S)$ is a C^*-subalgebra of $\mathcal{B}(S)$, that is, $\mathcal{C}(S)$ is closed under addition, multiplication, scalar multiplication, complex conjugation, and uniform limits. If S is locally compact, then $\mathcal{C}_0(S)$ denotes the C^*-subalgebra of $\mathcal{C}(S)$ consisting of the functions that vanish at infinity. If S is a convex subset of a locally convex topological vector space, we denote by $\mathcal{AF}(S)$ the space of bounded, continuous, complex-valued, affine functions on S. Note that $\mathcal{AF}(S)$ is a norm closed, conjugate closed, linear subspace of $\mathcal{C}(S)$.

The *dual* of a continuous mapping $\theta : S \to T$ from a topological space S into a topological space T is the mapping $\theta^* : \mathcal{C}(T) \to \mathcal{C}(S)$ defined by $\theta^*(f) := f \circ \theta, f \in \mathcal{C}(T)$. Clearly, $\theta^* \in \mathcal{L}(\mathcal{C}(T), \mathcal{C}(S))$.

Contents

Chapter One

Semigroups

The theory of semigroups arose in an effort to generalize ring theory and group theory and in particular the theory of groups of transformations of a set. The subject has grown enormously over the last 50 years and draws on many areas of mathematics, including algebraic topology, manifolds, and functional analysis. In this chapter we give a brief introduction to the subject, focusing on those aspects that will be needed in later sections. The reader who wishes to pursue the subject in more detail should consult the references mentioned in the Notes section at the end of the chapter.

The first two sections of the chapter deal only with the algebraic theory of semigroups. The remaining sections treat semigroups with a topology that is to some degree compatible with the algebraic structure.

1 ALGEBRAIC THEORY: BASIC CONCEPTS

1.1 Definition.

A *semigroup* is a pair $(S, \cdot)$, where S is a nonempty set and $(\cdot)$ is an associative (binary) operation $(s, t) \to s \cdot t : S \times S \to S$. Associativity means that

$$r \cdot (s \cdot t) = (r \cdot s) \cdot t \qquad (r, s, t \in S).$$

A semigroup with only one element is called *trivial*.

The operation on S will usually be called *multiplication*, and $s \cdot t$ will be called the *product* of s and t. Other notations for $s \cdot t$ are $s + t$ and $s \circ t$, the choice (usually) depending on the context. We shall generally drop the symbol for multiplication and denote the product of s and t by st. If $s \in S$ and n is a natural number, we shall write s^n for $ss \ldots s$ (n factors).

It may be shown [as in Petrich (1973), for example] that every semigroup satisfies the *general associative law*, which asserts that the value of the product of n members of the semigroup is independent of the positioning of the parentheses.

1.2 Notation.

For each member t of a semigroup S, define $\rho_t : S \to S$ and $\lambda_t : S \to S$ by

$$\rho_t(s) = st, \lambda_t(s) = ts \qquad (s \in S).$$

For subsets A, B of S define

$$At = \rho_t(A), \qquad tA = \lambda_t(A),$$
$$At^{-1} = \rho_t^{-1}(A), \qquad t^{-1}A = \lambda_t^{-1}(A),$$

and

$$AB = \bigcup_{t \in B} At = \bigcup_{t \in A} tB = \{st : s \in A, t \in B\}.$$

If $A_1, A_2, \ldots, A_n$ are subsets of S, define $A_1A_2 \cdots A_n$ inductively by $A_1A_2 \cdots A_n = (A_1A_2 \cdots A_{n-1})A_n$. If each $A_i = A$, we write A^n for $A_1A_2 \cdots A_n$. Finally, if S is a group, define

$$A^{-1} = \{s^{-1} : s \in A\}.$$

1.3 Definition.

Elements s, t in a semigroup S are said to *commute* if $st = ts$. The *center* of S is the set $Z(S)$ of all members of S that commute with every member of S. S is said to be *commutative* or *abelian* if $Z(S) = S$, that is, if any two elements of S commute.

The standard examples of commutative semigroups are $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^+$, $\mathbb{Q}^+$, and $\mathbb{Z}^+$ under ordinary addition or ordinary multiplication. Moreover, $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{C}, +)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{C} \setminus \{0\}, \cdot)$, and $(\mathbb{T}, \cdot)$ are commutative groups, and $(\mathbb{D}, \cdot)$ is a commutative semigroup.

An important example of a noncommutative semigroup is the set $M(n, \mathbb{C})$ of all $n \times n$ matrices over $\mathbb{C}$ under matrix multiplication ($n \geq 2$). For another example, let X be a set with cardinality greater than 1. Then the set X^X of all functions from X into X is a noncommutative semigroup with composition of functions as the semigroup operation.

1.4 Definition.

An element e of a semigroup S is called a *right* (respectively, *left*) *identity* for S if $se = s$ (respectively, $es = s$) for all $s \in S$. A right identity that is also a left identity is called an *identity*. Identities will frequently be denoted by the symbol 1. If S is a semigroup with identity 1, we define $s^0 = 1$ for any $s \in S$.

A semigroup may have many right identities. For example, in the semigroup consisting of all matrices of the form

$$\begin{bmatrix} 0 & x \\ 0 & 1 \end{bmatrix} \qquad (x \in \mathbb{R}),$$

every element is a right identity. However, if a semigroup has a right identity and a left identity, then the two coincide and the common element is an identity (Exercise 1.30). In particular, a semigroup can have at most one identity.

If a semigroup S lacks an identity, one may adjoin a new symbol 1 to S and define $1s = s1 = s$ for all $s \in S \cup \{1\}$; if the original product is retained for pairs from S, then $S \cup \{1\}$ is a semigroup with identity 1.

1.5 Notation.

If a semigroup S does not have an identity, then S^1 will denote S with an identity adjoined in the manner described in Definition 1.4. If S already has an identity, then we set $S^1 = S$.

1.6 Definition.

An element z in a semigroup S is a *right zero* if $sz = z$ for all $s \in S$. If every member of S is a right zero, then S is called a *right zero semigroup*. *Left zero* and *left zero semigroup* are defined analogously. A right zero that is also a left zero is called a *zero*. Zeros are frequently denoted by the symbol 0. If S has a zero and $st = 0$ for all $s, t \in S$, then S is called a *null semigroup*.

If S has a left zero and a right zero, then the two are equal and the common element is a zero (Exercise 1.30). Thus, a semigroup has at most one zero.

Note that *any* nonempty set may be given a multiplication relative to which it is a right zero semigroup. A similar comment applies for the left zero and null cases.

1.7 Example.

Under matrix multiplication, the set consisting of the matrices

$$a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$c = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad d = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is a semigroup with a left zero that is not a right zero and a left identity that is not a right identity. This is easily seen from the following "multiplication table" for S:

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

1.8 Definition.

An element e of a semigroup S is said to be an *idempotent* if $e^2 = e$. The set of all idempotents of S is denoted by $E(S)$. If $E(S) = S$, then S is called an *idempotent semigroup* or a *band*. A commutative idempotent semigroup is called a *semilattice*.

Left zero and right zero semigroups are simple examples of bands. If S_1 and S_2 are arbitrary nonempty sets, then the Cartesian product $S_1 \times S_2$ with multiplication

$$(s_1, s_2)(t_1, t_2) = (s_1, t_2)$$

is an idempotent semigroup that, except in trivial cases, is neither a left zero nor a right zero semigroup. Note that if S_1 is given left zero multiplication and S_2 is given right zero multiplication, then the product of two members of $S_1 \times S_2$ may be viewed as the result of multiplying coordinatewise (see Definition 1.26).

A simple example of a semilattice is the set $\{0, 1\}$, where 0 is a zero and 1 an identity. Another example is any totally ordered set with multiplication $xy = \min\{x, y\}$.

1.9 Definition.

Let S be a semigroup and let T be a nonempty subset of S. T is said to be

(a) a *subsemigroup* of S if $T^2 \subset T$, that is, if T is a semigroup with respect to multiplication in S;

(b) a *subgroup* of S if T is a group relative to multiplication in S;

(c) a *left ideal* of S if $ST \subset T$;

(d) a *right ideal* of S if $TS \subset T$; and

(e) a (*two-sided*) *ideal* of S if T is both a left ideal and a right ideal.

If, in any of these definitions, $T \neq S$, then T is said to be *proper*.

In the matrix semigroup of Example 1.7, $\{a, b, c\}$ is a left ideal that is not a right ideal, $\{b\}$ is a right ideal that is not a left ideal, $\{a, b\}$ is a subsemigroup that is neither a left ideal nor a right ideal, and $\{b, c\}$ is a proper ideal.

It is easy to verify that the intersection of a family of subsemigroups of a semigroup S is again a subsemigroup of S, provided the intersection is nonempty. The corresponding statements for left ideals, right ideals, and ideals also hold. An important special case is given in the next definition.

1.10 Definition.

Let A be a nonempty subset of a semigroup S. The intersection of all subsemigroups (respectively, left ideals, right ideals, ideals) of S that contain A is called the *subsemigroup* (respectively, *left ideal*, *right ideal*, *ideal*) *generated by* A, and the elements of A are called *generators*. The subsemigroup generated by A will be denoted by $\langle A \rangle$. If $S = \langle A \rangle$, we say that S is *generated by* A. A semigroup generated by a single element is said to be *cyclic*.

The subsemigroup generated by A may be concretely realized as the set of all products $s_1 s_2 \cdots s_n$, where $n \in \mathbb{N}$ and $s_i \in A$, $1 \leq i \leq n$. It is clearly the smallest subsemigroup of S containing the set A. Similarly, the left ideal (respectively, ideal) generated by A may be written $A \cup SA = S^1 A$ (respectively, $A \cup SA \cup AS \cup SAS = S^1 A S^1$).

If e is an idempotent in a semigroup S, then there is at least one subgroup of S containing e, namely $\{e\}$. The next result asserts that there is a largest such subgroup. First we give the following definition.

1.11 Definition.

Let e be an idempotent in a semigroup S. The union of all subgroups of S containing e is called the *maximal subgroup of* S *containing* e and is denoted by $H(e)$. If S has an identity 1, then $H(1)$ is called the *group of units* of S.

The following proposition justifies the use of the terminology of Definition 1.11.

1.12 Proposition.

Let S *be a semigroup and let* $e \in E(S)$. *Then* $H(e)$ *is a subgroup of* S *with identity* e.

Proof. Let T denote the subsemigroup of S generated by $H(e)$. Since $se = es = s$ for all $s \in H(e)$, e is an identity for T. Let $s \in T$. Then $s = s_1 s_2 \cdots s_n$, where $n \in \mathbb{N}$ and $s_i \in H(e)$, $i = 1, 2, \ldots, n$. For each i choose $t_i \in H(e)$ such that $s_i t_i = t_i s_i = e$, and set $t := t_n \cdots t_2 t_1$. Then $st = ts = e$, which shows that T is a group. Therefore, $H(e) = T$. □

It is easily verified that in general $H(e) = \{t \in eSe : e \in St \cap tS\}$. Our main interest is in the case when $H(e) = eSe$. Necessary and sufficient conditions for this to occur are given in the next section (Theorem 2.8).

1.13 Definition.

A semigroup S is called *left* (respectively, *right*) *simple* if it has no proper left (respectively, right) ideals. S is *simple* if it has no proper two-sided ideals.

A left zero semigroup with more than one element is left simple but not right simple. Groups are left simple, right simple, and simple. (For a converse, see Theorem 1.17.) Obviously, a semigroup that is left or right simple is simple. The following is an example of a semigroup that is simple but not a group.

1.14 Example.

Let S be the set of all matrices

$$\begin{bmatrix} x & 0 \\ y & 1 \end{bmatrix} \qquad (x, y \in (0, \infty)).$$

Under matrix multiplication, S is simple but neither left nor right simple. For instance, the subset of S defined by the condition $y > 1$ is a proper left ideal, and the subset defined by the condition $y > 2x$ is a proper right ideal.

The next proposition provides some easy tests for determining when a given semigroup is left simple, right simple, or simple.

1.15 Proposition.

A semigroup S is left (respectively, right) simple if and only if $St = S$ (respectively, $tS = S$) for all $t \in S$. S is simple if and only if $StS = S$ for all $t \in S$.

Proof. We prove only the left simple version. Since St is a left ideal, the necessity is clear. The sufficiency follows from the observation that if L is a left ideal and $t \in L$, then $St \subset L$. □

Remark. It is clear that the right simple version of 1.15 may be proved by making obvious modifications to the proof just given. This is an example of a situation that occurs frequently in semigroup theory: a "left" statement has a dual "right" statement, and the proof of one statement is the mirror image of that of the other. Hereafter, we shall record only one of the left/right statements and refer to the other as its "dual." (Beginning in Section 3 we shall encounter instances where this left/right duality fails, but these occur in a topological context and are easily distinguishable from the situations currently under discussion.)

1.16 Definition.

A semigroup S is *right* (respectively, *left*) *cancellative* if $r, s, t \in S$ and $sr = tr$ (respectively, $rs = rt$) imply $s = t$. A semigroup that is both left and right cancellative is said to be *cancellative*.