ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 12, Number 2, Spring 1982

GENERALIZED INVERSE SEMIGROUPS WITH INVOLUTION

H. E. SCHEIBLICH

ABSTRACT. This paper considers orthodox semigroups which have a normal band and which admit a unary involution operation. A structure theorem is proved and the free such semigroup is found. The partial order is also considered.

1. Introduction and preliminaries. Let S be a regular semigroup. Then S is *orthodox* provided the set $E = E(S)$ of idempotents of S is a subsemigroup of S. S is a *generalized inverse* semigroup provided S is orthodox and the band E is *normal*, i.e., $eghf = ehgf$ for all $e, g, h, f \in E$. The structure of all generalized inverse semigroups has been found by Yamada **[10]**, in terms of inverse semigroups and normal bands. Yamada's structure theorem will play an important role in this paper. It is described below.

In **[5]**, the authors consider a class of unary semigroups, i.e., semigroups which are equipped with a unary operator, $x \to x^*$. The * operator satisfies the axioms (1) $x^{**} = x$, (2) $xx^*x = x$, and an involution axiom (3) $(xy)^* = y^*x^*$. These semigroups are called regular * semigroups. In that same paper, a regular * semigroup S is shown to be orthodox ($EE \subset E$) if and only if S satisfies the identity $[(xx^*)(yy^*)(zz^*)]^2 = [(xx^*)(yy^*)(zz^*)]$. Thus, the class of all orthodox * semigroups forms a variety.

In **[1]**, it is shown that an orthodox * semigroup S is a generalized inverse semigroup (has a normal band) if and only if S satisfies the identity $a(xx^*)(x^*x)b = a(x^*x)(xx^*)b$. Thus, the generalized inverse * semigroups form a variety.

The purpose of this note is to study generalized inverse * semigroups. In Sections (2), (3), and (4) we shall: (2) specialize Yamada's structure theorem to the * case, (3) find the free generalized inverse * semigroup, and (4) consider the natural partial order on a generalized inverse semigroup.

2. The structure theorem. First, we will review Yamada's structure theorem for generalized inverse semigroups. Let S be an inverse semigroup

---

AMS (*MOS*) *subject classification* (1970): Primary 20M05, 20M10.

Key words and pharases: generalized inverse semigroup, involution operator, free, partial order.

Received by the editors on November 17, 1980, and in revised form on March 17, 1981.

with semilattice E. Let L, R be left and right normal bands, respectively, each with structure semilattice E. Thus, $L = \bigcup_{e \in E} L_e$ and $R = \bigcup_{e \in E} R_e$. Now let $Q = [L\colon S\colon R] = \{(a, x, b) \in L \times S \times R\colon a \in L_{xx^{-1}}$ and $b \in R_{x^{-1}x}\}$. Define multiplication in Q by the rule $(a, x, b)(c, y, d) = (au, xy, vd)$ where $u \in L_{(xy)(xy)^{-1}}$ and $v \in R_{(xy)^{-1}(xy)}$. Q is called the quasi direct product of L, S, and R.

THEOREM 2.1. [**10**]. *The quasi direct product* $[L\colon S\colon R]$ *is a generalized inverse semigroup. Conversely, if* T *is a generalized inverse semigroup, then the mapping* $a \to (R_{aa'}, a\mathscr{Y}, L_{a'a})$ *is an isomorphism of* T *onto the quasi direct product* $[E(T)/\mathscr{R}\colon T/\mathscr{Y}\colon E(T)/\mathscr{L}]$.

In the second part of the theorem just stated, $E(T)$ is the band of T, and $\mathscr{R}$, $\mathscr{L}$ denote the usual Green's relations on $E(T)$. Thus, $R_{aa'}$ is the $\mathscr{R}$ class of $E(T)$ which contains aa' where a' is an inverse of a. $\mathscr{Y}$ is the smallest inverse semigroup congruence on T.

Let S be an inverse semigroup with semilattice E. Let L be a left normal band with E as its structure semilattice. Let L^d be the band dual to L. Of course, L^d is right normal and E is the structure semilattice of L^d. Define the unary operation * on the quasi direct product $Q = [L\colon S\colon L^d]$ by $(a, x, b)^* = (b, x^{-1}, a)$.

THEOREM 2.2. *The quasi direct product* $Q = [L\colon S\colon L^d]$ *is a generalized inverse * semigroup. Conversely, if* T *is a generalized inverse * semigroup, then the mapping* $\phi\colon a \to (R_{aa^*}, a\mathscr{Y}, R_{a^*a})$ *is an isomorphism of* T *onto* $(E(T)/\mathscr{R}\colon T/\mathscr{Y}\colon (E(T)/\mathscr{R})^d]$.

PROOF. It will be shown first that $Q = [L\colon S\colon L^d]$ is a generalized inverse * semigroup. By Theorem 1.1, it will only be necessary to check that the unary operation * has the required properties. Notice that with L^d in the third coordinate, the multiplication in $Q = [L\colon S\colon L^d]$ becomes $(a, x, b)(c, y, d) = (au, xy, dv)$ where $u \in L_{(xy)(xy)^{-1}}$ and $v \in L_{(xy)^{-1}(xy)}$.

It is easy to see that $(a, x, b)^{**} = (a, x, b)$. Further $(a, x, b)(a, x, b)^* = (a, x, b)(b, x^{-1}, a) = (au, xx^{-1}, av)$ where $u, v \in L_{xx^{-1}}$. Noting that L_e is a left zero semigroup for all $e \in E$, $(au, xx^{-1}, av) = (a, xx^{-1}, a)$. Thus, $(a, x, b)(a, x, b)^*(a, x, b) = (a, xx^{-1}, a)(a, x, b) = (au, xx^{-1}x, bv) = (a, x, b)$ since $a, u \in L_{xx^{-1}}$ and $b, v \in L_{x^{-1}x}$. Finally, in order to compute $[(a, x, b)(c, y, d)]^*$, let $u \in L_{(xy)(xy)^{-1}}$ and $v \in L_{(xy)^{-1}(xy)}$. Then $[(a, x, b)(c, y, d)]^* = (au, xy, dv)^* = (dv, (xy)^{-1}, au) = (d, y^{-1}, c)(b, x^{-1}, a) = (c, y, d)^*(a, x, b)^*$. This completes the argument that $Q = [L\colon S\colon L^d]$ is a generalized inverse * semigroup.

Turning to the second part of the theorem, Theorem 2.1 says that the mapping $\theta\colon a \to (R_{aa^*}, a\mathscr{Y}, L_{a^*a})$ is an isomorphism of T onto $[E(T)/\mathscr{R}\colon T/\mathscr{Y}\colon E(T)/\mathscr{L}]$.

Consider now the map $f\colon L_p \to R_p$ of $E(T)/\mathscr{L}$ onto $E(T)/\mathscr{R}$ where p is a projection ($p = p^*$) in $E(T)$. This is a well defined map as each $\mathscr{L}[\mathscr{R}]$ class contains a unique projection [5, Theorem 2.2]. Also f is an antimorphism since $f[L_pL_q] = f[L_{pq}] = f[L_{(pq)^*(pq)}] = f[L_{q^*p^*pq}] = f[L_{qpq}] = R_{qpq} = R_{qp} = R_qR_p = f(L_q)f(L_p)$. Thus f is an isomorphism from $E(T)/\mathscr{L}$ onto $[E(T)/\mathscr{R}]^d$. It now follows that $\phi\colon a \to (R_{aa^*}, a\mathscr{Y}, R_{a^*a})$ is an isomorphism of T onto $[E(T)/\mathscr{R}\colon T/\mathscr{Y}\colon [E(T)/\mathscr{R}]^d]$. This ismorphism ϕ also preserves $*$ since $[\phi(a)]^* = (R_{aa^*}, a\mathscr{Y}, R_{a^*a})^* = (R_{a^*a}, (a\mathscr{Y})^{-1}, R_{aa^*}) = (R_{a^*a}, a^*\mathscr{Y}, R_{aa^*}) = \phi(a^*)$.

COROLLARY 2.3. *Let T be a generalized inverse semigroup. Then T admits an involution $*$ if and only if $E(T)$ admits an involution $*$. If T admits an involution $*$, it admits only one (up to isomorphism).*

PROOF. The uniqueness of $*$ on T, if it exists, follows directly from Theorem 2.2. Suppose that $E(T)$ admits $*$. Then $E(T) \cong E(T)/\mathscr{R} \otimes [E(T)/\mathscr{R}]^d$, the spined product of $E(T)/\mathscr{R}$ with $[E(T)/R]^d$. This is a part of the proof of Theorem 2.2. Alternately, it is proved directly in [9]. It follows that $T = [L\colon S\colon L^d]$ by Theorem 2.1. Thus T admits $*$ by Theorem 2.2.

3. The free generalized inverse * semigroup. The purpose of this section is to give a characterization of the free generalized inverse * semigroup. The construction depends upon the free inverse semigroup, so let us review that construction before proceeding.

Let X be a non-empty set and let $Y = X \cup X^{-1}$. Let F be the free semigroup on X and let G be the free group on X. Let R be the set of all reduced words in Y (x never stands next to x^{-1}) so that $G = R \cup \{1\}$ where 1 is the empty word. For each $y \in Y$, let $\bar{y}\colon G \to G$ be defined by

$$(v)\bar{y} = \begin{cases} 1 & \text{if } v = 1 \\ y^{-1} & \text{if } v = y \\ y^{-1}\cdot v & \text{otherwise.} \end{cases}$$

For $w = y_1y_2 \cdots y_n \in G$, let $\bar{w} = \bar{y}_1\bar{y}_2 \cdots \bar{y}_n$. When $A \subset G$, let $A\bar{w} = \{a\bar{w}\colon a \in A\}$.

Now let E be the set of all non-empty finite subsets of R which are closed under the operation of taking initial segments. Let $I = \{(A, w) \in E \times G\colon w \in A^1\}$. Define multiplication in I by $(A, w)(B, v) = (A \cup B(\bar{w})^{-1}, w\cdot v)$. Then $\mathscr{F} = (I, \varepsilon)$ is a free inverse semigroup on X where $\varepsilon\colon X \to I$ by $\varepsilon(x) = (\{x\}, x)$ [7].

The idempotents of I are the sets $(A, 1)$ where $A \in E$. The idempotent $(A, 1)$ is used to represent the product $\Pi_{w\in A}\ ww^{-1}$. For example, if $A = \{x, xy, xz^{-1}\}$, then $(A, 1) \equiv (xx^{-1})(xyy^{-1}x^{-1})(xz^{-1}zx^{-1}) = (xyy^{-1}x^{-1})(xz^{-1}zx^{-1})$. Notice that in a generalized inverse semigroup where the band is normal, these two idempotents would still commute since $(xy^{-1}yx^{-1})$

$(xz^{-1}zx^{-1}) = (xx^{-1})\ (xy^{-1}yx^{-1})\ (xz^{-1}zx^{-1})\ (xx^{-1}) = (xx^{-1})\ (xz^{-1}zx^{-1})\ (xy^{-1}yx^{-1})\ (xx^{-1}) = (xz^{-1}zx^{-1})\ (xy^{-1}yx^{-1})$. Of course, two idempotents ww^{-1} and vv^{-1} would always commute when w and v have the same first letter. Now let $C = \{x, xy, xz^{-1}, r, rs\}$. In an inverse semigroup, $\Pi_{w \in C}\ ww^{-1}$ would be $(xyy^{-1}x^{-1})(xz^{-1}zx^{-1})(rss^{-1}r^{-1})$. However, in a generalized inverse semigroup, these three idempotents would form four distinct products, namely $(xyy^{-1}x^{-1})\,(xz^{-1}zx^{-1})\,(rss^{-1}r^{-1})$, $(xyy^{-1}x^{-1})\,(rss^{-1}r^{-1})\,(xz^{-1}zx^{-1})$, $(rss^{-1}r^{-1})\,(xyy^{-1}x^{-1})\,(xz^{-1}zx^{-1})$, and $(rss^{-1}r^{-1})\,(xyy^{-1}x^{-1})\,(xz^{-1}zx^{-1})(rss^{-1}r^{-1})$. The order in which the idempotents appear really depends only on the first letter of the first word and the last letter of the last word.

Now let $L = \{(x, A) \in Y \times E\colon x \in A\}$. Define multiplication on L by $(x, A)(y, B) = (x, A \cup B)$. Then L is a left normal band with structure semilattice $E = E(I)$. L is constructed in such a way that the spined product $L \otimes L^d$ of L with the dual of L will play the role of the band of idempotents in the free generalized inverse * semigroup.

Let L be the left normal band constructed above, and let I be the free inverse semigroup. Let $[L\colon I\colon L^d]$ be the quasi direct product of L, I, and L dual. Define $i\colon X \to [L\colon I\colon L^d]$ by $i(x) = [(x, \{x\}), \varepsilon(x), (x^{-1}, \{x^{-1}\})]$.

THEOREM 3.1. *($[L\colon I\colon L^d]$, i) is a free generalized inverse * semigroup on the set X.*

PROOF. Let T be any generalized inverse * semigroup, and let $f\colon X \to T$. By Theorem 2.2, there is a left normal band M and an inverse semigroup S such that $T = [M\colon S\colon M^d]$.

Since $f\colon X \to [M\colon S\colon M^d]$, f factors into coordinate maps, $f = (\alpha, g, \beta)$, where $\alpha, \beta\colon X \to M = \bigcup_{e \in E} M_e$ and $g\colon X \to S$. Now extend $\alpha\colon X \to M$ to $\alpha\colon Y \to M$ by $\alpha(x^{-1}) = \beta(x)$ when $x \in X$.

Since $g\colon X \to S$ and (I, ε) is a free inverse semigroup on X, there is a homomorphism $\phi\colon I \to S$ such that $\phi \circ \varepsilon = g$.

Now define a mapping $\theta\colon [L\colon I\colon L^d] \to [M\colon S\colon M^d]$ by

$$\theta[(x, A), (A, w), (y, A\bar{w})] = [\alpha(x)u, \phi(A, w), \alpha(y)v]$$

where $u \in M_{\phi(A,1)}$ and $v \in M_{\phi(A\bar{w},1)}$.

It must be checked, of course, that the right hand member of equation (1) actually belongs to the * semigroup $[M\colon S\colon M^d]$. To do this, it will be necessary to show that $\alpha(x)u \in M_{[\phi(A,w)][\phi(A,w)]^{-1}}$ when $u \in M_{\phi(A,1)}$, that $\alpha(y)v \in M_{[\phi(A,w)]^{-1}[M(A,w)]}$ when $v \in M_{\phi(A\bar{w},1)}$, and that the products $\alpha(x)u$, $\alpha(y)v$ are independent of the choice of u, v. Notice that $[\phi(A, w)][\phi(A,w)]^{-1} = \phi[(A, w)(A, w)^{-1}] = \phi(A, 1)$ and $[\phi(A, w)]^{-1}[\phi(A, w)] = \phi[(A,w)^{-1}(A, w)] = \phi(A\bar{w}, 1)$. It will be shown that when $x \in A$, then $\alpha(x) \in M_e$ where $e \in E(S)$ and $e \geqq \phi(A, 1)$. The same argument will say that when $y \in A\bar{w}$, then $\alpha(y) \in M_f$ where $f \in E(S)$ and $f \geqq \phi(A\bar{w}, 1)$ From this will follow the desired result.

Suppose first that $x \in X$. Since $f(x) = [\alpha(x), g(x), \beta(x)] \in [L: S: L^d]$, $\alpha(x) \in M_{[g(x)][f(x)]^{-1}}$. Thus, it is enough to show that $[g(x)][g(x)]^{-1} \geqq \phi(A, 1)$. From $x \in A$ follows that $(\{x\}, x)(\{x\}, x)^{-1} \geqq (A, 1)$ so that $[\phi(\{x\}, x)][\phi(\{x\}, x)]^{-1} \geqq \phi(A, 1)$. Thus, $[g(x)][g(x)]^{-1} \geqq \phi(A, 1)$.

Now let $x = t^{-1}$ where $t \in X$. Since $f(t) = [\alpha(t), g(t), \beta(t)]$, $\alpha(x) = \beta(t) \in M_{[g(t)]^{-1}[g(t)]}$. So, it must be shown that $[g(t)]^{-1}[g(t)] \geqq \phi(A, 1)$. Now $(\{t\}, t)^{-1}(\{t\}, t) = (\{t^{-1}\}, 1) \geqq (A, 1)$ so that $[\phi(\{t\}, t)]^{-1}[\phi(\{t\}, t)] \geqq \phi(A, 1)$. Thus $[g(t)]^{-1}[g(t)] \geqq \phi(A, 1)$ as required.

It is now routine (although tedious) to calculate that ϕ is a homomorphism and that ϕ preserves the involution *. Further, when $x \in X$ then $(\theta \circ i)(x) = \theta[(x, \{x\}), (\{x\}, x), (x^{-1}, \{x^{-1}\})] = [\alpha(x)u, \phi(\{x\}, x), \alpha(x^{-1})v] = [\alpha(x)u, g(x), \alpha(x^{-1})v]$ where $u \in M_{[g(x)][g(x)]^{-1}}$ and $v \in M_{[g(x)]^{-1}[g(x)]}$. But $\alpha(x) \in M_{[g(x)][g(x)]^{-1}}$ so that $\alpha(x)u = \alpha(x)$. Similarly, $\alpha(x^{-1})v = \alpha(x^{-1})$. Thus, $(\theta \circ i)(x) = [\alpha(x), g(x), \alpha(x^{-1})] = [\alpha(x), g(x), \beta(x)] = f(x)$ so that $\theta \circ i = f$. Finally, it is easy to see that $i(X)$ generates $[L: I: L^d]$ since $\varepsilon(X)$ generates I. Thus, the homomorphism θ is unique.

COROLLARY 2.3. *The band of the free generalized inverse * semigroup* $[L: I: L^d]$ *is isomorphic to* $\{(x, A, y) \in Y \times E \times Y: x, y \in A\}$ *with multiplication* $(x, A, y)(r, B, s) = (x, A \cup B, s)$.

REMARK 2.4. Suppose now that $X = \{x\}$, a singleton set. Let I be the free semigroup on X with semilattice E and let G be the free generalized inverse * semigroup on X with band B. Of course, B has E as its structure semilattice. Every maximal rectangular subband of B will have either one or four elements. To see how this works, let us take two examples of elements of E.

If $A = \{x, x^2, x^3\} \in E$, then $(A, 1) \equiv x^3x^{-3}$ in E. In B, the rectangular band at A is just $(x, A, x) \equiv x^3x^{-3}$.

Now let $C = \{x, x^2, x^3, x^{-1}, x^{-2}\}$. Then $(C, 1) \equiv (x^3x^{-3})(x^{-2}x^2)$ in E. In B, the 2×2 rectangular band at A consists of

$$(x, A, x) \equiv (x^3x^{-3})(x^{-2}x^2)(x^3x^{-3})$$
$$(x, A, x^{-1}) \equiv (x^3x^{-3})(x^{-2}x^2)$$
$$(x^{-1}, A, x) \equiv (x^{-2}x^2)(x^3x^{-3})$$
$$(x^{-1}, A, x^{-1}) \equiv (x^{-2}x^2)(x^3x^{-3})(x^{-2}x^2).$$

One might wish to compare this with the band of H, where H is the free orthodox * semigroup on $X = \{x\}$. In the band of H, the sizes of the maximal rectangular subbands are unbounded [**8**].

4. The natural partial order. The aim of this section will be to discuss the natural partial order on a generalized inverse semigroup T. We shall also characterize this partial order when T has an involution.

First let S be any semigroup and let $E = E(S)$ be the set of idempotents of S. Recall that the set E has a partial order defined by $e \leqq f$ means $e = ef = fe$. In the case where S is an inverse semigroup with semilattice E the partial order on E extends to a partial order on S defined by $a \leqq b$ means $aa^{-1} = ba^{-1}$. This partial order is compatible with multiplication and inversion. It is equivalent to define $\leqq$ by $a \leqq b$ means $a^{-1}a = a^{-1}b$ [**2**].

Now let S be any regular semigroup with set E of idempotents. The partial order on E extends to a partial order on S defined by $a \leqq b$ means $a = eb = bf$ for some $e, f \in E$ [**4**]. This partial order $\leqq$ is compatible with multiplication if and only if S is pseudo-inverse, i.e., eSe is an inverse semigroup for each $e \in E$ [**6**].

Generalized inverse semigroups are pseudo-inverse [**6**]. Thus, the natural partial order on a generalized inverse semigroup is compatible with multiplication. Let us derive an alternate characterization of $\leqq$. In what follows, $V(a)$ will denote the set of inverses of a.

LEMMA 4.1. [**3**, Lemma 2.1]. *Let T be a generalized inverse semigroup and let $a, b \in T$. The following are equivalent.*

1. *There exists $e \in E(T)$ such that $a = be$.*
2. *For each $a' \in V(a)$, $a = ba'a$.*

COROLLARY 4.2. *Let T be a generalized inverse semigroup and let $a, b \in T$. The following are equivalent.*

1. There exists $e \in E(T)$ such that $a = be$.
2. There exists $a' \in V(a)$ such that $a = ba'a$, $aa' = ba'$.
3. For each $a' \in V(a)$, $a = ba'a$, $aa' = ba'$.

PROPOSITION 4.3. *Let T be a generalized inverse semigroup, The natural partial order $\leqq$ on T may be characterized by* (1) *$a \leqq b$ means there exists $a' \in V(a)$ such that $aa' = ba'$, $a'a = a'b$, or equivalently by* (2) *$a \leqq b$ means for each $a' \in V(a)$ then $aa' = ba'$, $a'a = a'b$.*

REMARK 4.4. Let T be a generalized inverse semigroup, say $T = [L: S: R]$ for a left (right) normal band $L(R)$ and an inverse semigroup S. Let $A = (a, x, b)$ and $B = (c, y, d) \in T$. We shall see that $A \leqq B$ if and only if $a \leqq c$, $x \leqq y$, and $b \leqq d$.

Assume first that $A \leqq B$. Let $A' = (r, x^{-1}, s) \in T$. It is routine to check that $A' \in V(A)$. Now $AA' = (a, x, b)(r, x^{-1}, s) = (au, xx^{-1}, vs)$ where $u \in L_{xx^{-1}}$, $v \in R_{xx^{-1}} = (a, xx^{-1}, s)$. Also $BA' = (c, y, d)(r, x^{-1}, s) = (cU, xy^{-1}, Vs)$ where $U \in L_{(yx^{-1})(yx^{-1})^{-1}}$ and $V \in R_{(yx^{-1})^{-1}(yx^{-1})}$. Since $AA' = BA'$, then $xx^{-1} = yx^{-1}$ and $a = cY$. From this follows $x \leqq y$ and $a \leqq c$. Similarly, $A'A = A'B$ implies $b \leqq d$.

Assume now that $a \leqq c$, $x \leqq y$, and $b \leqq d$. As before, let $A' = (r, x^{-1}, s) \in V(A)$. Since $x \leqq y$, then $xx^{-1} = yx^{-1}$. Thus $BA' = (c, y, d)(r, x^{-1}, s) = (ca, yx^{-1}, ss)$ (since $a \in L_{xx^{-1}}$ and $s \in R_{xx^{-1}}$) $= (a, xx^{-1}, s) = AA'$. Similarly, $A'A = A'B$.

COROLLARY 4.5. *Let* $(T, \cdot, {}^*)$ *be a generalized inverse* * *semigroup. The relation* $\leqq$ *defined by* $a \leqq b$ *means* $aa^* = ba^*$ *and* $a^*a = a^*b$ *is a partial order on* T *which is compatible with* $\cdot$ *and* *, *and which extends the natural partial order on* E.

ACKNOWLEDGEMENT. Corollary 2.3 is due to Celia L. Adair.

REFERENCES

1. C.L. Adair, *Varieties of* * *orthodox semigroups*, Ph.D. Thesis, University of South Carolina (1979).

2. A.H. Clifford and G.B. Preston, *The Algebraic Theory of Semigroups*, II, The Amer. Math. Soc. Math. Surveys **7** (1967).

3. S. Madhaven, *Some results on generalized inverse semigroups*, Semigroup Forum **16** (1978), 355–367.

4. K.S.S. Nambooripad, *The natural partial order on a regular semigroup*, Proc. Edinburgh Math. Soc., to appear.

5. T.E. Nordahl and H.E. Scheiblich, *Regular* * *semigroups*, Semigroup Forum **16** (1978), 369–377.

6. F. Pastijn, *The structure of pseudo-inverse semigroups*, preprint.

7. H.E. Scheiblich, *Free inverse semigroups*, Proc. of the Amer. Math. Soc. **38** (1973).

8. ———, *The free elementary* * *orthodox semigroup*, Semigroups, Academic Press (1980), 191–206.

9. ———, *Projective and injective bands with involution*, submitted.

10. Yamada, M., *Regular semigroups whose idempotents satisfy permutation identities*, Pacific Journal of Mathematics **21** (1967), 371–392.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA, SC 29208

ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 12, Number 2, Spring 1982

ON THE EXISTENCE OF UNIQUE EIGENSETS OF MONOTONE PROCESSES

WEI SHEN HSIA* AND BALAKRISH R. NATARAJAN

ABSTRACT. A Sufficient condition is given to guarantee the existence of a unique eigenset of a monotone process. Then, a special class of monotone processes is proved to have unique eigensets through this condition and the Perron-Frobenius Theorem.

1. Introduction. Rockafellar [**4**, p. 69, Theorem 4] proved a theorem which provides necessary and sufficient conditions for the existence of unique eigensets of monotone processes. Since those necessary and sufficient conditions must be satisfied by every pair of non-singular monotone sets in P_n and P_n^*, it is almost impossible to verify that a certain monotone process actually satisfies these conditions. In this paper, a sufficient condition in a simpler form is given to guarantee the existence of a unique eigenset. This sufficient condition in fact is a modification of Rockafellar's conditions. Then, a special class of monotone processes is proved to have unique eigensets through this modified condition and the Perron-Frobenius Theorem [**2**].

We shall only give the definitions of monotone sets, monotone processes, and eigensets of a monotone process. For more detailed definitions (e.g., positively homogeneous, sub-additive, non-singular, etc.), examples, and properties of monotone processes see [**3**], [**4**], and the references therein.

DEFINITION 1.1. [**4**, p. 11]. A monotone set of concave type in P_n, the nonnegative orthant of R^n, is a non-empty closed bounded convex set C such that $0 \leqq y_1 \leqq y_2 \in C$ implies $y_1 \in C$. A monotone set of convex type is a non-empty closed convex set such that $y_1 \geqq y_2 \in C$ implies $y_1 \in C$.

DEFINITION 1.2. [**4**, p. 9]. A monotone process of concave type from P_n to P_m is a nonnegative process T which is positively homogeneous, sub-additive, closed, and satisfies

(a) $T(x)$ is a monotone set of concave type for all $x \in P_n$, and

(b) $0 \leqq x_1 \leqq x_2$ implies $T(x_1) \subseteqq T(x_2)$.

AMS (MOS) subject classification (1970): Primary 15A18, 90A99.

*Sponsored by the University of Alabama under Project Number 951.

Received by the editors on March 25, 1980, and in revised form on Dec. 20, 1980.

Dually, T is a monotone process of convex type if conditions (a) and (b) are replaced by

(a′) $T(x)$ is a monotone set of convex type for all $x \in P_n$, and
(b′) $x_1 \geqq x_2 \geqq 0$ implies $T(x_1) \subseteqq T(x_2)$.

DEFINITION 1.3. [**4**, p. 58]. Let T be a non-singular monotone process from P_n to P_n. A non-singular monotone set C in P_n (of the same type as T) will be termed an eigenset of T if, for some $\lambda > 0$, $T(C) = \lambda C$.

Since this paper is mainly an extension of Rockafellar's result, we will adopt his notation and terminology freely.

2. Existence of unique eigensets of monotone processes. Let C and D be two monotone sets in P_n. We say $C \leqq D$ if and only if $\langle C, x^*\rangle \leqq \langle D, x^*\rangle$ for all $x^* \in P_n^*$, where P_n^* is the set of all nonegative linear functional on R^n [**4**, p. 16]. It is known that if C and D are monotone of concave type, then $C \leqq D$ if and only if $C \subseteqq D$; and if both are of convex type, then $C \leqq D$ if and only if $C \supseteqq D$.

We now define uniform convergence of a sequence of monotone sets. It is essentially the same as Rockafellar's Definition [**4**, p. 69], but it covers both concave and convex types.

DEFINITION 2.1. A sequence $C_1, C_2, \ldots$ of non-singular monotone sets of the same type in P_n converges uniformly to a set C_0 of the same type as each C_k if for every $\varepsilon > 0$ there exists a $k_0 = k_0(\varepsilon)$ such that

$$(1) \qquad (1 + \varepsilon)^{-1}C_0 \leqq C_k \leqq (1 + \varepsilon)C_0$$

for all $k \geqq k_0$.

If the sets C_k are of concave type, then (1) is equivalent to $(1 + \varepsilon)^{-1} C_0 \subseteqq C_k \subseteqq (1 + \varepsilon)C_0$ for all $k \geqq k_0$ as given by Rockafellar [**4**, p. 69]. If the sets C_k are of convex type, then (1) is equivalent to $(1 + \varepsilon)C_0 \subseteqq C_k \subseteqq (1 + \varepsilon)^{-1}C_0$ for all $k \geqq k_0$. In either case, we shall write $\lim_{k\to\infty} C_k = C_0$.

LEMMA 2.1. *If* $\lim_{k\to\infty} C_k = C_0$, *then* $\lim_{k\to\infty}\langle C_k, y^*\rangle = \langle C_0, y^*\rangle$ *for all* $y^* \in P_n^*$.

PROOF. Assume that the sets C_k and C_0 are of concave type. Then, given any $\varepsilon < 0$, there exists $k_0 = k_0(\varepsilon)$ such that $(1 + \varepsilon)^{-1}C_0 \subseteqq C_k \subseteqq (1 + \varepsilon)C_0$, for all $k \geqq k_0$. Therefore, for any $y^* \in P_n^*$, and for all $k \geqq k_0$,

$$(1 + \varepsilon)^{-1} \sup_{y \in C_0} \langle y, y^*\rangle \leqq \sup_{y \in C_k} \langle y, y^*\rangle \leqq (1 + \varepsilon) \sup_{y \in C_0} \langle y, y^*\rangle.$$

Hence, for all $y^* \in P_n^*$ and all $k \geqq k_0$, we have

$$(1 + \varepsilon)^{-1} \langle C_0, y^*\rangle \leqq \langle C_k, y^*\rangle \leqq (1 + \varepsilon) \langle C_0, y^*\rangle.$$

Thus, $\lim_{k\to\infty}\langle C_k, y^*\rangle = \langle C_0, y^*\rangle$ for all $y^* \in P_n^*$.

The same argument holds with the change of the direction of inequalities if C_k and C_0 are of convex type.

DEFINITION 2.2. A sequence $T_1, T_2, \ldots$ of monotone processes of the same type from P_n to P_n is said to converge uniformly to a monotone process T_0 of the same type if, given any $\varepsilon > 0$, there exists a $k_0 = k_0(\varepsilon)$ such that

$$(1 + \varepsilon)^{-1}T_0 \leqq T_k \leqq (1 + \varepsilon)T_0, \quad \text{for all } k \geqq k_0,$$

where $T_i \leqq T_j$ if and only if $T_i(x) \leqq T_j(x)$ [**4**, p. 17] for all $x \in P_n$.

In this event we write $\lim_{k\to\infty} T_k = T_0$. It is to be noted that in Definition 2.2, k_0 is independent of $x \in P_n$.

Let T be a non-singular monotone process of either type from P_n to P_n. Let $\bar{C}$ be a non-singular monotone set of the same type as T and $\bar{D}^*$ be a non-singular monotone set of type opposite to $\bar{C}$. Define a process T_0 from P_n to P_n by $T_0(x) = \langle x, \bar{D}^*\rangle\bar{C}$; then T_0 is a non-singular monotone process of the same type as T [**4**].

LEMMA 2.2. *Let T, T_0, $\bar{C}$, $\bar{D}^*$ be given as above. Let T^k be defined inductively by $T^k(x) = \bigcup_{y\in T^{k-1}(x)} T(y)$. If* $\lim_{k\to\infty} T^k = T_0$, *then*

(a) $\lim_{k\to\infty} T^k(x) = \langle x, \bar{D}^*\rangle \bar{C}$, *for all* $x \in P_n$, *and*

(b) $\lim_{k\to\infty} \langle T^k(C), D^*\rangle = \langle \bar{C}, D^*\rangle \cdot \langle C, \bar{D}^*\rangle$,

for all non-singular monotone sets C and D^* of types the same as $\bar{C}$ and $\bar{D}^*$, respectively.

PROOF. Given any $\varepsilon > 0$, there exists $k_0 = k_0(\varepsilon)$ such that for all $x \in P_n$ and for all $k \geqq k_0$, we have

$$(1 + \varepsilon)^{-1} T_0(x) \leqq T^k(x) \leqq (1 + \varepsilon)T_0(x). \tag{2}$$

By Definition 2.1, (2) implies $\lim_{k\to\infty} T^k(x) = T_0(x) = \langle x, \bar{D}^*\rangle \bar{C}$, for all $x \in P_n$.

Let C be any non-singular monotone set in P_n of the same type as $\bar{C}$. Then from (2), we have

$$(1 + \varepsilon)^{-1} \bigcup_{x\in C} T_0(x) \leqq \bigcup_{x\in C} T^k(x) \leqq (1 + \varepsilon) \bigcup_{x\in C} T_0(x),$$

for all $k \geqq k_0$. This implies $(1 + \varepsilon)^{-1}T_0(C) \leqq T^k(C) \leqq (1 + \varepsilon) T_0(C)$, for all $k \geqq k_0$. Hence, $\lim_{k\to\infty} T^k(C) = T_0(C)$. But, in [**4**, p. 69], it is shown that $T_0(C) = \langle C, \bar{D}^*\rangle \cdot \bar{C}$. Therefore,

$$\lim_{k\to\infty} T^k(C) = \langle C, \bar{D}^*\rangle \cdot \bar{C}. \tag{3}$$

Now, let D^* be any non-singular monotone set in P_n^* of the same type as $\bar{D}^*$. Then, (3) implies

$$\langle \lim_{k\to\infty} T^k(C), D^*\rangle = \langle C, \bar{D}^*\rangle \cdot \langle \bar{C}, D^*\rangle.$$

Applying Lemma 2.1 to the left hand side of this equation, we have

$$\lim_{k\to\infty} \langle T^k(C), D^*\rangle = \langle C, \bar{D}^*\rangle \cdot \langle C, D^*\rangle$$

The conclusions (a) and (b) in Lemma 2.2 are equivalent to the necessary and sufficient conditions given by Rockafellar [**4**, p. 69, Theorem 4] for the existence of a unique eigenset of a monotone process. Lemma 2.2 actually shows that uniform convergence of the sequence $T, T^2, \ldots$ to T_0 guarantees the existence of a unique eigenset for T. This conclusion can be rewritten as the following theorem.

THEOREM 2.1. *Let T be a non-singular monotone process of either type from P_n to P_n. If there exist non-singular monotone sets $\bar{C}$ and $\bar{D}^*$ of suitable types and there exists a scalar $\lambda > 0$ such that*

$$\lim_{k\to\infty} \left(\frac{1}{\lambda} T\right)^k = T_0, \tag{4}$$

where $T_0(\cdot) = \langle \cdot, \bar{D}^*\rangle \bar{C}$, then, aside from positive multiples, $\bar{C}$ and $\bar{D}^*$ are the unique non-singular eigensets of T and the adjoint of T, T^*, respectively. This means $T(\bar{C}) = \lambda\bar{C}$, and $T^*(\bar{D}^*) = \lambda\bar{D}^*$.

The scalar λ is known as the growth rate in the literature (e.g., [**4**]). Since a monotone process is possitively homogeneous, we can replace $[(1/\lambda)T]$ by T and assume $\lambda = 1$. For this reason, (4) is the sufficient condition in Lemma 2.2.

In the next section, we shall describe a special class of monotone processes satisfying (4) whose members therefore have unique eigensets.

3. Application. Let A be an $n \times n$ matrix such that none of its rows is identical to the zero vector. Define a process $A^\wedge$ from P_n to P_n by $A^\wedge(x) = (Ax)^\wedge = \{y \mid 0 \leqq y \leqq Ax\}$. Then $A^\wedge$ is a monotone process of concave type, and the adjoint of $A^\wedge$, $(A^\wedge)^*$, is a monotone process of convex type where $(A^\wedge)^*(x)^* = (A^t x^*)^\vee = \{y^* \mid y \geqq A^t x^*\}$ [**4**, p. 9].

In this section, we shall prove that the monotone process $A^\wedge$ defined by a nonnegative matrix A has unique eigenset.

First let us cite several known results in matrix theory.

LEMMA 3.1. [**1**]. *Let P be an $n \times n$ irreducible stochastic matrix. Then*

(a) *$P^\infty = \lim_{k\to\infty} P^k$ exists, i.e., for every $\varepsilon > 0$, there exists $k_0 = k^0(\varepsilon)$ such that for all $k \geqq k_0$ and $i, j = 1, \ldots, n$,*

$$(1 - \varepsilon)P^\infty_{ij} \leqq P^k_{ij} \leqq (1 + \varepsilon)P^\infty_{ij}$$

(b) *Furthermore, there exists a vector $\pi = (\pi_1, \ldots, \pi_n)$, where $\sum_{j=1}^n \pi_j = 1$ and $\pi_j > 0$ for $j = 1, \ldots, n$, such that each row of P^∞ is equal to π.*

THEOREM 3.1. (PERRON-FROBENIUS THEOREM [2]). *Let A be an $n \times n$ irreducible nonnegative matrix. Then A has a "maximal" positive eigenvalue λ_0 that is a simple root of the characteristic equation such that $|\lambda| \leqq \lambda_0$ for other eigenvalues λ of A. Furthermore, to this λ_0, there corresponds an eigenvector $z^0 = (z_1^0, \ldots, z_n^0)$ such that each component of z^0 is greater than zero.*

Now, if A is an $n \times n$ irreducible nonnegative matrix, then A is similar to some matrix $(\lambda_0 P)$, where λ_0 is the "maximal" eigenvalue of A given in Theorem 3.1, and P is an irreducible probability matrix. In fact, c.f. [2], we have

$$A = Z(\lambda_0 P)Z^{-1}, \tag{5}$$

where Z is a diagonal matrix with diagonal elements $z_1^0, \ldots, z_n^0$ and $(z_1^0, \ldots, z_n^0)$ is a positive eigenvector of A corresponding to λ_0.

From (5), we have $[(1/\lambda_0)A]^k = ZP^kZ^{-1}$, for all positive integers k. Therefore,

$$\lim_{k\to\infty}\left(\frac{1}{\lambda_0} A\right)^k = \left(\frac{1}{\lambda_0} A\right)^\infty = ZP^\infty Z^{-1} \tag{6}$$

exists.

If we apply (a) of Lemma 3.1 to (6), it is not difficult to prove that, given $\varepsilon > 0$, there exists $k_0 = k_0(\varepsilon)$ such that for all $k \geqq k_0$ and $i, j = 1, \ldots, n$,

$$(1 - \varepsilon)\left(\frac{1}{\lambda_0} A\right)_{ij}^\infty \leqq \left(\frac{1}{\lambda_0} A\right)_{ij}^k \leqq (1 + \varepsilon)\left(\frac{1}{\lambda_0} A\right)_{ij}^\infty \tag{7}$$

If A is an $n \times n$ irreducible nonnegative matrix, then so is A^t. By the Perron-Frobenius Theorem, there exists a vector w^0 with all components positive such that $A^t w^0 = \lambda_0 w^0$. Therefore $([(1/\lambda_0]A)^k)^t\, w^0 = w^0$, for all integers k. Hence, we have

$$\left(\left(\frac{1}{\lambda_0} A\right)^\infty\right)^t w^0 = w^0. \tag{8}$$

Applying (6) in (8), using the representation of P^∞ described in Lemma 3.1 and then equating the components on both sides in (8), we get $\langle z^0, w^0\rangle \cdot \pi_j / z_j^0 = w_j^0$ for $j = 1, \ldots, n$.
Hence,

$$\pi_j = z_j^0 w_j^0 / \langle z^0, w^0\rangle, \text{ for } j = 1, \ldots, n. \tag{9}$$

Now, we are ready to prove the main result of this section.

THEOREM 3.2. *Let A be an $n \times n$ irreducible non-negative matrix. Then the monotone process $T = A^\wedge$ defined by A and its adjoint, T^*, have unique non-singular eigensets, expect for positive scalar multiples.*