Semigroup Forum Volume 45 1992

Printed in the United States of America

INDEX TO VOL. 45

Semigroup Forum Vol. 45 (1992) 1–8

RESEARCH ARTICLE

On a Multiplication of Semigroup Varieties

Ju. G. Koŝelev

Communicated by J. S. Ponizovskii

The role of multiplication of group varieties is well known [2], as well as the role of wreath products in such a construction [1]. In the present paper the idea of exploiting wreath products is used for defining a multiplication for semigroup varieties which differs from the operations given in [3]. In section 5 we also consider some elementary properties of the defined multiplication. Some results of the paper were announced in [9].

1. Preliminaries

For any semigroup A we denote by A^1 a semigroup with an adjoined identity 1 (even in the case when A possesses an identity itself). For each set I and for a semigroup A, A^I denotes a semigroup of all maps $I \to A$ under pointwise multiplication.

If I is a set, then $(I)P$ ($P(I)$) stands for a semigroup of all transformations of I acting on elements of I on the right (on the left). *A representation* (*an antirepresentation*) of a semigroup B by transformations of I is a semigroup homomorphism

$$\theta : B \to (I)P \quad (\theta : B \to P(I)) .$$

A representation $\theta : B \to (B)P$ defined by the rule

$$(x)(\theta(b)) = xB \quad (\forall\, x \in B)$$

is called *the standard representation*; the same rule defines a representation $\theta^1 : B \to (B^1)P$ called *the extended standard representation.* If there is no possibility for confusion, and for brevity, we write ib (respectively bi) instead of $(i)(\theta(b))$ (instead of $(\theta((b))(i))$) in the case when θ is a representation (an antirepresentation) of B.

Let A and B be semigroups, and let θ be a representation of B by transformations of a set I. Consider the set $A^I \times B$, and define a multiplication in $A^I \times B$ as follows:

$$(f_1, b_1)(f_2, b_2) = (f_1 \cdot b_1 f_2 \cdot b_2) \quad (\forall \;\; f_1, f_2 \in A^I, b_1, b_2 \in B) .$$

Then $(A^I \times B, \cdot)$ becomes a semigroup denoted by $AWr(\theta)B$ and called *a wreath product of A and B corresponding to a representation θ.* In the case when θ is standard (extended standard) the semigroup $AWr(\theta)B$ is denoted by $AWrB$ (AWr^1B) and called *a standard* (*an extended standard*) *wreath product of A and B.*

A direct product of semigroups A_l ($l \in L$) is denoted by $\prod_{l \in L} A_l$.

Let A and B be semigroups, and let $h : B \to \operatorname{End} A$ be an antirepresentation of B. Then $A\lambda(h)B$ is a semigroup $(A \times B, \cdot)$ under the multiplication

$$(A_1, b_1)(a_2, b_2) = (a_1 \cdot b_1 a_2, b_1 b_2) ;$$

this semigroup is called *a semidirect product of A and B with respect to h*. Note that $AWr(\theta)B$ is a semidirect product of A^I and B.

Let K be a class of semigroups. Define operations S, C, Q on K as follows:

$$SK = \left\{T/T \text{ is a subsemigroup of some } A \in K\right\} ;$$

$$CK = \left\{T/T = \prod_{l \in L} A_l, \; A_l \in K \text{ for all } 1 \in L \text{ and for some } L\right\} ;$$

$$QK = \left\{T/T \text{ is an epimorphic image of some } A \in K\right\} .$$

It is well known that, for any class K of semigroups the class $QSCK$ is a variety. A class of semigroups which is closed under operations S, Q and under finite direct proucts is called a *pseudo-variety.*

Let $X = \{x_1, x_2, \ldots\}$ be a countable alphabet, and $F(X)$ a corresponding free semigroup. Then a semigroup identity is a formal equality of the type $u = v$ with $u, v \in F(X)$.

2. The multiplication of semigroup varieties

Consider two semigroup varieties V_1, V_2 and define $K = \{AWr^1B/A \in V_1, B \in V_2\}$. We define a product of V_1, V_2 as a class of semigroups $V_1 \cdot V_2 = QSK$.

Theorem 1. *A product of two semigroup varieties is again a semigroup variety.*

Proof. We have only to prove that $QSK = ASCK$. Evidently, $QSK \subset QSCK$. Conversely, if $M \in QSCK$, then $M = \lambda(M_0)$, where $\lambda : M_0 \to M$ is an epimorphism of semigroups and M_0 is a subsemigroup of some $\prod\limits_{1 \in L} (A_1 Wr^1 B_1)$ with $A_l \in V_1$, $B_1 \in V_2$ for all $l \in L$. Suppose that there exists a monomorphism (an embedding in what follows)

$$\mu : \prod_{l \in L} (A_l Wr^1 B_l) \to A = \left(\prod_{l \in L} A_1\right) Wr^1 \left(\prod_{l \in L} B_l\right) .$$

Then the restriction of μ to M_0 is an embedding of M_0 into A. Now we have $\prod\limits_{l \in L} A \in V_1$ and $\prod\limits_{l \in L} B \in V_2$, whence $\mu(M_0) \in SK$ and $M = (\lambda\mu^{-1})(\mu(M_0)) \in QSK$. The following lemma completes the proof of Theorem 1.

Lemma 1. *Let L be a set, and let A_l and B_l be semigroups for all $l \in L$. Then there exists an embedding*

$$\mu : \prod_{l \in L} (A_l Wr^1 B_l) \to \left(\prod_{l \in L} A_1\right) Wr^1 \left(\prod_{l \in L} B_l\right) .$$

Proof. From now on, we assume that for $1 \in \left(\prod_{l\in L} B_l\right)^1$, one has $(l)1 = 1_1 \in B_l$ for all $l \in L$. Suppose that for any $v \in \prod_{l\in L}(A_l Wr^1 B_l)$ and for any $l \in L$, we have $(l)v(f_l, b_l)$; if this is the case, we put

$$v^\mu = (\varphi, b) \in \left(\prod_{l\in L} A_l\right) Wr^1 \left(\prod_{l\in L} B_l\right) ,$$

where $b \in \prod_{l\in L} B_l$ is such that $(l)b = b_l$ for all $l \in L$, and $\varphi \in \left(\prod_{l\in L} A_l\right)^{\left(\prod_{l\in L} B_1\right)^1}$ is such that

$$(l)((\beta)\varphi) = ((\beta)\varphi) = ((1)\beta)\, f_l = (\beta_l) f_l \quad \text{for any} \quad \beta \in \left(\prod_{l\in L} B_l\right)^1$$

and any $l \in L$. We claim that μ is a semigroup homomorphism. Take $v_1 \in \prod_{l\in L} \left(A_l Wr^1 B_l\right)$ and assume

$$(l)v_1 = \left(f_l^1, b_l^1\right),\ v_1^\mu = \left(\varphi^1, b^1\right),\ (l)b^1 = b_l^1 ,$$
$$(l)\left((\beta)\varphi^1\right) = ((l)\beta)\, f_l^1 = (\beta_l)\, f_l^1$$

for all $l \in L$ and all $\beta \in \left(\prod_{l\in L} B_l\right)^1$. Then

$$(l)vv_1 = (l)v \cdot (l)v_1 = (f_l, b_l)\left(f_l^1, b_l^1\right) = \left(f_l \cdot b_l f_l^1, b_l b_l^1\right) .$$

Hence $(vv_1)^\mu = (\varphi', b')$, where $(l)b' = b_l b_l^1 = b_l'$ and

$$(l)\left((\beta)\varphi'\right) = ((l)\beta)\left(f_l \cdot b_l f_l^1\right) = ((l)\beta)\, f_l \cdot ((l)\beta b)\, f_l^1$$

for all $\beta \in \left(\prod_{l\in L} B_l\right)^1$. But $v^\mu v_1^\mu = (\varphi, b)\left(\varphi^1, b^1\right) = \left(\varphi \cdot b\varphi^1, bb^1\right)$, and $(l)bb^1 = $
$= (l)b \cdot (l)b^1 = b_l \cdot b_l^1 = b_l'$. For any $\beta \in \left(\prod_{l\in L} B_l\right)^1$ we have $(l)\left((\beta)\left(\varphi \cdot b\varphi^1\right)\right) =$
$(l)\left((\beta)\varphi \cdot (\beta b)\varphi^1\right) = (l)\left((\beta)\varphi\right)(l)\left((\beta b)\varphi^1\right) = ((l)\beta)\, f_l \cdot ((l)(\beta b))\, f_l^1 = (l)\left((\beta)\varphi'\right)$.
Thus, $\varphi \cdot b\varphi^1 = \varphi'$ and $(vv_1)^\mu = v^\mu v_1^\mu$ as desired. Now we have to prove that μ is monic. Assume that $(\varphi, b) = v^\mu = v_1^\mu = \left(\varphi^1, b^1\right)$ in the above notation. Then $b_l = (l)b = (l)b^1 = b_l^1$ for any $l \in L$. Also we have

$$((l)\beta)\, f_l = (l)\left((\beta)\varphi\right) = (l)\left((\beta)\varphi^1\right) = ((l)\beta)\, f_l^1$$

for all $\beta \in \left(\prod_{l\in L} B_l\right)$. For any $b_l \in B_l^1$, we can take β such that $b_l = (l)\beta$, whence $f_l = f_l^1$. Thus $(l)v = (l)v_1$ for all $l \in L$. That means that $v = v_1$, and μ is monic.

This completes the proof of Lemma 1 as well as that of Theorem 1.

3. The associativity of multiplication

Theorem 2. *The multiplication of semigroup varieties (defined in Section 2) is associative.*

Let V_1, V_2 and V_3 be semigroup varieties. We have to show that $(V_1 \cdot V_2) \cdot V_3 = V_1 \cdot (V_2 \cdot V_3)$. The inclusion $(V_1 \cdot V_2) \cdot V_3 \subset V_1 \cdot (V_2 \cdot V_3)$ is immediate from the following

Lemma 2. *Let A, B and C be semigroups. There exists an embedding*

$$\mu : (AWr^1B)Wr^1C \to AWr^1(BWr^1C) .$$

Proof. Define

$$\mu \begin{cases} (AWr^1B)Wr^1C \to AWr^1(BWr^1C) \\ v \to v^\mu \end{cases}$$

by the rule: if $v = (g_1 < c_1) \in (AWr^1B)Wr^1C$ is such that $(c)g_1 = \left(f_c^1, b_c^1\right)$ for all $c \in C^1$; then $v^\mu = (\varphi_1, (\psi_1, c_1))$, where $(c)\psi_1 = b_c^1$ and $(\psi, c)\varphi_1 = ((1)\psi)\, f_c^1$ for all $c \in C^1$ and for all $(\psi, c) \in (BWr^1C)^1$. Note that we put $1 = (1,1) \in (BWr^1C)^1$ as well as assume $(c)1 = 1 \in B^1$ for all $1 \in (B^{C^1})^1$ and for all $c \in C^1$. We are going to show that μ is a semigroup homomorphism.

Let $v_1 = v = (g_1, c_1)$ be as defined above, and let $v_2 = (g_2, c_2)$ such that $(c)g_2 = (f_c^2, b_c^2)$ for all $c \in C^1$. Then $v_2^\mu = (\varphi_2, (\psi_2, c_2))$ with $(c)\psi_2 = b_c^2$ and $(\psi, c)\varphi_2 = ((1)\psi)\, f_c^2$ for all $c \in c^1$ and for all $(\psi, c) \in (BWr^1C)^1$. Now $v_1v_2 = (g_1, c_1)(g_2, c_2) = (g_1 \cdot c_1g_2, c_1c_2)$ and $(c)(g_1 \cdot c_1g_2) = (c)g_1 \cdot (cc_1)\, g_2 = \left(f_c^1, b_c^1\right)\left(f_{cc_1}^2, b_{cc_1}^2\right) = \left(f_c^1 \cdot b_c^1 f_{cc_1}^2, b_c^1 \cdot b_{cc_1}^2\right)$ for any $c \in C^1$. Then $(v_1v_2)^\mu = (\varphi', (\varphi', c'))$ where $c' = c_1c_2$, and $(\psi, c)\varphi' = (1)\psi)\left(f_c^1 \cdot b_c^1 f_{cc_1}^2\right) = ((1)\varphi)\, f_c^1 \cdot \left((1)\psi b_c^1\right) f_{cc_1}^2$ for all $c \in C^1$ and for all $(\psi, c) \in (BWr^1C)^1$. But we have

$$\begin{aligned} v_1v_2 &= (\varphi_1, (\psi_1, c_1))(\varphi_2, (\psi_2, c_2)) = \\ &= (\varphi_1)\psi_1, c_1)\varphi_2, (psi_1 \cdot c_1\psi_2, c_1c_2)) , \end{aligned}$$

and $(c)(\psi_1, c_1\psi_2) = (c)\psi_1(cc_1)\psi_2 = b_c^1 b_{cc_1}^2$ for all $c \in C^1$, and

$$\begin{aligned} (\psi, c)(\varphi_1, (\psi_1, c_1)\varphi_2) &= (\psi, c)\varphi_1 \cdot (\psi \cdot c\psi_1, cc_1)\, \varphi_2 = \\ &= ((1)\psi)\, f_c^1 \cdot ((1)(\psi \cdot c\psi_1))\, f_{cc_1}^2 = \\ &= ((1)\psi)\, f_c^1\, ((1)\psi \cdot (c)\psi_1)\, f_{cc_1}^2 = \\ &= ((1)\psi)\, f_c^1 \cdot \left((1)\psi b_c^1\right) f_{cc_1}^2 \end{aligned}$$

for any $(\psi, c) \in (BWr^1C)^1$. Hence $v_1^\mu v_2^\mu = (v_1v_2)^\mu$ as desired. Let us show that μ is monic. Suppose (in the above notation) that $v_1^\mu = (\varphi_1, (\psi_1, c_1)) = (\varphi_2, (\psi_2, c_2)) = v_2^\mu$. Then $c_1 = c_2$, and $b_c^1 = (c)\psi_1 = (c)\psi_2 = b_c^2$ for all $c \in C^1$, and $((1)\psi)\, f_c^1 = (\psi, c)\varphi_1 = (\psi, c)\varphi_2 = ((1)\psi)\, f_c^2$ for each $(\psi, c) \in (BWr^1C)^1$. Since the last chain of equalities holds for all ψ, we see that $f_c^1 = f_c^2$. Hence $(c)v_1 = (c)v_2$ for all $c \in C^1$ and $v_1 = v_2$.

The proof of Lemma 2 is complete.

The converse inclusion $V_1 \cdot (V_2 \cdot V_3) \subset (V_1 \cdot V_2) \cdot V_3$ follows from Lemmas 2, 3.

Let A, B and C be semigroups. Consider a wreath product $AWr(\theta)(BWr^1C)$ corresponding to a representation θ of BWr^1C by transformations of a set $(B^{C^1})\times C^1$ defined so that the equality $(\beta, c(\psi_1, c_1) = (\beta\psi_1, cc_1)$ holds for all $(\beta, c) \in (B^{C^1}) \times C^1$ and for all $(\psi_1, c_1) \in BWr^1C$. Hence the set of elements of a semigroup $AWr(\theta)(BWr^1C)$ is a set

$$A^{(B^{C^1})\times C^1} \times B^{C^1} \times C \ .$$

Lemma 3. *Let A, B and C be semigroups. There exists an embedding*

$$\mu : AWr(\theta)(BWr^1C) \to (AWr^1B^{C^1})Wr^1C \ .$$

Proof. Define

$$\mu \begin{cases} AWr(\theta)(BWr^1C) \longrightarrow (AWr^1B^{C^1})Wr^1C \\ W = (\varphi_1, (\psi_1, c_1)) \longmapsto w^\mu = (g_1, c_1) \end{cases}$$

by the rule: $(c)g_1 = (f_c^1, c\psi_1)$, $(\beta)f_c^1 = (\beta, c)\varphi_1$ for all $c \in C^1$ and for all $\beta \in (B^{C^1})^1$. Let

$$w_i = (\varphi_i, (\psi_i c_i)), \quad i = 1, 2 \ .$$

We have to prove that $(w_1w_2)^\mu = w_1^\mu w_2^\mu$. We obtain

$$\begin{aligned} w_1w_2 &= (\varphi_1, (\psi_1, c_1))(\varphi_2, (\psi_2, c_2)) = \\ &= (\varphi_1(\psi_1, c_1)\varphi_2, (\psi_1 \cdot c_1\psi_2, c_1c_2)) \ , \end{aligned}$$

and $(w_1w_2)^\mu = (g, c_1c_2)$ with $(c)g = (f_c, c\psi_1cc_1\psi_2)$,

$$\begin{aligned} (\beta)f_c &= (\beta, c)(\varphi_1 \cdot (\psi_1, c_1)\psi_2) = \\ &= (\beta, c)\varphi_1 \cdot (\beta \cdot c\psi_1, cc_1)\varphi_2 \end{aligned}$$

for all $c \in C^1$ and all $\beta \in (B^{C^1})^1$. On the other hand, $w_1^\mu w_2^\mu = (g_1, c_1)(g_2, c_2) = (g_1 \cdot c_1g_2, c_1c_2)$. For any $c \in C^1$, we have

$$\begin{aligned} (c)(g_1 \cdot c_1g_2) &= (c)g_1 \cdot (cc_1)g_2 = \left(f_c^1, c\psi_1\right)\left(f_{cc_1}^2, cc_1\psi_2\right) = \\ &= \left(f_c^1 \cdot c\psi_1 f^2 cc_1, c\psi_1 \cdot cc_1\psi_2\right) \ . \end{aligned}$$

We also have $(\beta)\left(f_c^1 \cdot c\psi_1 f_{cc_1}^2\right) = (\beta)f_c^1 \cdot (\beta \cdot c\psi_1)f_{cc_1}^2 = (\beta, c)\varphi_1(\beta \cdot c\psi_1, cc_1)\varphi_2$. Therefore, $w_1^\mu w_2^\mu = (w_1w_2)^\mu$.

Suppose now that $w_1^\mu = (g_1, c_1) = (g_2, c_2) = w_2^\mu$. Then $c_1 = c_2$, and $(f_1^1, \psi_1) = (1)g_1 = (1)g_2 = (f_1^2, \psi_2)$. Hence $\psi_1 = \psi_2$. Take

$$(\beta, c) \in \left(B^{C^1}\right) \times C^1 \ .$$

We have $\beta, c)\varphi_1 = (\beta)f_c^1 = (\beta)f_c^2 = (\beta, c)\varphi_2$. That means $w_1 = w_2$ and μ is monic. The proof of Lemma 3 is complete.

Lemma 4. *The semigroup $AWr^1(BWr^1C)$ is a homomorphic image of the semigroup $AWr(\theta)(BWr^1C)$.*

Proof. Define

$$\mu \begin{cases} AWr(\theta)(BWr^1C) \longrightarrow AWr^1(BWr^1C) \\ W = (\varphi, (\psi, c)) \longmapsto w^\mu = (\overline{\varphi}, (\psi, c)) \end{cases}$$

where $\overline{\varphi}$ is a restriction of φ to $\left(B^{C^1} \times C\right)^1$. We put $1 = (1,1) \in \left(B^{C^1}\right) \times C^1$ for $1 \in \left(B^{C^1} \times C\right)^1$.

Clearly μ is epic. Let us show that μ is a semigroup homomorphism. In fact, let

$$w_i = (\varphi_i, (\psi_i, c_i)), \; w_i^\mu = (\overline{\varphi}_i, (\psi_i, c_i)), \quad i = 1, 2 .$$

We have

$$(w_1 w_2)^\mu = \left(\overline{\varphi_1 \cdot (\psi_1, c_1)\, \varphi_2}, (\psi_1 \cdot c_1 \psi_2, c_1 c_2)\right) ,$$
$$w_1^\mu w_2^\mu = (\overline{\varphi}_1 \cdot (\psi_1, c_1)\, \overline{\varphi}_2, (\psi_1 \cdot c_1 \psi_2, c_1 c_2)) .$$

Claim that

$$\overline{\varphi_1 \cdot (\psi_1, c_1)\, \varphi_2} = \overline{\varphi}_1 \cdot (\psi_1, c_1)\, \overline{\varphi}_2 .$$

If (β, c) is not the identity of $\left(B^{C^1}\right)^1 \times C^1$, then $(\beta\psi_1, cc_1)$ is also not the identity. Hence

$$\begin{aligned}(\beta, c)\left(\overline{\varphi_1 \cdot (\psi_1, c_1)\, \varphi_2}\right) &= (\beta, c)\,(\varphi_1, c_1)\varphi_2) = \\ &= (\beta, c)\varphi_1 \cdot (\beta\psi_1, cc_1)\, \varphi_2 = (\beta, c)\overline{\varphi}_1\, (\beta\psi_1, cc_1)\, \overline{\varphi}_2 = \\ &= (\beta, c)\, (\overline{\varphi}_1 (\psi_1, c_1) \overline{\varphi}_2) .\end{aligned}$$

We have also, for $1 \in \left(B^{C^1} \times C\right)^1$,

$$\begin{aligned}(1)\left(\overline{\varphi_1 \cdot (\psi_1, c_1)\, \varphi_2}\right) &= (1)\,(\varphi_1 \cdot (\psi_1, c_1)\, \varphi_2) = (1)\varphi_1 \cdot (\psi_1, c_1)\, \varphi_2 = \\ &= (1)\overline{\varphi}_1 \cdot (\psi_1, c_1)\, \overline{\varphi}_2 = (1)\,(\overline{\varphi}_1 \cdot (\psi_1, c_1)\, \overline{\varphi}_2) ,\end{aligned}$$

since ψ_1 is not the identity of $\left(B^{C^1}\right)^1$ as well as c_1 is not the identity of C^1. This completes the proof of Lemma 4.

Thus the proof of Theorem 2 is complete, too.

4. Other characterizations of the multiplication of semigroup varieties

For any two semigroup varieties V_1 and V_2, we introduce a class K in Section 2. Consider also classes K_1 and K_2 where $K_1\,(K_2)$ is a class of all semigroups of the form $A\lambda(h)B$ (of the form $AWr(\theta)B$) with $A \in V_1$, $B \in V_2$.

Theorem 3. *Let V_1 and V_2 be semigroup varieties. Then $V_1 \cdot V_2 = QSK_1 = QSK_2$.*

Proof. Let A and B b semigroups, let I be a set, and θ a representation of B by transformations of I. It has been observed in Section 1 that in such a case a wreath product $AWr(\theta)b$ is a semidirect product of semigroups A^I and B. Hence, $K_2 \subset K_1$, so that $QSK_2 \subset QSK_1$. In [8] it was shown that any semidirect product $A\lambda(h)B$ of A and B can be embedded into the extended standard wreath product AWr^1B. Thus, up to isomorphisms, $K_1 \subset SK$. Therefore, $QSK_1 \subset QSK$. Clearly, $K \subset K_2$, so that, in particular $QSK \subset QSK_2$. We obtain

$$QSK_2 \subset QSK_1 \subset QSK \subset QSK_2 ,$$

hence $V_1 \cdot V_2 = QSK = QSK_1 = QSK_2$. The proof of Theorem 3 is complete.

In [6], an associative multiplication of monoid varieties was defined by using the standard wreath product. But this idea does not work for arbitrary semigroup varieties because such operations turn out not to be associative in general.

It is worth mentioning that the multiplication of semigroup varieties defined in the present paper can be extended to semigroup pseudovarieties. Indeed, let V_1 and V_2 be two semigroup pseudovarieties. Now one can define classes K, K_1 and K_2 as has been done for semigroup varieties, with one extension: in the definition of K_2, one can use finite sets I only. Then $V_1 \cdot V_2 = QSK = QSK_1 = QSK_2$ is a pseudovariety too. This follows easily from the fact that, using Lemma 1, one can confine oneself to finite direct products; as for Theorem 3, its statement holds for pseudovarieties as well.

Observe that this approach to the multiplication of semigroup varieties (pseudovarieties) is constructive. It makes it possible to do first steps in the problem of the investigation of the monoid of semigroup varieties (pseudovarieties). This approach differs from that of [6], where a product of some semigroup classes (in particular, semigroup varieties and semigroup pseudovarieties) was introduced, though in cases of varieties and pseudovarieties both methods coincide.

5. Some elementary properties of the semigroup of varieties

Let $\mathcal{M}$ be a set of all semigroup varieties. Clearly, the trivial variety $V(x = y)$ is the left identity and the variety $V(x = x)$ is the zero under the multiplication of semigroup varieties defined in the present paper. Thus $\mathcal{M}$ is a semigroup with zero.

In [7], the following system of semigroup idetities was considered:

$$(1) \qquad \begin{cases} s_{\infty\infty} : x = x \\ s_{ll} x_1 x_2 \ldots x_l = x_1 x_2 \ldots x_l x_{l+1} \\ s_{nl} : x_1 x_2 \ldots x_n x_{n+1} \ldots x_l = x_1 x_2 \ldots x_n x_{l+1} \ldots x_{2l-n} \end{cases}$$

where n and l are integers, $0 \leq n < l$.

Say that a semigroup variety V is closed under extended standard wreath products if and only if $A, B \in V$ implies $AWr^1B \in V$.

Theorem 4 [7]. *A semigroup variety is closed under extended standard wreath product if and only if it is defined by one of the identities of* (1).

Corollary. *A variety $V \in \mathcal{M}$ is an idempotent of $\mathcal{M}$ if and only if $V = V(s_{ij})$, where s_{ij} is an identity from* (1).

Denote by $\mathcal{A}$ the set of all supercommutative semigroup varieties (i.e. semigroup varieties containing a variety $V = V(xy = yx)$). Using results from [6], one can prove the following

Theorem 5. *$\mathcal{A}$ is a prime ideal of $\mathcal{M}$, and the multiplication in $\mathcal{A}$ is a zero one.*

References

[1] Neumann, B. H. and P. M. Neumann, *Wreath products and varieties of groups*, Math. Z. **80** (1962), 44–62.

[2] Neumann, H., *Varieties of groups*, Springer-Verlag, 1967.

[3] Malcev, A. I., *On the multiplication of classes of algebraic systems*, Sib. Math. J. **8** (1967), 346–465.

[4] Malcev, A. I., *Algebraic systems*, Moscow, 1970.

[5] Eilenberg, S., *Automata languages and machines*, Vol. B., Academic Press, London–New York, 1967.

[6] Košelev, Yu. G., *On one associative operation on a set of all monoid varieties*, Sovrem. Alg. **4** (1975), 107–117 (Russian).

[7] Košelev, Yu. G., *Varieties preserved under wreath products*, Semigroup Forum **12** (1976), 95–107.

[8] Košelev, Yu. G., *A remark on the regularity of semigroups*, Semigroups with additional structures, Leningrad, Len. Gos. Ped. Inst. im. Gerz. (1986), 28–31 (Russian).

[9] Košelev, Yu. G., *An associative multiplication of semigroup varieties*, Inf. Conf. A. I. Malcev, Abstracts, Novosibirsk (1978), **63** (Russian).

Novosibirsk Pedagogical Institute
Department of Mathematics
630126, Novosibirsk, Viljuiskaja, 28

Received September 20, 1989
and in final form November 20, 1989

Semigroup Forum Vol. 45 (1992) 9–22

RESEARCH ARTICLE

Injective Endomorphisms of $\mathcal{G}_X$-normal Semigroups: Infinite Defects

Inessa Levi

Communicated by G. J. Lallement

Let X be an infinite set and $\mathcal{G}_X$ be the symmetric group on X. A semigroup S of transformations of X is said to be $\mathcal{G}_X$-*normal* if for every $h \in \mathcal{G}_X$, $hSh^{-1} \subseteq S$. Amongst examples of $\mathcal{G}_X$-normal semigroups is the semigroup of all transformations of X, the semigroup of all one-to-one transformations of X, and the Baer-Levi semigroup [2]. It was shown in [3] and [4] that all automorphisms of a $\mathcal{G}_X$-normal semigroup are *inner*, that is, of the form $f \to hfh^{-1}$, where $f \in S$ and h is a fixed permutation of X. For a mapping $f : X \to X$, the *defect* of f is def $f = |X - R(f)|$, where $R(f) = f(X)$ is the range of f. In [7] we gave a description of all injective endomorphisms of $\mathcal{G}_X$-normal semigroups of one-to-one transformations with *finite* defects. The present paper presents a complete description of injective endomorphisms of a $\mathcal{G}_X$-normal semigroup S of one-to-one transformations with *infinite* defects smaller than $|X|$. Each injective endomorphism of S determines a family $\{h_i\}$ of one-to-one functions from X into a subset W of X, $|W| = |X|$, and a group homomorphism ξ of S, and every such pair, suitably linked, determines an injective endomorphism of S (Theorem 1.1). This gives a complete description of injective endomorphisms of S since the group homomorphisms of S are described in [9]. The main ingredients involved in a description of injective endomorphisms in [7] and the present paper are similar: a set of one-to-one mappings $\{h_i\}_{i\in I}$ and a homomorphism ξ (which is a group homomorphism in the present paper and may or may not be in [7]). However, the way in which $\{h_i\}_{i\in I}$ and ξ are linked depends a lot on whether S consists of transformations with finite or infinite defects. Moreover, the techniques that are used in describing $\{h_i\}_{i\in I}$ and ξ in [7] and in the present work are completely different: in both cases the difference sets $D(s,t) = \{x \in X : s(x) \neq t(x)\}, s,t \in S$, are involved, but in [7] the images $\varphi(s)$ and $\varphi(t)$, under injective endomorphism φ of an s and t that differ on a finite number of points, also differ on a finite number of points. This is not the case in the present paper. As an example of our results, we present a description of all injective endomorphisms of a Baer-Levi semigroup $BL(p,q)$, $p > q$. For more results on $\mathcal{G}_X$-normal semigroups the reader may see [5], [6], [8], [9], while [10] and [14] present results on injective endomorphisms.

1. The Main Results

In the sequel S is a $\mathcal{G}_X$-normal semigroup of one-to-one transformations with infinite defects smaller than $|X|$. Let σ-def S = $\{\text{def}\, f : f \in S\}$ be the

This research was partially supported by a grant from the University of Louisville.

spectrum of defects of S, and

$$\eta = \min\{\alpha : \alpha > \operatorname{def} f, \text{ for all } f \in S\} .$$

For a transformation f of X let $S(f) = \{x \in X : f(x) \neq x\}$ and the *shift* of f, shift $f = |S(f)|$. Let σ-shift $S = \{\text{shift } f : f \in S\}$ be the *spectrum of shifts* of S, and

$$\mu = \min\{\alpha : \alpha > \text{shift} f, f \in S\} .$$

For every $\alpha \in \sigma$-def S, let $S(\alpha) = \{f \in S : \operatorname{def} f = \alpha\}$ and

$$\mu(\alpha) = \min\{\beta : \beta > \text{shift} f, f \in S(\alpha)\} .$$

Given $f, g \in S$, let $D(f,g) = \{x \in X : f(x) \neq g(x)\}$ be the difference set of f and g. If α is an infinite cardinal, let Δ_α be the congruence on S such that $(f,g) \in \Delta_\alpha$ if $|D(f,g)| < \alpha$. Some properties of Δ_α are described in [9] where it is shown specifically that θ is a group congruence on S if and only if $\theta = \Delta_\alpha$ with $\alpha \geq \eta$.

Set $S/\Delta_\eta = \{V_\beta : \beta \in \Omega\}$, where Ω is an index set and let σ-def $V_\beta = \{\operatorname{def} f : f \in V_\beta\}$, σ-shift $V_\beta = \{\text{shift } f : f \in V_\beta\}$. In the sequel we view Ω as a semigroup isomorphic to S/Δ_η.

Let $\mathcal{I}_X$ be the semigroup of all total one-to-one transformations of X and

$$S(X, \alpha, \beta) = \{f \in \mathcal{I}_X : \text{shift} f \leq \alpha, \operatorname{def} f = \beta\} .$$

If $\alpha \geq \beta \geq \aleph_0$, then $S(X, \alpha, \beta)$ is a $\mathcal{G}_X$-normal semigroup, and it is shown in [8] that

$$S = \bigcup\{S(X,\alpha,\beta) : \alpha \in \sigma\text{-shift } S, \beta \in \sigma\text{-def } S\} \quad \text{and} \tag{1}$$
$$S \supseteq S(X,\beta,\beta) \quad \text{for every} \quad \beta \in \sigma\text{-def } S .$$

Theorem 1.1. *Let φ be an injective endomorphism of S. There exist*

(i) *a subset W of X with $|W| = |X|$ and a partition $\{X_i : i \in I\}$ of W such that $|X_i| = |X|, |I| \in \sigma$-def S or $|I| \leq \nu = \min \sigma$-def S,*

(ii) *a homomorphism $\tau : \Omega \to \mathcal{G}_I$ such that for every $\beta \notin \ker \tau, \mu(\gamma|I|) = |X|^+$, for each $\gamma \in \sigma$-def V_β,*

(iii) *a set of bijections $h_i : X \to X_i$, $i \in I$, and*

(iv) *a group homomorphism $\xi : S \to G_\mu = \{h \in \mathcal{G}_{X-W} : \text{shift } h < \mu\}$ with the associated congruence Δ_α on S, $\alpha \geq \eta$, satisfying shift $\xi(g) < \mu(\gamma|I|)$, $g \in V$, a Δ_α-class, $\gamma \in \sigma$-def V, such that given $f \in V_\beta$, $x \in X$,*

$$\varphi(f)(x) = \begin{cases} h_{\tau(\beta)(i)} f h_i^{-1}(x), & \text{if} \quad x \in X_i , \\ \xi(f)(x), & \text{if} \quad x \in X - W . \end{cases} \tag{2}$$

Conversely, given (i) - (iv) *the mapping defined in* (2) *is an injective endomorphism of S.*

A proof of the above theorem is given in Section 2.

Remark 1.2. If $|X| \notin \sigma$-shift S, that is, shift $f < |X|$, for every $f \in S$, then $\ker \tau = \Omega$. Therefore, given an injective endomorphism φ of such an S, there exist (i), (iii), and (iv) as in Theorem 1.1 such that for an $f \in V_\beta$, $x \in X$,

$$\varphi(f)(x) = \begin{cases} h_i f h_i^{-1}(x), & \text{if} \quad x \in X_i , \\ \xi(f)(x), & \text{if} \quad x \in X - W . \end{cases} \tag{3}$$

Conversely, given (i), (iii), and (iv), the mapping φ defined in (3) is an injective endomorphism of S.

Remark 1.3. If $|X| \in \sigma\text{-shift}\, S$, then there exists the smallest $\beta \in \sigma\text{-def}\, S$ such that $S(X, |X|, \beta) \subseteq S$. Hence for every $\gamma \in \sigma\text{-def}\, S$, $\gamma > \beta$, we have that $S(X, |X|, \gamma) \subseteq S$. Therefore in a construction of an injective endomorphism the restriction on τ described in (ii) of Theorem 1.1 can be satisfied by choosing $|I| \geq \beta$.

Next we consider the restriction on the choice of a group homomorphism ξ as stated in Theorem 1.1 (iv).

Proposition 1.4. i) *Let* $\nu = \min \sigma\text{-def S}$. *If* $\sigma\text{-shift}\, S(\nu) = \sigma\text{-shift}\, S$, *then any group homomorphism* $\xi : S \to G_\mu$ *satisfies the condition* (iv) *of Theorem 1.1.* ii) *Assume* $|X|$ *is regular. If* $|X| \in \sigma\text{-shift}\, S$ *and* $|X - W| < |X|$, *then* ξ *is trivial. If* $|X - W|^+ < |X|$, *then* ξ *is trivial.*

While the proof of i) follows easily from (2), the proof of ii) is non-trivial and will be given in Section 2.

An important example of a $\mathcal{G}_X$-normal semigroup of one-to-one transformations with infinite defects is a Baer-Levi semigroup $BL(p,q)$, where $p = |X|$, $\aleph_0 \leq q \leq p = |X|$. $BL(p,q)$ consists of all one-to-one transformations of X with defect q. Below we present an example of applying Theorem 1.1 to obtain a description of injective endomorphisms of $BL(p,q)$, $p > q \geq \aleph_0$.

Proposition 1.5. *Let* φ *be an injective endomorphism of* $BL(p,q)$, $p > q$. *There exist*

i) *a subset* W *of* X *with* $|W| = |X|$ *together with a partition* $\{X_i : i \in I\}$ *of* W *such that* $|X_i| = p$ *and* $|I| \leq q$;

ii) *a homomorphism* $\tau : \Omega \to \mathcal{G}_I$ *(where* $\Omega \cong S/\Delta_{q^+}$*)*;

iii) *a set of bijections* $h_i : X \to X_i$, $i \in I$, *and*

iv) *a homomorphism* $\xi : S \to \mathcal{G}_{X-W}$ *such that* $\ker \xi = \Delta_\alpha$, *for some* $\alpha \geq q^+$, *and for every* $f \in V_\beta$, $x \in X$,

$$\varphi(f)(x) = \begin{cases} h_{\tau(\beta)(i)} f h_i^{-1}(x), & \text{if } x \in X_i\,, \\ \xi(f)(x), & \text{if } x \in X - W\,. \end{cases}$$

Conversely, given i)–iv), *the mapping defined above is an injective endomorphism of* $BL(p,q)$.

The next example of a $\mathcal{G}_X$-normal semigroup with a structure richer than that of a Baer-Levi semigroup will be used to illustrate the critical ideas of the proof.

Example 1.6. Let $\alpha_1 < \alpha_2 < \alpha_3$ be infinite cardinals, X be a set of cardinality α_3, and

$$\begin{aligned} T &= S(X, \alpha_1, \alpha_1) \cup S(X, \alpha_3, \alpha_2) \\ &= \{f \in \mathcal{I}_X :\ \text{either def } f = \alpha_1 \leq \text{shift } f \text{ or def } f = \alpha_2 \leq \text{shift } f \leq \alpha_3\}\,. \end{aligned}$$

Note that $\nabla\text{-def}\, T = \{\alpha_1, \alpha_2\}$ and $\eta = \min\{\alpha : \alpha > \mathrm{def\,f}, \mathrm{t} \in \mathrm{T}\} = \alpha_2^+$, the cardinal successor of α_2. Also $\mu = \min\{\alpha > \mathrm{shift\,f}, \mathrm{f} \in \mathrm{T}\} = \alpha_3^+$ and $\gamma = \min \nabla\text{-def}\,\mathrm{T} = \alpha_1$.

2. Proofs

In the sequel let φ be an injective endomorphism of a $\mathcal{G}_X$-normal semigroup S of one-to-one transformations f of X with $\aleph_0 \leq \operatorname{def} f < |X|$. We outline the basic ideas involved in our description of φ. With the endomorphism φ, we associate a partition of X into sets W and U with W being the union of all sets $D(\varphi(t),\varphi(s))$ such that $t,\ s \in S$ with $|D(t,s)| = 1$. For an $f \in S$, we obtain the description of $\varphi(f)$ via considering $\varphi(f)|_W$ (results numbered 2.1–2.22) and $\varphi(f)|_U$ (results numbered 2.23, 2.24) separately. The description of $\varphi(f)|_W$ is given in terms of a family of suitably interlocked one-to-one transformations $\{h_i\}_{i\in I}$ from X into W. The determination of these transformations constitutes the largest part of the proofs and is developed in 2.1–2.21. We give now an informal account of the main ideas involved in the determination of $\{h_i\}_{i\in I}$. Through a sequence of Lemmas (2.1–2.7) we show that W is in fact a disjoint union of the sets $D(\varphi(t),\varphi(s))$, with $|D(t,s)| = 1$, $t,s \in S$. Moreover, we establish the existence of an order-monomorphism M between difference sets $D = D(t,s)$ of transformations in S such that given $D \subseteq X$ with $|D| < \eta = \min\{\alpha : \alpha > \operatorname{def} f, \text{for all } f \in S\}$, M maps D onto M_D, where M_D is the difference set $D(\varphi(t),\varphi(s))$ of some mappings $t,s \in S$ such that $D(t,s) = D$ (Lemma 2.10). The sets M_D play a very important role in defining $\{h_i\}_{i\in I}$. We show that if f and g differ on fewer than η points, then the difference set of $\varphi(f)$ and $\varphi(g)$ is the image of $D(f,g)$ under M, $M_{D(f,g)}$ (Lemmas 2.13, 2.14, and 2.24). To define $\{h_i\}_{i\in I}$, we associate with φ a partition $\{X_i\}_{i\in I}$ of W orthogonal to the partition $\{M_x\}_{x\in X}$ (Lemmas 2.16–2.19) and define $h_i(y) = M_y \cap X_i$. Proposition 2.22 asserts the description of $\varphi(f)|_W$ presented in Theorem 1.1. Now, the set of all restrictions of the form $\varphi(f)|_U$, $U = X - W$, $f \in S$, forms a subgroup $G_\mu = \{h \in \mathcal{G}_U : \operatorname{shift} h < \mu\}$ of the symmetric group on U (Lemmas 2.23, 2.24) and $\varphi(f)|_U$ is determined by the congruence class Δ_α of S (Theorem 1.1 (iv)).

We start our description of φ by associating with it a certain lattice monomorphism between subsets of $\mathcal{P}_X$, the power set of X. Let $f,g \in S$. Recall that $\eta = \min\{\alpha : \alpha > \operatorname{def} f, \text{for all } f \in S\}$ and observe that in view of (1), $(f,g) \in \Delta_\eta$, the minimal group congruence on S [9], if and only if there exists a $t \in S$ such that $ft = gt$ (indeed, choose a $t \in S$ with $X - R(t) \supseteq D(f,g)$). Let $\mathcal{D}$ be the set of all difference sets $D(f,g)$ of all $(f,g) \in \Delta_\eta$, so that

$$\mathcal{D} = \{D \subseteq X : |D| < \eta\}\ .$$

In the following three results we show that φ preserves difference sets of certain pairs of transformations in Δ_η (Proposition 2.3). That will enable us to define an order-monomorphism M from $\mathcal{D}$ into $\mathcal{P}_X$. We start with the following technical lemma.

Lemma 2.1. *Let $(f,p) \in \Delta_\eta$ with $D(f,p) = T$. There exist $s,t \in S$ such that $sf = tp$ and $D(s,t) = f(T) \cup p(T)$.*

Proof. Let $f(T) = A$, $p(T) = B$ and $C = X - (A \cup B)$. Since $|A| = |B| = |T| < \eta \le |X|$, we have that $|C| = |X|$. Let $\alpha = \max\{\operatorname{def} f, \operatorname{def} p, |T|\}$, and $\beta \in \sigma\text{-def}\, S$ such that $\beta \ge \alpha$. Choose a subset E of C such that $|E| = |X|$ and $|C - E| = \beta$, together with a bijection k from C onto E such that $\operatorname{shift} k = \beta$. Without loss of generality assume that $|A - B| \ge |B - A|$. Let $u : A \to B$ be such that $uf(y) = p(y)$, for every $y \in T$, $v : B - A \to A - B$ be a one-to-one mapping. Define

$$s(x) = \begin{cases} k(x), & \text{if } \ x \in C\ , \\ u(x), & \text{if } \ x \in A\ , \\ v(x), & \text{if } \ x \in B - A\ , \end{cases}$$

$$t(x) = \begin{cases} k(x), & \text{if } \ x \in C\ , \\ x, & \text{if } \ x \in A \cup B\ . \end{cases}$$

Then $s,t \in S(X,\beta,\beta) \subseteq S$ (see (1)), $sf = tp$, and since $A \cup B \subseteq S(s)$, $D(s,t) = A \cup B$. ■

The proof of the following lemma is straightforward.

Lemma 2.2. *Let $f,g,t \in S$. Then $D(tf,tg) = D(f,g)$ and $t(D(ft,gt)) = D(f,g) \cap R(t)$.*

Proposition 2.3. *Let $(f,g),(p,q) \in \Delta_\eta$ such that $D(f,g) = D = D(p,q)$ and $f(D) \cap g(D) = \emptyset = p(D) \cap q(D)$. Then $D(\varphi(f),\varphi(g)) = D(\varphi(p),\varphi(q))$.*

Proof. Consider the following two cases.

Case 1: $g = q$.

Let $D(f,p) = T$. Observe that $T \subseteq D$ and choose s and t as in Lemma 2.1, that is, $sf = tp$ and $D(s,t) = f(T) \cup p(T)$. Since $f(D) \cap g(D) = \emptyset = p(D) \cap q(D)$, we have that $D(s,t) \subseteq f(D) \cup p(D) \subseteq X - R(g)$, so that $sg = tg$. Then by Lemma 2.2, $D(\varphi(f),\varphi(g)) = D(\varphi(s)\varphi(f),\varphi(s)\varphi(g)) = D(\varphi(t)\varphi(p),\varphi(t)\varphi(g)) = D(\varphi(p),\varphi(g))$.

Case 2: $g \neq q$.

Using (1) choose $k,\ell \in S$ such that $kf = \ell p$. Then by Lemma 2.2, $D(\varphi(f),\varphi(g)) = D(\varphi(k)\varphi(f),\varphi(k)\varphi(g)) = D(\varphi(\ell)\varphi(p),\varphi(k)\varphi(g))$. Also, $D(\varphi(p),\varphi(q)) = D(\varphi(\ell)\varphi(p),\varphi(\ell)\varphi(q))$. Moreover, $D(\ell p,kg) = D(kf,kg) = D(f,g) = D$, $D(\ell p,\ell q) = D(p,q) = D$, and $\ell p(D) \cap kg(D) = kf(D) \cap kg(D) = k(f(D) \cap g(D)) = \emptyset$, $\ell p(D) \cap \ell q(D) = \ell(p(D) \cap q(D)) = \emptyset$. Thus by Case 1, $D(\varphi(f),\varphi(g)) = D(\varphi(\ell)\varphi(p),\varphi(k)\varphi(g)) = D(\varphi(\ell)\varphi(p),\varphi(\ell)\varphi(q)) = D(\varphi(p),\varphi(q))$. ■

In view of Proposition 2.3, we can define now a map $M : \mathcal{D} \to \mathcal{P}_X$ such that for $D \in \mathcal{D}$, $M : D \to M_D$ where $M_D = D(\varphi(f),\varphi(g))$ such that $D(f,g) = D$ with $f(D) \cap g(D) = \emptyset$. Notice that if $D = \{x\} = D(f,g)$, then $\{f(x)\} \cap \{g(x)\} = \emptyset$ so that $D(\varphi(f),\varphi(g)) = M_{\{x\}}$. In the sequel we write M_x for $M_{\{x\}}$. Observe that $\mathcal{D}$ is a distributive lattice (with respect to set-theoretic unions and intersections). Moreover, $\mathcal{D}$ is *relatively complemented* (that is, every $D \in \mathcal{D}$ has a complement in every interval in $\mathcal{D}$ which contains D). Therefore, $\mathcal{D}$ is a *generalized Boolean algebra.* Our aim now is to show that M is an order-monomorphism. We start with the following result.

Lemma 2.4. *Let $D \in \mathcal{D}$, $f \in S$.*

i) $\varphi(f)(M_D) = M_{f(D)} \cap R(\varphi(f))$,

ii) $R(\varphi(f)) \cap M_D \neq \emptyset$ *implies* $R(f) \cap D \neq \emptyset$,

iii) *if* $D = X - R(f)$, *then* $M_D \subseteq X - R(\varphi(f))$.

Proof. i) Choose $s,t \in S$ such that $D(s,t) = f(D)$ and $s(f(D)) \cap t(f(D)) = \emptyset$. Then $D(\varphi(s),\varphi(g)) = M_{f(D)}$. Also $D(sf,tf) = D$ and $D(\varphi(s)\varphi(f),\varphi(t)\varphi(f)) = M_D$. Therefore, $\varphi(f)(M_D) = M_{f(D)} \cap R(\varphi(f))$ (Lemma 2.2).

ii) Choose $p,q \in S$ such that $D(p,q) = D$, $p(D) \cap q(D) = \emptyset$. Then $D(\varphi(p),\varphi(q)) = M_D$, so $\varphi(p)\varphi(f) \neq \varphi(q)\varphi(f)$, hence $pf \neq qf$ and so $\emptyset \neq R(f) \cap D(p,q) = R(f) \cap D$.

iii) Choose $p,q \in S$ as in ii). Then $pf = qf$, so $D(\varphi(p)\varphi(f),\varphi(q)\varphi(f)) = \emptyset$, or $M_D \cap R(\varphi(f)) = \emptyset$. ■

The next two lemmas are used to prove that M maps disjoint sets to disjoint sets (Lemma 2.7). For $0 \leq \alpha, \beta \leq |X|$ such that $\alpha + \beta = |X|$, let

$$\mathcal{A}_{\alpha,\beta} = \{A \subseteq X : |A| = \alpha, |X - A| = \beta\} .$$

For a cardinal α, let α^+ denote the cardinal successor of α. In the sequel we assume the generalized continuum hypothesis.

Lemma 2.5. *Let $|X| = \gamma$. Then $|\mathcal{A}_{\gamma,\gamma}| = \gamma^+$.*

Proof. Observe that for every $\alpha \leq \gamma$, $|\bigcup\{\mathcal{A}_{\alpha,\beta} : \alpha + \beta = \gamma\}| = \gamma^\alpha$ ([11]). Moreover, for $\alpha < \gamma$, $\gamma^\alpha = \gamma$ only if $\alpha < cf(\gamma)$ and $\gamma^\alpha = \gamma^+$ otherwise. Therefore, if γ is regular, then $\gamma^\alpha = \gamma$ and so $|\bigcup\{\mathcal{A}_{\alpha,\gamma} : \alpha < \gamma\}| = |\{\mathcal{A}_{\gamma,\beta} : \beta < \gamma\}| = \gamma$. Since $\mathcal{A}_{\gamma,\gamma} = \mathcal{P}_X - \bigcup\{\mathcal{A}_{\alpha,\gamma} : \alpha < \gamma\} - \bigcup\{\mathcal{A}_{\gamma,\beta} : \beta < \gamma\}$ and the power set $\mathcal{P}_X$ of X consists of γ^+ elements, $|\mathcal{A}_{\gamma,\gamma}| = \gamma^+$.

Now assume that γ is singular and $\alpha < \gamma$ is such that $cf(\alpha) = cf(\gamma)$, so that $\mathcal{A}_{\alpha,\gamma} = \gamma^+$. Partition X into sets X_1 and X_2 such that $|X_1| = |X_2| = \gamma$. Let $\mathcal{B}_{\alpha,\gamma} = \{B \subseteq X_1 : |B| = \alpha\}$. By the previously stated observation, $|\mathcal{B}_{\alpha,\gamma}| = \gamma^+$. Now, $\mathcal{A}_{\gamma,\gamma} \supset \{B \cup X_2 : B \in \mathcal{B}_{\alpha,\gamma}\}$, so that $|\mathcal{A}_{\gamma,\gamma}| \geq |\mathcal{B}_{\alpha,\gamma}| = \gamma^+$, hence $|\mathcal{A}_{\gamma,\gamma}| = \gamma^+$. ∎

Lemma 2.6. *Let Y be an infinite set and $\mathcal{B}$ be a collection of distinct subsets of Y with $|\mathcal{B}| = \gamma$, where $\gamma = |Y|^+$. There exist γ^+ subcollections $\mathcal{C}$ of $\mathcal{B}$ such that $|\mathcal{C}| = \gamma$ and $\bigcap \mathcal{C} \neq \emptyset$.*

Proof. Given $a \in Y$, let $\mathcal{K}(a) = \{B \in \mathcal{B} : a \in B\}$. Then $\mathcal{B} = \bigcup\{\mathcal{K}(a) : a \in \bigcup \mathcal{B}\}$. Since $|\mathcal{B}| = \gamma$ is a cardinal successor, it is regular and so there exists $\mathcal{K}(a)$ with $|\mathcal{K}(a)| = \gamma$. Now for any subset $\mathcal{C}$ of $\mathcal{K}(a)$, $\bigcap \mathcal{C} \ni a$, so $\bigcap \mathcal{C} \neq \emptyset$. There are γ^+ subsets $\mathcal{C}$ of $\mathcal{K}(a)$ of cardinality γ (Lemma 2.5), as required. ∎

Lemma 2.7. *Let D, $T \in \mathcal{D}$ with $D \cap T = \emptyset$. Then $M_D \cap M_T = \emptyset$.*

Proof. We will show that if $M_D \cap M_T \neq \emptyset$, then $|M_D| = |X|$, a contradiction to Lemma 2.4 iii). Let D, T be as stated. Assume without loss of generality that $|D| \leq |T|$ and let $|T| = \alpha$. We show firstly that for every $K \in \mathcal{A}_{\alpha,|X|}$ with $K \cap D = \emptyset$, we have that $M_D \cap M_K \neq \emptyset$. Indeed, since $D, K \in \mathcal{D}$, there exists an $f \in S$ such that $f(T) = K$, $f(D) = D$. Then by Lemma 2.4 i), $\emptyset \neq \varphi(f)(M_D \cap M_T) \subseteq M_{f(D)} \cap M_{f(T)} = M_D \cap M_K$.

Next we show that if $K \in \mathcal{A}_{\alpha,|X|}$ is such that K, D and T are pairwise disjoint, then $M_D \cap M_K \neq M_D \cap M_T$. Let $\beta \in \sigma\text{-def}\, S$ with $\beta \geq \alpha$ and choose $g \in S(X, \beta, \beta) \subseteq S$ such that $T \subseteq X - R(g)$, $g(K) = K$, $g(D) = D$. If $M_D \cap M_K = M_D \cap M_T$, then $\emptyset \neq \varphi(g)(M_D \cap M_K) \subseteq M_{g(D)} \cap M_{g(K)} = M_D \cap M_K = M_D \cap M_T$, so by Lemma 2.4 ii), $T \cap R(g) \neq \emptyset$, a contradiction.

Now we prove that the above observations imply that $|M_D|^+ = |X|^+$, and so $|M_D| = |X|$ (see [13]). Choose a subset Y of X such that $|Y| = |X| = |X - Y|$, $D \subseteq X - Y$, together with a partition $\mathcal{Y}$ of Y into $|X|$ sets of cardinality α. Let $\mathcal{K} = \{M_D \cap M_K : K \in \mathcal{Y}\}$. Since the sets in $\mathcal{Y}$ and D are pairwise disjoint, $|\mathcal{K}| = |X|$. Let $|M_D| = \beta$, then $\beta^+ = |\mathcal{P}_{M_D}| \geq |\mathcal{K}| = |X|$. Assume $\beta^+ = |X|$. Let $\mathcal{B} = \{\mathcal{C} \subseteq \mathcal{K} : |\mathcal{C}| = |X|, \bigcap \mathcal{C} \neq \emptyset\}$. By Lemma 2.6, $|\mathcal{B}| = |X|^+$. We show that if $\mathcal{C}_1, \mathcal{C}_2 \in \mathcal{B}$ with $\mathcal{C}_1 \neq \mathcal{C}_2$, then $\bigcap \mathcal{C}_1 \neq \bigcap \mathcal{C}_2$. Therefore the function $\lambda : \mathcal{B} \to \mathcal{P}_{M_D}$, given by $\lambda : \mathcal{C} \to \bigcap \mathcal{C}$ is one-to-one, so $\beta^+ \geq |X|^+$ and hence $\beta^+ = |X|^+$. Now assume that $\mathcal{C}_1 \neq \mathcal{C}_2$ and let $C_i = \{K \in \mathcal{Y} : M_K \cap M_D \in \mathcal{C}_i\}$, $i = 1, 2$. Since $\mathcal{C}_1 \neq \mathcal{C}_2$, we have that $C_1 \neq C_2$. Let $L \in C_2 - C_1$ so that $L \cap (\bigcup C_1) = \emptyset$. Since $\bigcup C_i \subseteq Y \subseteq X$, $i = 1, 2$, with

$|X-Y|=|X|$, there exists $t \in S$ such that t is the identity on $(\bigcup C_1) \cup D$ and $L \subseteq X - R(t)$. Assume $\bigcap \mathcal{C}_1 = \bigcap \mathcal{C}_2$. Then

$$\begin{aligned}\varphi(t)(\bigcap \mathcal{C}_1) &= \varphi(t)(\bigcap\{M_D \cap M_K : K \in C_1\}) \subseteq \bigcap\{M_{t(D)} \cap M_{t(K)} : K \in C_1\} \\ &= \bigcap\{M_D \cap M_K : K \in C_1\} = \bigcap \mathcal{C}_1 = \bigcap \mathcal{C}_2 \subseteq M_D \cap M_L ,\end{aligned}$$

so $M_L \cap R(t) \neq \emptyset$, a contradiction. ■

In the next two results we show that given $f \in S$, $D \in \mathcal{D}$, $\varphi(f)(M_D) = M_{f(D)}$ (compare with Lemma 2.4 i)).

Lemma 2.8. *Assume there exists an $f \in S$ and a $D \in \mathcal{D}$ such that $\varphi(f)(M_D) = M_{f(D)}$. Then for every $g \in S$ with* $\operatorname{def} g \leq \operatorname{def} f$, *every $K \in \mathcal{D}$ with $|K| = |D|$ we have that $\varphi(g)(M_K) = M_{g(K)}$.*

Proof. Since $\operatorname{def} g \leq \operatorname{def} f$ and because of (1), there exist $s,t \in S$ such that $f = tgs$ and $s(D) = K$. Then $M_{f(D)} = \varphi(f)(M_D) = \varphi(t)\varphi(g)\varphi(s)(M_D) \subseteq \varphi(t)\varphi(g)(M_{s(D)}) = \varphi(t)\varphi(g)(M_K) \subseteq \varphi(t)(M_{g(K)}) \subseteq M_{tg(K)} = M_{tgs(D)} = M_{f(D)}$ (by Lemma 2.4,i). Therefore, $\varphi(t)\varphi(g)(M_K) = \varphi(t)(M_{g(K)})$, so $\varphi(g)(M_K) = M_{g(K)}$, since $\varphi(t)$ is one-to-one. ■

Lemma 2.9. *Let $f \in S$. For $D \in \mathcal{D}$, $\varphi(f)(M_D) = M_{f(D)}$.*

Proof. Let $\{D_i : i \in I\}$ be a partition of X into $|X|$ sets of cardinality $|D|$ each. Then $|X| > \operatorname{def} \varphi(f) \geq |\bigcup\{M_{f(D_i)} - \varphi(f)(M_{D_i}) : i \in I\}|$ (by Lemma 2.7). Thus there exists D_i such that $M_{f(D_i)} = \varphi(f)(M_{D_i})$. The result now follows from Lemma 2.8. ■

Lemma 2.10. *The mapping $M : D \to M_D$ is an order-monomorphism from $\mathcal{D}$ into $\mathcal{P}_X$.*

Proof. Let $D, T \in \mathcal{D}$. We show that $D \subseteq T$ if and only if $M_D \subseteq M_T$. Assume $D \subseteq T$ and choose $f,s,t \in S$ with $D(s,t) = T$, $s(T) \cap t(T) = \emptyset$ and $D(sf,tf) = D$ with $f(D) = D$. Then $D(\varphi(s),\varphi(t)) = M_T$, $D(\varphi(s)\varphi(f),\varphi(t)\varphi(f)) = M_D$, so by Lemmas 2.2 and 2.9, $M_T \cap R(\varphi(f)) = \varphi(f)(M_D) = M_{f(D)} = M_D$, and $M_D \subseteq M_T$. Now assume $M_D \subseteq M_T$ while $x \in D - T$. By the previous argument $M_x \subseteq M_D$, while by Lemma 2.7, $M_x \cap M_T = \emptyset$. Thus $M_x \subseteq M_D - M_T$, a contradiction. Finally, we show that M is one-to-one. If $D \neq T$ with $x \in D - T$, we have the $M_x \subseteq M_D - M_T$, as above, so that $M_D \neq M_T$. ■

Lemma 2.11. *Let $f \in S$ be such that $S(f) \in \mathcal{D}$. Then $S(\varphi(f)) \subseteq M_{S(f)}$.*

Proof. Let $S(f) = A$ and $B \subseteq X - A$ such that $|B| = |A| < |X|$. Choose a bijection $p : A \to B$ moving $|A|$ points. Let $\beta \in \sigma\text{-def}\, S$ such that $|A| \leq \beta$ and $C \subseteq X - A - B$ with $|C| = \beta$. Choose a bijection $q : X - A \to X - A - B - C$ with $\operatorname{shift} q = \beta$ and define s,t such that for $x \in X$,

$$s(x) = \begin{cases} q(x), & \text{if } x \in X - A , \\ x, & \text{if } x \in A , \end{cases} \qquad t(x) = \begin{cases} q(x), & \text{if } x \in X - A , \\ p(x), & \text{if } x \in A . \end{cases}$$

Then $s,t \in S(X,\beta,\beta) \subseteq S$, $D(s,t) = A$ and $s(A) \cap t(A) = A \cap B = \emptyset$. Also, $D(s,tf) = A$ and $s(A) \cap tf(A) \subseteq s(A) \cap t(A) = \emptyset$. Thus $D(\varphi(s),\varphi(t)) = M_A = D(\varphi(s),\varphi(t)\varphi(f))$. If $x \in X - M_A$, then $\varphi(t)\varphi(f)(x) = \varphi(s)(x) = \varphi(t)(x)$, so $\varphi(f)(x) = x$ and $S(\varphi(f)) \subseteq M_A$. ■

Our next lemma demonstrates that M preserves unions and intersections of the sets in $\mathcal{D}$.

Lemma 2.12. *M is a lattice monomorphism.*

Proof. We show that for $D, T \in \mathcal{D}$, $M_{D\cup T} = M_D \cup M_T$, $M_{D\cap T} = M_D \cap M_T$. Observe that since $D, T \subseteq D \cup T$ and M is an order-monomorphism we have that M_D, $M_T \subseteq M_{D\cup T}$, so that $M_D \cup M_T \subseteq M_{D\cup T}$. To verify the reverse inclusion, we assume firstly that $D \cap T = \emptyset$. Let $\beta \in \sigma\text{-def}\, S$ such that $|D|$, $|T| \leq \beta$. Recall that $\beta < |X|$ and choose disjoint sets $A, B \subseteq X - D - T$ such that $|A| = |D|$ and $|B| = \beta$. Choose bijections $u : D \to A$ and $v : A \cup B \to B_1$ where $B_1 \subseteq B$ with $|B_1| = |B - B_1| = \beta$. Let $f, g \in S(X, \beta, \beta) \subseteq S$ be as follows. For $x \in X$,

$$f(x) = \begin{cases} v(x), & \text{if } \ x \in A \cup B\ , \\ x, & \text{otherwise}\ , \end{cases}$$

$$g(x) = \begin{cases} v(x), & \text{if } \ x \in A \cup B\ , \\ u(x), & \text{if } \ x \in D\ , \\ x, & \text{otherwise}\ . \end{cases}$$

Then $D(f, g) = D$, $f(D) \cap g(D) = \emptyset$ and $S(f)$, $S(g) \subseteq X - T$. Similarly, choose $s, t \in S(X, \beta, \beta) \subseteq S$ such that $D(s, t) = T$, $s(T) \cap t(T) = \emptyset$ and $s(T)$, $t(T)$, $S(s)$, $S(t) \subseteq X - D - A - B$. Observe that $D(sf, tg) = D \cup T$. Indeed, if $d \in D$, then $sf(d) = s(d) = t(d) \neq tg(d)$, since $g(d) \in A \subseteq X - D$. If $d \in T$, then $f(d) = g(d) = d$, so $sf(d) = s(d) \neq t(d) = tg(d)$. Also, if $d \in X - (D \cup T)$, then $f(d) = g(d)$ and either $f(d) \notin T$, or $d = f(d) \in T$ (since $S(f) \subseteq X - T$), a contradiction to the choice of d. Moreover, $sf(D \cup T) \cap tg(D \cup T) = (D \cup s(T)) \cap (A \cup t(T)) = \emptyset$. Therefore, $D(\varphi(sf), \varphi(tg)) = M_{D\cup T}$.

Now we show that $D(\varphi(s)\varphi(f), \varphi(t)\varphi(g)) \subseteq M_D \cup M_T$. By Lemmas 2.11 and 2.7, $S(\varphi(f)) \cap S(\varphi(s)) \subseteq M_{S(f)} \cap M_{S(s)} = \emptyset$ and similarly, $S(\varphi(f)) \cap S(\varphi(t)) = \emptyset$. Now let $x \notin M_D \cup M_T$, we show that $\varphi(s)\varphi(f)(x) = \varphi(t)\varphi(g)(x)$. Since $x \notin M_D$, we have that $\varphi(f)(x) = \varphi(g)(x) = y$, for some $y \in X$. If $x = y$, then since $x \notin M_T$, we have that $\varphi(s)\varphi(f)(x) = \varphi(s)(x) = \varphi(t)(x) = \varphi(t)\varphi(g)(x)$. If $x \neq y$, then $x, y \in S(\varphi(f))$, so $y \notin S(\varphi(s))$, $S(\varphi(t))$, hence $\varphi(s)\varphi(f)(x) = \varphi(s)(y) = y = \varphi(t)(y) = \varphi(t)\varphi(g)(x)$. Therefore, $X - (M_D \cup M_T) \subseteq X - D(\varphi(s)\varphi(f), \varphi(t)\varphi(g)) = X - M_{D\cup T}$, and the required inclusion follows.

Now assume that $D \cap T \neq \emptyset$. Without loss of generality assume that $T - D \neq \emptyset$. Then $M_{D\cup T} = M_{D \dot\cup (T-D)} = M_D \cup M_{T-D} \subseteq M_D \cup M_T$, by the previous case and Lemma 2.10.

Next we show that M preserves intersections. Observe that $M_{D\cap T} \subseteq M_D \cap M_T$ since M is an order-monomorphism. To show the reverse inclusion write $M_{D\cup T} = M_{D-T} \,\dot\cup\, M_{D\cap T} \,\dot\cup\, M_{T-D}$ (note that some of these sets may be empty). Then $M_{D\cap T} = M_{D\cup T} - M_{D-T} - M_{T-D} = M_D \cup M_T - (M_{D-T} \cup M_{T-D}) = (M_D - M_{D-T}) \cup (M_T - M_{T-D}) \supseteq M_D \cap M_T$, since $M_{D-T} \subseteq X - M_T$, $M_{T-D} \subseteq X - M_D$.

It follow from the above lemma and [1], p. 8, that $M(\mathcal{D})$ is a generalized Boolean algebra.

If a lattice L has a 0, then an element $a \in L$ is an *atom* if $a \neq 0$ but $x \leq a$ implies $x = 0$ or $x = a$, $x \in L$. A distributive lattice L with 0 is atomic if each non-zero element of L is a sum of atoms. Clearly $\mathcal{D}$ is atomic, and we show the following.

Lemma 2.13. *$M(\mathcal{D})$ is atomic.*

Proof. Since $M(\mathcal{D})$ is a generalized Boolean algebra, it suffices to show that for each $M_A \in M(\mathcal{D})$ there exists an atom $\emptyset \neq M_B \in M(\mathcal{D})$ such that $M_B \subseteq M_A$ ([1]), Theorem III.1.5). The last statement follows from an observation that for every $x \in X$, M_x is an atom of $M(\mathcal{D})$. Indeed, assume $M_C \subseteq M_x$, for some $C \in \mathcal{D}$, then by Lemma 2.10, $C \subseteq \{x\}$, so $C = \{x\}$, as required. ■

Let $W = \bigcup\{M_x : x \in X\}$ (see Theorem 1.1(i)). We associate with φ a set of one-to-one transformations from X into W that will describe the behavior of $\varphi(f)$ on W, for an $f \in S$. Recall that $\eta = \min\{\alpha : \alpha > \operatorname{def} f, \text{for all } f \in S\}$.

Lemma 2.14. *If $(f,g) \in \Delta_\eta$ and $A \in \mathcal{D}$ with $A \subseteq X - D(f,g)$, then $\varphi(f)|_{M_A} = \varphi(g)|_{M_A}$.*

Proof. Choose $t \in S$ such that $ft = gt$ and $t(A) = A$. Then $\varphi(t)(M_A) = M_A$ and for an $a \in M_A$, there exists $b \in M_A$ such that $\varphi(t)(b) = a$, so that $\varphi(f)(a) = \varphi(f)\varphi(t)(b) = \varphi(g)\varphi(t)(b) = \varphi(g)(a)$. ■

It was shown in [9] that Δ_η is a group congruence with the kernel $V_0 = \{f \in S : \operatorname{shift} f < \eta\}$, a $\mathcal{G}_X$-normal subsemigroup of S. Let $S/\Delta_\eta = \{V_\beta : \beta \in \Omega\}$, where Ω is an index set. Note that for every $\beta \in \Omega$, V_β is transitive. Now fix an $x \in X$ and for every $\beta \in \Omega$, $u \in M_x$, let

$$Y_{u,\beta} = \{\varphi(f)(u) : f \in V_\beta\} \subseteq W .$$

Lemma 2.15. *If u and v are distinct points in M_x, then $Y_{u,\beta} \cap Y_{v,\beta} = \emptyset$.*

Proof. Suppose $z \in Y_{u,\beta} \cap Y_{v,\beta}$, so that there exist $f, g \in V_\beta$ such that $\varphi(f)(u) = \varphi(g)(v) \in M_{f(x)} \cap M_{g(x)}$ (by Lemma 2.9). Hence by Lemma 2.7, $f(x) = g(x)$ and by Lemma 2.14, $\varphi(f)|_{M_x} = \varphi(g)|_{M_x}$, so that $u = v$. ■

Observe that $\{M_y : y \in X\}$ forms a partition of W (Lemma 2.7). Partitions $\mathcal{A}$ and $\mathcal{B}$ of a set Z are said to be *orthogonal* if for every $A \in \mathcal{A}$, $B \in \mathcal{B}$, $|A \cap B| = 1$.

Lemma 2.16. *For every $\beta \in \Omega$, $\{Y_{u,\beta} : u \in M_x\}$ forms a partition of W orthogonal to $\{M_y : y \in X\}$.*

Proof. To show that $\{Y_{u,\beta} : u \in M_x\}$ forms a partition of W, it suffices to show that $\bigcup\{Y_{u,\beta} : u \in M_x\} = W$ (by Lemma 2.15). Now let $v \in W$ and $w \in X$ be such that $v \in M_w$. Choose $f \in V_\beta$ with $f(x) = w$ (recall that V_β is transitive). Then $\varphi(f)(M_x) = M_w$, so for some $u \in M_x$, $\varphi(f)(u) = v$ and $v \in Y_{u,\beta}$.

Now we show that $\{Y_{u,\beta} : u \in M_x\}$ is orthogonal to $\{M_y : y \in X\}$ by verifying that $|Y_{u,\beta} \cap M_y| = 1$ for every $u \in M_x$ and $y \in X$. Choose $g \in V_\beta$ with $g(x) = y$. Then $\varphi(g)(u) \in Y_{u,\beta} \cap M_y$. If also $a \in Y_{u,\beta} \cap M_y$, then $a = \varphi(t)(u)$ for some $t \in V_\beta$ with $t(x) = y$ and so by Lemma 2.14, $a = \varphi(t)(u) = \varphi(g)(u)$. ■

Next we show that the partition $\{Y_{u,\beta} : u \in M_x\}$ of W does not depend on the choice of the initial point x. Let $z \in X - \{x\}$ and let $Z_{u,\beta} = \{\varphi(f)(u) : f \in V_\beta\}$, for every $u \in M_z$.

Lemma 2.17. *Let $\alpha, \beta \in \Omega$. There exist $\delta, \gamma \in \Omega$ such that for any $f \in V_\alpha$, $k \in V_\delta, e \in V_\gamma$ there exists $g \in V_\beta$ such that $fk = ge$.*

Proof. It suffices to show that there exist δ, $\gamma \in \Omega$ such that $V_\alpha V_\delta = V_\beta V_\gamma$. The latter follows from the fact that $S/\Delta\eta$ is a group. ■

Lemma 2.18. *Given $v \in M_z$, there exists a unique $u \in M_x$ such that $Y_{u,\beta} = Z_{v,\beta}$, for every $\beta \in \Omega$.*

Proof. Let $y \in Z_{v,\beta}$, $y = \varphi(f)(v)$, for some $f \in V_\beta$. Choose $k, g, \ell \in V_\beta$ such that $k(z) = z$, $\ell(z) = x$ and $fk = g\ell$ (Lemma 2.17). There exists $w \in M_z$ such that $\varphi(k)(w) = v$ and so $y = \varphi(f)(v) = \varphi(f)\varphi(k)(w) = \varphi(g)\varphi(\ell)(w) \in Y_{u,\beta}$, where $u = \varphi(\ell)(w) \in M_{\ell(z)} = M_x$. If y' is another element of $Z_{v,\beta}$, $y' = \varphi(f')(v)$, $f' \in V_\beta$, let k and ℓ be the mappings chosen above, while $g' \in V_\beta$ such that $fk = g'\ell$. Then $y' = \varphi(f')(v) = \varphi(f')\varphi(k)(w) = \varphi(g')\varphi(\ell)(w) = \varphi(g')(u) \in Y_{u,\beta}$. ∎

Next we show that if $\alpha, \beta \in \Omega$, then the sets $\{Y_{u,\alpha} : u \in M_x\}$ and $\{Y_{u,\beta} : u \in M_x\}$ are the same.

Lemma 2.19. *Given $v \in M_x$, $\alpha, \beta \in \Omega$, there exists a unique $u \in M_x$ such that $Y_{v,\alpha} = Y_{u,\beta}$.*

Proof. Let $y \in Y_{v,\alpha} \cap M_w$, for some $w \in X$. Choose $f \in V_\alpha$ with $f(x) = w$ and $\varphi(f)(v) = y$ and $k, \ell \in S$ such that $k(x) = x$, $\ell(x) = x$. By Lemma 2.17 there exists $g \in V_\beta$ such that $fk = g\ell$. Then $\varphi(k)(M_x) = M_x$, so $\varphi(k)(a) = v$, for some $a \in M_x$ and $y = \varphi(f)(v) = \varphi(f)\varphi(k)(a) = \varphi(g)\varphi(\ell)(a) = \varphi(g)(u)$, where $u = \varphi(\ell)(a) \in M_{\ell(x)} = M_x$. Thus $y \in Y_{u,\beta}$. If y' is another element of $Y_{v,\alpha}$, choose $f' \in V_\alpha$ so that $y' = \varphi(f')(v)$ and $g' \in V_\beta$ such that $f'k = g'\ell$. As above, $y' = \varphi(g')(u) \in Y_{u,\beta}$. ∎

Our next two results aid us in defining the homomorphism τ (Theorem 1.1 (ii)). Recall that $V_0 = \ker \Delta_\eta = \{f \in S : \operatorname{shift} f < \eta\}$.

Lemma 2.20. *Let $f \in V_0$ with $f(x) = x$. Then $\varphi(f)|_{M_x}$ is the identity transformation of M_x.*

Proof. Since $\eta > \operatorname{shift} f$, there exists $\alpha \in \sigma\text{-def}\, S$ with $\operatorname{shift} f \le \alpha$. Let $x \in X - S(f)$ and choose $t \in S(X, \alpha, \alpha) \subseteq S$ with $S(f) \subseteq X - R(t)$ and $t(x) = x$. Then $ft = t$, $\varphi(f)\varphi(t) = \varphi(t)$, so for every $u \in M_x$, there exists $v \in M_x$ such that $u = \varphi(t)(v)$ and $\varphi(f)(u) = \varphi(f)\varphi(t)(v) = \varphi(t)(v) = u$. ∎

Let $\alpha \in \Omega$ be such that $V_\alpha = V_0$ and define $\lambda : \Omega \to \mathcal{G}_{M_x}$ by $\lambda(\beta) = h$ if $Y_{u,\beta} = Y_{h(u),\alpha}$, for every $u \in M_x$.

Lemma 2.21. *λ is a homomorphism.*

Proof. By Lemma 2.19, given a $\beta \in \Omega$, there exists a function $h : M_x \to M_x$ such that $Y_{u,\beta} = Y_{h(u),\alpha}$ for every $u \in M_x$. Again by Lemma 2.19, there exists a function $p : M_x \to M_x$ such that $Y_{p(u),\beta} = Y_{u,\alpha}$ for every $u \in M_x$. Then $Y_{u,\beta} = Y_{h(u),\alpha} = Y_{ph(u),\beta}$, so $u = ph(u)$. Similarly we show that $u = hp(u)$ and so h and p are inverses of each other, hence $h \in \mathcal{G}_{M_x}$. Therefore, λ is well-defined. To show that λ is a homomorphism, let $\beta, \delta \in \Omega$, $\lambda(\beta) = h$, $\lambda(\delta) = p$ and $\lambda(\beta\delta) = r$. We have that for a $u \in M_x$,

$$\begin{aligned} Y_{r(u),\alpha} = Y_{u,\beta\delta} &= \{\varphi(st)(u) : s \in V_\beta, t \in V_\delta\} \\ &= \{\varphi(s)(Y_{u,\delta}) : s \in V_\beta\} \\ &\supseteq \{\varphi(s)(Y_{u,\delta} \cap M_x) : s \in V_\beta\} \\ &= \{\varphi(s)(Y_{p(u),\alpha} \cap M_x) : s \in V_\beta\} \\ &= \{\varphi(s)(p(u)) : s \in V_\beta\}, \quad \text{by Lemma 2.20} \\ &= Y_{p(u),\beta} = Y_{hp(u),\alpha}. \end{aligned}$$

Therefore $r(u) = hp(u)$ for every $u \in M_x$, so that $\lambda(\beta\delta) = r = hp = \lambda(\beta)\lambda(\delta)$. ∎

Let I be a set with $|I| = |M_x|$ (recall that $|M_x| < \eta$). Choose a one-to-one correspondence $\rho : I \to M_x$ and denote $Y_{u,\alpha} = Y_{\rho(i),\alpha}$ by X_i, $u \in M_x$, $\alpha \in \Omega$ such that $V_\alpha = V_0$. Let $\delta : \mathcal{G}_{M_x} \to \mathcal{G}_I$ be given by $\delta(h) = \rho^{-1} h \rho$. Clearly, δ is an isomorphism and $\tau = \delta\lambda : \Omega \to \mathcal{G}_I$ is a homomorphism such that for a $v \in M_x$ and $\beta \in \Omega$, $Y_{v,\beta} = Y_{\lambda(\beta)(v),\alpha} = X_{\delta\lambda(\beta)(j)} = X_{\tau(\beta)(j)}$, where $j = \rho^{-1}(v)$ (see Theorem 1.1(ii)).

Now we are in a position to define bijections $h_i : X \to X_i$, $i \in I$ (Theorem 1.1(iii)). Let $y \in X$, and $h_i(y) = M_y \cap X_i$. Lemma 2.16 ensures that h_i is a bijection for every $i \in I$. The next result describes the behavior of $\varphi(f)|_W$, $f \in S$.

Proposition 2.22. *Given $f \in V_\beta$, $y \in X_i$, $\varphi(f)(y) = h_{\tau(\beta)(i)} f h_i^{-1}(y)$.*

Proof. Let $y = M_z \cap X_i$, for some $z \in X$, so that $y = h_i(z)$. Now, $\varphi(f)(y) \in \varphi(f)(M_z) = M_{f(z)}$. Choose $k \in V_0$ such that $k(x) = z$. Then $\varphi(k)(M_x) = M_z \ni y$, so $\varphi(k)(v) = y$, for some $v \in M_x$. But then $y \in Y_{v,\alpha} = X_{\rho^{-1}(v)}$, so $v = \rho(i)$ and we have $\varphi(f)(y) = \varphi(f)\varphi(k)(v) = \varphi(fk)(v) \in Y_{v,\beta}$, since $(fk, f) \in \Delta_\eta$. Now $Y_{v,\beta} = X_{\tau(\beta)(\rho^{-1}(v))} = X_{\tau(\beta)(i)}$, so that $\{\varphi(f)(y)\} = M_{f(z)} \cap X_{\tau(\beta)(i)} = \{h_{\tau(\beta)(i)} f(z)\} = \{h_{\tau(\beta)(i)} f h_i^{-1}(y)\}$. ■

Denote $X - W$ by U. Let $\varphi(f)|_U = \widetilde{f}$ and $\widetilde{S} = \{\widetilde{f} : f \in S\}$.

Lemma 2.23. *$\widetilde{S}$ is a semigroup of one-to-one transformations of U.*

Proof. It suffices to show that for every $f \in S$, $\varphi(f)(U) \subseteq U$. Let $y \in U$, $\varphi(f)(y) = z$. If $z \in W$, $z \in M_v$ for some $v \in X$, then $v \in R(f)$, say $f(u) = v$. But this in turn implies that $\varphi(f)(y) \in \varphi(f)(M_u) = M_v \ni z$, so $y \in M_u \subseteq W$, a contradiction. ■

Observe that the mapping $\xi : S \to \widetilde{S}$ given by $\xi : f \to \widetilde{f}$ is a homomorphism (see Theorem 1.1(iv)). Let θ be the congruence on S produced by ξ. Since whenever $(f, g) \in \Delta_\eta$ with $D(f, g) = D$, $D(\varphi(f), \varphi(g)) = M_D = \bigcup\{M_x : x \in D\} \subseteq W$, we have that $\theta \supseteq \Delta_\eta$. Also θ is left cancellative ($\widetilde{S}$ is a semigroup of one-to-one transformations) and it was shown in [9] that a congruence on S is left cancellative if and only if it is equal to Δ_α, for some $\aleph_0 \le \alpha \le |X|^+$. This together with the fact that $\theta \supseteq \Delta_\eta$ implies that $\theta = \Delta_\alpha$ for some $\eta \le \alpha \le |X|^+$. It is proven in [9] that every such congruence on S is a group congruence; moreover, every group congruence on S is of the form Δ_α, for some $\eta \le \alpha \le |X|^+$. Now recall that $\mu = \min\{\alpha : \alpha > \mathrm{shift}\, f, f \in S\}$, $G_\mu = \{h \in \mathcal{G}_U : \mathrm{shift}\, h < \mu\}$. The above observations prove the following result.

Lemma 2.24. *$\xi : S \to G_\mu$ is a group homomorphism with the associated congruence Δ_α, $\alpha \ge \eta$, on S.*

Remark 2.25. Observe that if V is a congruence class of Δ_α, $\alpha \ge \eta$, and V contains a mapping f with $\mathrm{shift}\, f < \alpha$, then V is the kernel of Δ_α and for every $g \in V$, $\mathrm{shift}\, g < \alpha$. If there exists an $f \in V$ with $\mathrm{shift}\, f \ge \alpha$, then for every $g \in V$, $\mathrm{shift}\, g = \mathrm{shift}\, f$ (else assume $\mathrm{shift}\, f < \mathrm{shift}\, g$, then $\alpha > |D(f, g)| \ge \mathrm{shift}\, g$, a contradiction).

Recall that for $\alpha \in \sigma\text{-def}\, S$, $\mu(\alpha) = \min\{\beta : \beta > \mathrm{shift}\, f \text{ for every } f \in S(\alpha)\}$.

Lemma 2.26. *Given $f \in V_\beta \in S/\Delta_\eta$,*
i) $\operatorname{def}\varphi(f) = |I| \operatorname{def} f$,
ii) *if $\beta \in \ker\tau$, then* $\operatorname{shift}\varphi(f) = |I| \operatorname{shift} f + \operatorname{shift}\xi(f)$,
iii) *if $\beta \notin \ker\tau$, then $\mu(\gamma|I|) = |X|^+$, for each $\gamma \in \sigma\text{-}\operatorname{def} V_\beta$.*

Proof. Observe that $\operatorname{def}\varphi(f) = \left|W - R(\varphi(f)|_W)\right| = \left|\dot\bigcup\{X_{\tau(\beta)(i)} - h_{\tau(\beta)(i)} f h_i^{-1}(X_i) : i \in I\}\right| = \sum_{i \in I} \left|h_{\tau(\beta)(i)}(X) - h_{\tau(\beta)(i)} f(X)\right| = |I| \operatorname{def} f$. If $\tau(\beta)(i) \neq i$ for some $i \in I$, then $\varphi(f)(X_i) = h_{\tau(\beta)(i)} f h_i^{-1}(X_i) \subseteq h_{\tau(\beta)(i)}(X) = X_{\tau(\beta)(i)}$, so $X_i \subseteq S(\varphi(f))$ and $\operatorname{shift}\varphi(f) = |X_i| = |X|$. ∎

Corollary 2.27. *If $|X| \notin \sigma\text{-}\operatorname{shift} S$, then for every $f \in S$, $y \in X_i$, $\varphi(f)(y) = h_i f h_i^{-1}(y)$.*

Since $\operatorname{def}\varphi(f) = |I| \operatorname{def} f$, $f \in S$, and $\varphi(f) \in S$, we have the following.

Lemma 2.28. *Given $f \in S$, $\operatorname{shift}\xi(f) < \mu(\gamma|I|)$, $\gamma \in \sigma\text{-}\operatorname{def} V$, $V \in S/\xi = S/\Delta_\alpha$.*

Proof of Theorem 1.1. The set W is defined prior to Lemma 2.14. The sets X_i, $i \in I$, the homomorphism $\tau : \Omega \to \mathcal{G}_I$ and the bijections $h_i : X \to X_i$, $i \in I$, are described prior to Proposition 2.22. The group homomorphism ξ is defined in Lemma 2.24, while (iv) is proved in Proposition 2.22 and Lemmas 2.24 and 2.28. By Lemma 2.26, for an $f \in V_\beta$, $\operatorname{def}\varphi(f) = |I| \operatorname{def} f$, so either $|I| \in \sigma\text{-}\operatorname{def} f$ or $|I| \leq \nu = \min \sigma\text{-}\operatorname{def} S$, as required in (i). Lemma 2.26 ii) proves the statement in (ii).

Conversely, assume (i)–(iv) are as given, and let $f \in S$, $f \in V_\beta$. We show firstly that $\varphi(f) \in S$. Indeed, $\operatorname{def}\varphi(f) = |I| \operatorname{def} f$, and $\operatorname{shift}\varphi(f) = \operatorname{shift}(\varphi(f)|_W) + \operatorname{shift}(\varphi(f)|_U) = \operatorname{shift} h_{\tau(\beta)(i)} f h_i^{-1} + \operatorname{shift}\xi(f) < \mu(|I| \operatorname{def} f)$, as required. Now let $f, g \in S$, $f \in V_\beta$, $g \in V_\gamma$. Then for a $y \in X_i$,

$$\begin{aligned}
\varphi(f)\varphi(g)(y) &= \varphi(f) h_{\tau(\gamma)(i)} g h_i^{-1}(y) \\
&= h_{\tau(\beta)\tau(\gamma)(i)} f h^{-1}_{\tau(\gamma)(i)} h_{\tau(\gamma)(i)} g h_i^{-1}(y) \\
&= h_{\tau(\beta\gamma)(i)} f g h_i^{-1}(y) \\
&= \varphi(fg)(y)\,.
\end{aligned}$$

Since for a $y \in X - W$, $\varphi(fg)(y) = \xi(fg)(y) = \xi(f)\xi(g)(y) = \varphi(f)\varphi(g)(y)$, φ is a homomorphism. Finally, we show that φ is one-to-one. Assume $\varphi(f) = \varphi(g)$, $f \in V_\beta$, $g \in V_\gamma$. Then for a $y \in X_i$, $\varphi(f)(y) = h_{\tau(\beta)(i)} f h_i^{-1}(y)$, $\varphi(g)(y) = h_{\tau(\gamma)(i)} g h_i^{-1}(y)$, so that $\tau(\beta)(i) = \tau(\gamma)(i)$, for every $i \in I$, or $\tau(\beta) = \tau(\gamma)$. But then $f h_i^{-1}(y) = g h_i^{-1}(y)$, so $f = g$ because of the arbitrariness of the choice of y. ∎

The rest of this section constitutes a proof of Proposition 1.4 ii).

Lemma 2.29. *Let Z be an infinite set and $T = S(Z, \alpha, \beta)$. Then*

$$|T| = \begin{cases} |Z|^+, & \textit{if} \quad \alpha \geq cf(|Z|)\,, \\ |Z|, & \textit{otherwise}\,. \end{cases}$$

Proof. Let $\mathcal{A} = \{A \subseteq Z : |A| = \alpha\}$. Observe that for every $A \in \mathcal{A}$, there exists an $f \in T$ such that $A = S(f)$, so that $|T| \geq |\mathcal{A}|$. Moreover, for every A

as above, there are at most $|Z|^\alpha$ ways to choose a one-to-one mapping $f : Z \to Z$ with $A = S(f)$. Therefore $|T| \le |\mathcal{A}||Z|^\alpha$.

If $\alpha \ge cf(|Z|)$, then $|\mathcal{A}| = |Z|^+$, $|Z|^\alpha = |Z|^+$ and so $|T| = |Z|^+$ (see Lemma 2.5). If $\alpha < cf(|Z|)$, then $|\mathcal{A}| = |Z| = |Z|^\alpha$ and so $|T| = |Z|$. ∎

Corollary 2.30. *If $|Z|$ is regular then*

$$|T| = \begin{cases} |Z|^+, & \text{if } \alpha = |Z| , \\ |Z|, & \text{otherwise} . \end{cases}$$

Lemma 2.31. *Let $\gamma \ge \eta$, then $|S(X,\alpha,\beta)/\Delta_\gamma| = |S(X,\alpha,\beta)|$ or 1.*

Proof. Since $\gamma \ge \eta$, Δ_γ is a group congruence. Assume Δ_γ is a non-trivial group congruence. Then $S(X,\alpha,\beta)/\Delta_\gamma \cong S(X,\alpha,0)/\Delta_\gamma$ (see [9]) and $S(X,\alpha,0) \supset S(X,\alpha,0)$ ([13], 11.5.2). The result follows by Lemma 2.29. ∎

Corollary 2.32. *Let $\gamma \ge \eta$. Then $|S/\Delta_\gamma| = |S|$ or 1.*

A proof of Proposition 1.4 ii) follows from 2.29–2.32.

Example 1.6 (revisited). We describe an injective endomorphism φ of T. Note that $\mathcal{D} = \{D \subseteq X : |D| \le \alpha_2\}$, and $S/\Delta_\eta = \{V_\beta : \beta \in \Omega\}$ consists of $V_0 = \ker \Delta_\eta = \{f : \text{shift} f \le \alpha_2\} = S(X,\alpha_1,\alpha_1) \cup S(X,\alpha_2,\alpha_2)$, with the rest of the classes of S/Δ_η subsets of $S(X,\alpha_3,\alpha_2) - S(X,\alpha_2,\alpha_2)$. Therefore, for every $\alpha \ge \eta$, $V \in S/\Delta_\alpha$, $V \ne \ker \Delta_\alpha$, we have that $\sigma\text{-def}\, V = \{\alpha_2\}$.

Let I be an index set and for some $x \in X$, let $\rho : I \to M_x$ be a one-to-one onto correspondence. We construct a partition $\{X_i\}_{i\in I}$ of W such that

$$X_i = \{\varphi(f)(\rho(i)) : f \in V_0\} ,$$

and define $h_i : X \to X_i$ by $h_i(y) = M_y \cap X_{i,i\in I}$. Define a homomorphism $\tau : \Omega \to \mathcal{G}_I$ such that for $\beta \in \Omega$ and $u \in M_x$,

$$X_{\tau(\beta)(i)} = \{\varphi(g)(\rho(i)) : g \in V_\beta\} .$$

Finally, we let ξ be the restriction mapping $\xi : \varphi(f) \to \varphi(f)|_{X-W}$.

To construct a specific example of an injective endomorphism of T, partition $X = W \cup U$ such that $|W| = |U| = |X|$. Let $|I| = \alpha_2$, partition W into α_2 sets X_i of cardinality $|X|$, and choose a set of bijections $h_i : X \to X_i$, $i \in I$. We note that for any homomorphism $\tau : \Omega \to \mathcal{G}_I$, if $\beta \notin \ker\tau$, then $V_\beta \ne V_0$ and so $\sigma\text{-def}\, V_\beta = \{\alpha_2\}$, hence the condition of (ii) in Theorem 1.1 is satisfied. Let σ be the trivial homomorphism. Finally, let $\xi : T \to G_{\alpha_2^+}$ be such that $\xi(T) \cong T/\Delta_{\alpha_2^+}$. Then the mapping $\varphi : T \to T$ such that for $f \in T$, $x \in X$,

$$\varphi(f)(x) = \begin{cases} h_i f h_i^{-1}(x), & \text{if } x \in X_i , \\ \xi(f)(x), & \text{if } x \in X - W \end{cases}$$

is an injective endomorphism of T, with $\varphi(T) \subseteq S(X,\alpha_3,\alpha_2)$.

References

[1] Balbes, R. and P. Dwinger, *Distributive Lattices*, University of Missouri Press, 1974.

[2] Clifford, A. H. and G. B. Preston, *The Algebraic Theory of Semigroups*, Mathematical Surveys, No. 7, American Mathematical Society, Providence, R.I., Vol. I, 1961, Vol. II, 1967.

[3] Levi, I., *Automorphisms of normal transformation semigroups*, Proc. Edinburgh Math. Soc. **28** (1985), 185–205.

[4] Levi, I., *Automorphisms of normal partial transformation semigroups*, Glasgow Math. J. **29** (1987), 149–157.

[5] Levi, I. and B. M. Schein, *The semigroup of one-to-one transformations with finite defects*, Glasgow Math. J. **31** (1989), 243–249.

[6] Levi, I. and W. W. Williams, *Normal semigroups of partial transformations*, I, Proc. of Algebra Conference, Lisbon, Portugal, 1988.

[7] Levi, I., *Injective endomorphisms of $\mathcal{G}_X$-normal semigroups, finite defects*, submitted.

[8] Levi, I., *Normal semigroups of one-to-one transformations*, to appear in Proc. of Edinburgh Math. Soc.

[9] Levi, I. and S. Seif, *On congruences of $\mathcal{G}_X$-normal semigroups*, SR, **43**, 93–113.

[10] Magill, Jr., K. D., *Some homomorphism theorems for a class* of *semigroups*, Proc London Math. Soc. **15** (1965), 517–526.

[11] Rubin, J. E., *Set Theory for the Mathematician*, Holden-Day, 1967.

[12] Scott, W. R., *Group Theory*, Dover Publications, 1987.

[13] Sierpinski, W., *Cardinal and ordinal numbers*, second edition, Warszawa, 1965.

[14] Sullivan, R. P., *Monomorphisms of transformation semigroups*, Semigroup Forum **32** (1985), 233–239.

Mathematics Department
University of Louisville
Louisville, KY 40292

Received February 20, 1990
and in final form March 27, 1991

Semigroup Forum Vol. 45 (1992) 23–25

SHORT NOTE

Note on Compatible Well Orders on a Free Monoid

Yuji Kobayashi

Communicated by G. Lallement

Let Σ be a finite alphabet and Σ^* be the free monoid over Σ. Let σ be a (total) order on Σ^*. We say that σ is *compatible* if $x\ \sigma\ y$ implies $zxw\ \sigma\ zyw$ for all x, y, z, $w \in \Sigma^*$. A *well order* is an order σ without infinite chain x_1, $x_2, \ldots$ such that $x_{i+1}\ \sigma\ x_i$ for all i. Compatible well orders on free monoids play an important role in the theory of string rewriting systems [1].

For a word $x \in \Sigma^*$, $|x|$ denotes the length of x. An order σ is called *length-sensitive* if $|x| < |y|$ implies $x\ \sigma\ y$ for all x, $y \in \Sigma^*$. Clearly, a length-sensitive order is a well order. In [2] we introduced a notion of strong compatibility: An order σ is *strongly compatible*, if $x_1x_2\ \sigma\ y_1y_2$, $|x_1| = |y_1|$ and $|x_2| = |y_2|$ imply $x_1zx_2\ \sigma\ y_1zy_2$ for all x_1, x_2, y_1, y_2, $z \in \Sigma^*$. Obviously a length-sensitive strongly compatible order is compatible.

Suppose an order $<$ on Σ is given. The (length-sensitive) *lexicographic order* $<_{lex}$ on Σ^* is defined as follows: Let $x = x_1 \ldots x_m$ and $y = y_1 \ldots y_n$ be in Σ^*, where x_i, $x_j \in \Sigma^*$. If $m < n$, then $x <_{lex} y$, if $m = n$ and $x_1 \ldots x_{i-1} = y_1 \ldots y_{i-1}$, $x_i < y_i$ for some $i \leq m$, then $x <_{lex} y$. The *anti-lexicographic* order $<_{al}$ is defined as: $x <_{al} y$ if and only if $\tilde{x} <_{lex} \tilde{y}$, where $\tilde{x}$ is the mirror image of x. The orders $<_{lex}$ and $<_{al}$ are typical length-sensitive strongly compatible orders.

From now on we suppose that the cardinality of Σ is 2 and $\Sigma = \{a, b\}$. In [2] we proved that there are only twelve different length-sensitive strongly compatible orders on Σ^* but there are infinitely (at least countably) many length-sensitive compatible orders on Σ^* which are not strongly compatible.

In this note we shall prove that there are actually uncountably many length-sensitive compatible orders on Σ^*.

Let s be a positive real number. We define a function f_s on Σ^* by

$$f_s(x) = \sum_{1 \leq i \leq n, a_i = b} s^{i-1} \quad \text{for} \quad x = a_1 \ldots a_n\ ,$$

where $a_i \in \Sigma$. As a special case $f_s(x)$ is considered to be 0 if $x = a^n$. Then we have

$$(*) \qquad f_s(xy) = f_s(x) + s^{|x|} f_s(y) \quad \text{for} \quad x, y \in \Sigma^*\ .$$

Now define an order σ_s as follows:

(i) if $|x| < |y|$, then $x\ \sigma_s\ y$,
(ii) if $|x| = |y|$ and $f_s(x) < f_s(y)$, then $x\ \sigma_s\ y$,
(iii) if $|x| = |y|$, $f_s(x) = f_s(y)$ and $x <_{lex} y$, then $x\ \sigma_s\ y$,

for $x, y \in \Sigma^*$. We can easily see that σ_s has the following properties (use $(*)$).

(1) σ_s is a length-sensitive compatible order on Σ^*,
(2) $\sigma_s = <_{lex}$ for $0 < s \leq 1/2$, and $\sigma_s = <_{al}$ for $s \geq 2$,
(3) σ_s is not strongly compatible for $1/2 < s < 1$ and for $1 < s < 2$.

Theorem. *For uncountably many real numbers s between 1 and 2, σ_s are different length sensitive compatible orders on Σ^*.*

Let $n \geq 2$ and let X be an indeterminate. For a vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{n-1})$ where ε_i is 0 or 1 and some of them are non-zero, let $P_\varepsilon(X)$ be the polynomial:

$$P_\varepsilon(X) = X^n - \sum_{i=1}^{n-1} \varepsilon_i X^i - 1 .$$

Lemma 1. *The equation $P_\varepsilon(X) = 0$ has a unique root in the interval $(1,2)$. With this unique root $\alpha(\varepsilon)$ we have $P_\varepsilon(t) < 0$ for $1 < t < \alpha(\varepsilon)$ and $P_\varepsilon(t) > 0$ for $t > \alpha(\varepsilon)$.*

Proof. Since $P_\varepsilon(1) > 0$ and $P_\varepsilon(2) > 0$, it suffices to show that $1 < s < t$ and $P_\varepsilon(s) \geq 0$ imply $P_\varepsilon(s) < P_\varepsilon(t)$. We shall show this by induction on n. If $n = 2$, then $P_\varepsilon(X) = X^2 - X - 1$ and the assertion holds. Suppose $n > 2$ and set $m = \min\{i \mid \varepsilon_i \neq 0\}$ and $\varepsilon' = (\varepsilon_{m+1}, \ldots, \varepsilon_{n-1})$. Then

$$P_\varepsilon(X) = X^m P_{\varepsilon'}(X) - 1 .$$

Since $P_\varepsilon(s) \geq 0$, we see $P_{\varepsilon'}(s) > 0$. By the induction hypothesis we have $P_{\varepsilon'}(s) < P_{\varepsilon'}(t)$. Consequently we find $P_\varepsilon(s) < P_\varepsilon(t)$ because $1 < s < t$.

Lemma 2. *Let* $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{n-1})$ *and* $\varepsilon_{(m)} = (\underbrace{0, \ldots, 0}_{m-1}, 1, \varepsilon_1, \ldots, \varepsilon_{n-1})$. *Then the sequence $\{\alpha(\varepsilon_{(m)})\}$ of the roots is monotone decreasing, and*

$$\lim_{m \to \infty} \alpha(\varepsilon_{(m)}) = \alpha(\varepsilon) .$$

Proof. Note that $P_{\varepsilon_{(m)}} = X^m P_\varepsilon(X) - 1$. For a real number s between $\alpha(\varepsilon)$ and 2, we see

$$P_{\varepsilon_{(m)}}(s) = s^m P_\varepsilon - 1 < s^{m+1} P_\varepsilon(X) - 1 = P_{\varepsilon_{m+1}}(s) .$$

By Lemma 1 we have

$$\alpha(\varepsilon_{(m)}) > \alpha(\varepsilon_{(m+1)} > \alpha(\varepsilon) > 1 .$$

Hence $\{\alpha(\varepsilon_{(m)}\}$ is a monotone decreasing bounded sequence and converges to some $\alpha_0 \geq \alpha$. Since $\alpha(\varepsilon_{(m)})^m \cdot P_\varepsilon(\alpha(\varepsilon_m)) = 1$, we find

$$P_\varepsilon(\alpha_0) = \lim_{m \to \infty} P_\varepsilon(\alpha(\varepsilon_{(m)})) = \lim_{m \to \infty} \frac{1}{\alpha(\varepsilon_{(m)})^m} = 0 .$$

This implies that α_0 is equal to the root $\alpha(\varepsilon)$ of $P_\varepsilon(X)$.

Let A be the subset of the interval $(1,2)$ given by

$$A = \{\alpha(\varepsilon) \mid \varepsilon = (\varepsilon_1, \ldots, \varepsilon_{n-1}) \in \{0,1\}^{n-1}, n \geq 2\}$$

and let $\overline{A}$ be the closure of A.

Lemma 3. *We have*

(1) A has no isolated points.

(2) $\overline{A}$ has the cardinality of continuum.

Proof. The assertion (1) follows from Lemma 2. By (1), $\overline{A}$ is a perfect compact set. If $\overline{A}$ contains an interval, then of course $\overline{A}$ has the cardinality of continuum. Otherwise, $\overline{A}$ is totally disconnected, and is homeomorphic to the Cantor set by [3, Corollary 30.4]. In any case $\overline{A}$ has the cardinality of continuum.

Proof of Theorem. We shall show that the set $\{\sigma_s \mid 1 < s < 2\}$ of orders on Σ^* has the cardinality of continuum. Let A_1 be the set of limits of monotone decreasing sequences in A and A_2 the set of limits of monotone increasing sequences in A. Since $\overline{A} = A_1 \cup A_2$, we see A_1 or A_2 has the cardinality of continuum by Lemma 3. Let it be B. In the sequence we show that σ_s are all different for $s \in B$.

Let $s, t \in B$ and suppose $s < t$. By the definition of the set B there exists some $r \in B$ such that $s < r < t$. The r is a root of $P_\varepsilon(X)$ for some $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{n-1})$. Let $u = a^n b$ and $v = bv_1 \ldots v_{n-1}a$ be in Σ^*, where $v_i = a$ if $\varepsilon_i = 0$ and $v_i = b$ if $\varepsilon_1 = 1$. Then we have

$$f_s(u) - f_s(v) = s^n - \sum_{i=1}^{n-1} \varepsilon_i s^i - 1 = P_\varepsilon(s) < 0 .$$

and

$$f_t(u) - f_t(v) = P_\varepsilon(t) > 0$$

by Lemma 1. These imply that $u \ \sigma_s \ v$ and $v \ \sigma_t \ u$, showing that σ_s and σ_t are different.

Remark 1. It is similarly proved that the set $\{\sigma_s \mid 1/2 < s < 1\}$ also has the cardinality of continuum.

Remark 2. It is interesting to know whether σ_s are different for all $s \in (1,2)$. This problem is related to the way how the roots $\alpha(\varepsilon)$ are distributed on $(1,2)$. Quite recenty, Professor M. Katura, Kyoto Sangyou University, has answered this question affirmatively.

References

[1] Jantzen, M., *Confluent String Rewriting*, Springer, 1988.

[2] Saito, T. T. Iwamoto, Y. Kobayashi and K. Kajitori, *Strongly compatible total orders on free monoids*, Semigroup Forum **43** (1991), 357–366.

[3] Willard, S., *General Topology*, Addison-Wesley, 1970.

Department of Information Science
Faculty of Science
Toho University
Miyama 2-2-1
Funabashi 274, Japan

Received March 10, 1990
and in final form January 24, 1991

Semigroup Forum Vol. 45 (1992) 26–37

RESEARCH ARTICLE

Interpolation of Semigroups and Integrated Semigroups*

Wolfgang Arendt, Frank Neubrander and Ulf Schlotterbeck**

Communicated by R. Nagel

0. Introduction

In 1971, S. G. Krein, G. I. Laptev and G. A. Cvetkova [K-L-C] proved that any linear (unbounded) operator A on a Banach space E such that the resolvent set contains a half-line (w, ∞), generates a C_o-semigroup on a certain (maximal) subspace Z of E (see also [Ka], [M-O-O], [Ne3]). This is a very general result, expressing the popular belief that linear dynamic systems are good-natured in some sense.

The fact that no information on the size of Z is available in general, suggests a comparison of these operators according to the actual size of Z; with $Z = E$ as the best possible case and with $Z = \{0\}$ as the worst, but still possible one (see [Be]). On the good side of this scale, the cases $D(A^k) \subset Z$ $(k \in \mathbb{N})$ are of particular interest. In this paper we want to show that this situation is actually characteristic for generators of k-times integrated semigroups which were introduced in order to treat the abstract Cauchy problem $u'(t) = Au(t)$, $u(0) = x$ in cases where the resolvent of the operator A exists and has polynomial growth in a right half-plane. Such operators frequently occur if one studies differential operators in $L^p(\mathbb{R}^n)$, $(1 \leq p \leq \infty)$, systems of linear partial differential equations or higher order Cauchy problems, to mention just a few instances (see [Ar1], [Ar2], [A-K], [dL], [K-H], [Ne1], [Ne2], [Ne3], [N-S], [Oh], [So], [T-M1], [T-M2], [Th]). It turned out that integrated semigroups share many properties with C_o-semigroups. The purpose of this paper is to show that every C_o-semigroup on a Banach space F induces integrated semigroups on continuously embedded subspaces and that, conversely, every integrated semigroup on F can be "sandwiched" by C_o-semigroups on extrapolation- and interpolation spaces.

In order to make this more precise we introduce some notation. An operator A on a Banach space E is the generator of k-times integrated semigroup (where $k \in \mathbb{N}_0$) if there exist $w \geq 0$ and $S(\cdot)\colon [0, \infty) \to \mathcal{L}(E)$ strongly continuous such that (w, ∞) is contained in the resolvent set of A, and

$$(\mu I - A)^{-1}x = \mu^k \int_0^\infty e^{-\mu t} S(t)x\,dt \qquad (x \in E, \mu > w).$$

The function $S(\cdot)$ is called *k-times integrated semigroup.* If there exists $M \geq 0$ such that $|S(t)| \leq Me^{wt}$ for all $t \geq 0$, then $S(\cdot)$ is called exponentially bounded of type w. Thus, if $S(\cdot)$ is of type w, it is also of type w' for all $w' > w$.

---

* A preliminary version of this paper appeared in: Semesterbericht Funktionalanalysis 15, Tübingen (1988/89).

** Research supported in part by NSF Grant DMS-8601983 and by DFG (Deutsche Forschungsgemeinschaft)

This notion differs slightly from the usual one in the case $k = 0$, but will be convenient for our purposes.

An operator A generates a 0-times integrated semigroup if and only if A generates a C_o-semigroup (see [Ar1]).

For Banach spaces E, F we write $E \hookrightarrow F$ if E is a subspace of F and the inclusion is continuous. We write $E \hookrightarrow_d F$ if $E \hookrightarrow F$ and E is dense in F.

Let B be an operator on F with domain $D(B)$. We denote by $\rho(B)$ the resolvent set and by $R(\mu, B) := (\mu I - B)^{-1}$ the resolvent of B in μ. If $E \hookrightarrow F$, then we denote by B_E the "part of B in E" defined by $D(B_E) := \{x \in D(B) \cap E : Bx \in E\}$, $B_E x := Bx$.

In particular, if $R \in \mathcal{L}(F)$ (the space of all bounded linear operators) such that $RE \subset E$, then R_E is the usual "restriction" of R to E. It follows from the closed graph theorem that $R_E \in \mathcal{L}(E)$.

If B is a closed operator on F, then $D(B^k)$ is a Banach space for the graph norm $|x|_{B^k} := |x| + |Bx| + ... + |B^k x|$, $(k \in \mathbb{N}_0)$. This Banach space is denoted by $[D(B)^k]$. Clearly, $[D(B^k)] \hookrightarrow F$. The part of B in $[D(B^k)]$ is denoted by B_k.

If B generates a C_o-semigroup $T(\cdot)$ on F, then B_k generates a C_o-semigroup $T_k(\cdot)$ on $[D(B^k)]$ and $T_k(t)$ coincides with the restriction of $T(t)$ to $[D(B^k)]$ for all $t \geq 0$ (see [Na]).

Now we can state the main results.

Theorem 0.1. (Interpolation Theorem) *Let B be the generator of a C_o-semigroup $T(\cdot)$ on a Banach space F.*

(a) *Assume that E is a Banach space such that $[D(B^k)] \hookrightarrow E \hookrightarrow F$ for some $k \in \mathbb{N}$. In the case $k \geq 2$ assume in addition that $R(\mu_o, B)E \subset E$ for some $\mu_o \in \rho(B)$. Then B_E generates an exponentially bounded, k-times integrated semigroup $S_E(\cdot)$ on E.*

(b) *Assume that $D(B) \neq F$. Then, given any $k \in \mathbb{N}$, there exists a Banach space E such that $[D(B^k)] \hookrightarrow E \hookrightarrow F$ and B_E generates an exponentially bounded, k-times integrated semigroup, but not a $(k-1)$-times integrated semigroup.*

The next result is a converse of Theorem 1. Given the generator A of a k-times integrated semigroup we construct a maximal inscribed space on which the part of A acts as a generator of a C_o-semigroup. An extrapolation space is obtained as well.

Theorem 0.2. (Extrapolation Theorem) *Let A be the generator of an exponentially bounded, k-times integrated semigoup of type $w > 0$ on a Banach space E. Let $\alpha > w$. Then there exists a generator B of C_o-semigroup of type α on a Banach space F such that*

(a) $[D(B^k)] \hookrightarrow E \hookrightarrow_d F$ *and* $A = B_E$*;*

(b) *the Banach space $[D(B^k)]$ is maximal unique in the following sense : If W is a Banach space such that $W \hookrightarrow E$ and A_W generates a C_o-semigroup of type α on W, then $W \hookrightarrow [D(B^k)]$.*

Note that the operator A is not necessarily densely defined. Theorem 0.2 says that A is "sandwiched" by the C_o-semigroup generators B and B_k. In combination with Theorem 0.1 one obtains the following characterization.

Corollary 0.3. *Let A be a densely defined operator on a Banach space E with non empty resolvent set. Then the following statements are equivalent.*

(a) *A generates an exponentially bounded, k-times integrated semigroup on E.*
(b) *There exists a Banach space G such that $[D(A^k)] \hookrightarrow G \hookrightarrow E$, and A_G generates a C_o-semigroup on G.*
(c) *There exists a generator B of a C_o-semigroup on a Banach space F such that $[D(B^k)] \hookrightarrow E \hookrightarrow F$, $R(\mu, B)E \subset E$ for some $\mu \in \rho(B)$ and $A = B_E$.*

These results show that the concepts of integrated semigroups and C_o-semigroups are the same up to the choice of the Banach space. However, in many instances it turns out that it is relatively easy to prove that an operator A generates an integrated semigroup on a "nice" Banach space E, whereas the construction of the inter- or extrapolation spaces on which C_o-semigroups are generated is quite tedious, if not impossible (see [A-K], [Ne1], [Ne2]).

The paper is organized as follows. Section 1 contains the basic properties of integrated semigroups which are needed later while in Section 2 and 3 the main results are proved.

1. Preliminaries

At first we define Laplace transforms of operator-valued functions. Let E be a Banach space and $S(\cdot) : [0, \infty) \to \mathcal{L}(E)$ be a strongly continuous function. For $\mu \in \mathbb{C}$ and $b \geq 0$ we define the operator $\int_0^b e^{-\mu t} S(t) dt \in \mathcal{L}(E)$ by

$$\left(\int_0^b e^{-\mu t} S(t) dt\right) x := \int_0^b e^{-\mu t} S(t) x dt \qquad (x \in E),$$

where $\int_0^b e^{-\mu t} S(t) x dt$ is the usual Riemann integral. Let $S_1(t) := \int_0^t S(s) ds$ $(t \geq 0)$. If $S(\cdot)$ is exponentially bounded, then the Laplace transform

$$\int_0^\infty e^{-\mu t} S(t) dt := \lim_{b \to \infty} \int_0^b e^{-\mu t} S(t) dt$$

of $S(\cdot)$ exists. The converse statement does not hold (see [Do], p. 38). We show that the Laplace transform of $S(\cdot)$ exists if and only if the once integrated function $S_1(\cdot)$ is exponentially bounded.

Proposition 1.1. a) *If $|S_1(t)| \leq Me^{wt}$ for some $M, w \geq 0$ and all $t \geq 0$, then*

$$\lim_{b \to \infty} \int_0^b e^{-\mu t} S(t) dt$$

exists in the operator norm for $\mu \in \mathbb{C}$ with $\operatorname{Re} \mu > w$ *and*

$$\int_0^\infty e^{-\mu t} S(t) dt = \mu \int_0^\infty e^{-\mu t} S_1(t) dt.$$

b) *Conversely, if* $\sup_{b \geq 0} \left| \int_0^b e^{-\mu t} S(t) dt \right| < \infty$, *then there exists $M \geq 0$ such that*

$|S_1(t)| \le Me^{\operatorname{Re}\mu t}$ *(if* $\operatorname{Re}\mu > 0$ *or* $\mu = 0$*) and* $|S_1(t)| \le M(1+t)$ *(if* $\operatorname{Re}\mu = 0$ *and* $\mu \neq 0$*) for all* $t \ge 0$.

Proof. a) From the assumptions it follows that

$$\left| \int_t^{t+h} e^{-\mu s} S_1(s)ds \right| \le M(e^{-(\operatorname{Re}\mu - w)t})/(\operatorname{Re}\mu - w) \to 0$$

for $t \to \infty$ uniformly in $h \ge 0$ for all $\mu \in \mathbb{C}$ with $\operatorname{Re}\mu > w$. Hence $\int_0^\infty e^{-\mu s} S_1(s)ds$ converges in the operator norm. Integrating by parts one obtains

$$\int_0^t e^{-\mu s} S(s)ds = e^{-\mu t} S_1(t) + \mu \int_0^t e^{-\mu t} S_1(s)ds \to \mu \int_0^\infty e^{-\mu s} S_1(s)ds$$

for $t \to \infty$ with respect to the operator norm.

b) By assumption, the operator family $B(t) := \int_0^t e^{-\mu t} S(s)ds$ is norm bounded. The statement follows from

$$S_1(t) = \int_0^t e^{\mu s} e^{-\mu s} S(s)ds = e^{\mu t} B(t) - \mu \int_0^t e^{\mu s} B(s)ds. \qquad \blacksquare$$

Next we define integrated semigroups and their generators.

Definition 1.2. Let $k \in \mathbb{N}_0$. An operator A on a Banach space E is called a *generator* of a k-times integrated semigroup if there exist $M, w \ge 0$ and a strongly continuous function $S(\cdot)\colon [0,\infty) \to \mathcal{L}(E)$ satisfying

$$\left| \int_0^t S(s)ds \right| \le Me^{wt} \qquad (t \ge 0)$$

such that $(w, \infty) \subset \rho(A)$ and $R(\mu, A) = \mu^k \int_0^\infty e^{-\mu t} S(t)dt$ for $\mu > w$. The function $S(\cdot)$ is called the *k-times integrated semigroup generated by A*. If there exists $C \ge 0$ and w' such that $|S(t)| \le Ce^{w't} (t \ge 0)$, then $S(\cdot)$ is called *exponentially bounded of type w'*. $\blacksquare$

We do not know whether in the situation of Definition 1.2 the function $S(\cdot)$ is automatically exponentially bounded. Thus the definition we give here might be more general than the one given in [Ar1] or [Ne1], where the function $S(\cdot)$ is always assumed to be exponentially bounded. However, by Proposition 1.1, if A generates a k-times integrated semigroup $S(\cdot)$, then A generates the exponentially bounded, $(k+1)$-times integrated semigroups $S_1(t) = \int_0^t S(s)ds$.

Moreover, the Cauchy problem with respect to A

$$\mathrm{CP}(A) \quad u'(t) = Au(t), \quad u(0) = x, \quad u(\cdot) \in C^1([0,\infty), E) \cap C([0,\infty), D(A))$$

has at most one solution for all $x \in E$ (see [Ar1] or [Ne1]).

Let A be the generator of a k-times integrated semigroup $S(\cdot)$ on E.

Define $T(t) \in \mathcal{L}([D(A^k)], E)$, $(t \geq 0)$ by

(1.1) $T(t)x := S(t)A^k x + (t^{k-1}/(k-1)!)A^{k-1}x + \ldots + tAx + x \qquad (x \in D(A^k))$.

For $x \in D(A^{k+1})$ the function $u(\cdot) := T(\cdot)x$ is the unique solution of CP(A). In fact, by the proof of Prop. 3.3 in [Ar1] one has

$$(1.2) \quad \int_0^t S(s)y\,ds \in D(A) \text{ and } A\int_0^t S(s)y\,ds = S(t)y - (t^k/k!)y \quad (y \in E, t \geq 0),$$

$$(1.3) \qquad S(t)Ay = AS(t)y \qquad (y \in D(A)).$$

In particular, for $y \in D(A)$, $S(\cdot)y \in C^1([0,\infty), E) \cap C([0,\infty), D(A))$ and

$$(1.4) \qquad \frac{d}{dt}S(t)y = S(t)Ay + (t^{k-1}/(k-1)!)y \qquad (y \in D(A)).$$

Now it follows immediately that $T(\cdot)x = (d^k/dt^k)S(\cdot)x$ for $x \in D(A^k)$ and that $T(\cdot)x$ is a solution of CP(A) whenever $x \in D(A^{k+1})$. Moreover, if $u(\cdot)$ is a solution of CP(A) with $|u(t)| \leq \text{const}\cdot e^{\alpha t}$ for some $\alpha \geq 0$, then for $\alpha < \mu \in \rho(A)$

$$(1.5) \qquad R(\mu, A)u(0) = \int_0^\infty e^{-\mu t}u(t)\,dt \qquad \text{(see [Ne1, 4.6])}.$$

Next we discuss the "rescaling" of integrated semigroups (see also [dL]).

Proposition 1.3. *Let $k \in \mathbb{N}_0$ and $r \in \mathbb{R}$. If A generates an exponentially bounded, k-times integrated semigroup, then $A - rI$ generates an exponentially bounded, k-times integrated semigroup.*

Proof. Let $S(\cdot)$ be of type w. There exists a polynomial $p(\cdot)$ of degree $k-1$ such that $\sum_{j=1}^{k}\binom{k}{j}r^j\mu^{-j} = \int_0^\infty e^{-\mu t}p(t)\,dt$, $(\mu > 0)$. Define a strongly continuous and exponentially bounded function $S_r(\cdot)\colon [0,\infty) \to \mathcal{L}(E)$ by

$$S_r(t) := e^{-rt}S(t) + \int_0^t p(t-s)e^{-rs}S(s)\,ds \qquad (t \geq 0).$$

Let $\mu > \max(0, w - r)$. It follows from Fubini's Theorem that

$$\begin{aligned}\int_0^\infty e^{-\mu t}S_r(t)\,dt &= \int_0^\infty e^{-(\mu+r)t}S(t)\,dt + \int_0^\infty e^{(-\mu+r)t}S(t)\,dt\int_0^\infty e^{-\mu s}p(s)\,ds \\ &= \left(1 + \sum_{j=1}^{k}\binom{k}{j}r^j\mu^{-j}\right)(\mu+r)^{-k}R(\mu+r, A) \\ &= (1 + r/\mu)^k(\mu+r)^{-k}R(\mu+r, A) = \mu^{-k}R(\mu, A - rI). \qquad \blacksquare\end{aligned}$$

2. Interpolation of semigroups

We first prove Theorem 0.1 for $k = 1$; i.e., we consider interpolation spaces between the given Banach space F and the domain of the C_o-semigroup generator B. For later purposes it is convenient to consider not only generators of C_o-semigroups, but also generators of integrated semigroups. Thus, for $k = 1$, Theorem 0.1 is the special case of the following theorem for $m = 0$.

Theorem 2.1. *Let $m \in \mathbb{N}_0$ and let B generate an m-times integrated semigroup $T(\cdot)$ on a Banach space F. Assume E is a Banach space such that $[D(B)] \hookrightarrow E \hookrightarrow F$. Then B_E generates an $(m+1)$-times integrated semigroup $S_E(\cdot)$ on E. Moreover, if $T(\cdot)$ is exponentially bounded, then $S_E(\cdot)$ is exponentially bounded.*

We will use the following lemma, which is easy to prove.

Lemma 2.2. *Let A be an operator on E and let B be an operator on F such that $E \hookrightarrow F$. Assume that there exists $\mu \in \rho(B)$ such that $R(\mu, B)E \subset E$. Then $A = B_E$ if and only if $\mu \in \rho(A)$ and $R(\mu, A) = R(\mu, B)_E$.*

Proof of Theorem 2.1. Let $T(\cdot): [0,\infty) \to \mathcal{L}(F)$ be the m-times integrated semigroup generated by B. Define $S(t) := \int_0^t T(s)ds \in \mathcal{L}(F)$ $(t \geq 0)$. By (1.2), $S(t)F \subset D(B)$ and so $S(t)E \subset E$ for $t \geq 0$. Set $S_E(t) := S(t)_E \in \mathcal{L}(E)$. By (1.2), $BS(t) = T(t) - (t^m/m!)I$. It follows that $S(\cdot)x \in C([0,\infty), [D(B)])$ for all $x \in F$. Consequently, $S_E(\cdot): [0,\infty) \to \mathcal{L}(E)$ is strongly continuous.

We show that $\int_0^t S_E(s)ds$ is exponentially bounded (in $\mathcal{L}(E)$). By Proposition 1.1, there exist $M, w > 0$ such that $|S(t)|_{\mathcal{L}(F)} \leq Me^{wt}$ $(t \geq 0)$. Let $x \in E$. Then we have $\int_0^t S_E(s)xds = \int_0^t S(s)xds \in D(B)$ and $B\int_0^t S_E(s)ds = S(t)x - (t^{m+1}/(m+1)!)x$. Since $[D(B)] \hookrightarrow E \hookrightarrow F$ we obtain

$$\begin{aligned}
\left|\int_0^t S_E(s)xds\right|_E &\leq \text{const} \cdot \left|\int_0^t S(s)xds\right|_B \\
&= \text{const} \cdot \left(\left|\int_0^t S(s)xds\right|_F + \left|B\int_0^t S(s)xds\right|_F\right) \\
&= \text{const} \cdot \left(\left|\int_0^t S(s)xds\right|_F + \left|S(t)x - (t^{m+1}/(m+1)!)x\right|_F\right) \\
&\leq \text{const} \cdot \left(\frac{M}{w}e^{wt}|x|_F + Me^{wt}|x|_F + (t^{m+1}/(m+1)!\,|x|_F\right).
\end{aligned}$$

Consequently, $\sup_{t\geq 0}\left|e^{-wt}\int_0^t S_E(s)xds\right|_E < \infty$, $(x \in E)$. By the uniform boundedness principle, $\left|\int_0^t S_E(s)ds\right|_{\mathcal{L}(E)} \leq Ce^{wt}$, $(t \geq 0)$ for C suitable. Similarily, one shows that $|T(t)|_{\mathcal{L}(F)} \leq Me^{wt}$, $(t \geq 0)$ implies that $|S_E(t)|_{\mathcal{L}(E)} \leq Ce^{wt}$.

Since $R(\mu, B)E \subset E$ for $\mu \in \rho(B)$, one has $\rho(B) \subset \rho(B_E)$ and $R(\mu, B_E) = R(\mu, B)_E$ for all $\mu \in \rho(B)$. For $\mu > w$ define

$$R(\mu) := \mu^{m+1}\int_0^\infty e^{-\mu t}S_E(t)dt \in \mathcal{L}(E).$$

Integrating by parts one obtains

$$R(\mu)x = \mu^m \int_0^\infty e^{-\mu t}T(t)xdt = R(\mu, B)x = R(\mu, B_E)x \quad (x \in E, \mu > w).$$

By definition, this means that B_E generates the $(m+1)$-times integrated semigroup $S_E(\cdot)$ on E. ∎

Next we prove statement (a) of Theorem 0.1 for arbitrary k. We will frequently use the following fact. Let B be a closed operator and $k \in \mathbb{N}$. If $\mu \in \rho(B)$, then $x \to |(\mu I - B)^k x|$ defines an equivalent norm on $[D(B^k)]$.

Proof of Theorem 0.1. We show that $R(\mu, B)E \subset E$ for all $\mu \in \rho(B)$. Let $\mu, \mu_0 \in \rho(B)$. Iterating the resolvent equation $R(\mu, B) = R(\mu_0, B) + (\mu_0 - \mu)R(\mu_0, B)R(\mu, B)$ yields

$$R(\mu, B) = \sum_{j=1}^{k-1}(\mu_0 - \mu)^{j-1}R(\mu_0, B)^j + (\mu_0 - \mu)^{k-1}R(\mu_0, B)^{k-1}R(\mu, B). \tag{2.1}$$

By assumption or by $R(\mu_0, B)F \subset D(B)$ (for $k = 1$), we have $R(\mu_0, B)E \subset E$ and $R(\mu_0, B)^{k-1}R(\mu, B)E \subset D(B^k) \subset E$. This proves the claim.

Let B generate a C_0-semigroup $T(\cdot)$ with $|T(t)| \leq Me^{wt}$ $(t \geq 0)$ for some $M, w > 0$. Considering $B - rI$ if necessary instead of B, we can assume that $0 \in \rho(B)$ (see Proposition 1.3). The exponentially bounded, k-times semigroup generated by B is given by $S(t) := \int_0^t ((t-s)^{k-1}/(k-1)!)T(s)ds$. It follows from (1.1) that

$$S(t) = B^{-k}T(t) - \sum_{i=0}^{k-1}(t^i/i!)B^{-k+i}. \tag{2.2}$$

Consequently, $S(t)E \subset E$, $S_E(t) := S(t)_E \in \mathcal{L}(E)$ and $S_E(\cdot): [0, \infty) \to \mathcal{L}(E)$ is strongly continuous. One obtains from (2.2) that

$$\begin{aligned}
|S_E(t)x|_E &\leq \sum_{i=0}^{k-1}(t^i/i!)\,|B^{-k+i}|_{\mathcal{L}(E)}\,|x|_E + |B^{-k}T(t)x|_E \\
&\leq \text{const} \cdot \left(e^{wt}\,|x|_E + |B^{-k}T(t)x|_{B^k}\right) \leq \text{const} \cdot e^{wt}\,|x|_E + \text{const} \cdot |T(t)x|_F \\
&\leq \text{const} \cdot e^{wt}\,|x|_E + \text{const} \cdot e^{wt}\,|x|_F \leq \text{const} \cdot e^{wt}\,|x|_E .
\end{aligned}$$

Hence, by the uniform boundedness principle, $S_E(\cdot)$ is exponentially bounded. Now one proceeds as in the proof of Theorem 2.1. ∎

We show by an example that in case $k > 1$ the hypothesis of E being invariant under the resolvent cannot be omitted in Theorem 0.1 (a).

Example 2.3. Let B generate a C_0-semigroup on a Banach space F. Assume that $D(B) \neq F$. Let $w \in F \backslash D(B)$ and $E := D(B^2) + \mathbb{R}.w$. Then E is a Banach space for the norm $|x + cw|_E := |x|_{B^2} + |cw|_F$. Clearly, $[D(B^2)] \hookrightarrow E \hookrightarrow F$. But B_E does not generate a k-times integrated semigroup for any $k \in \mathbb{N}$. In fact, assume that there exists $\mu \in \rho(B_E) \cap \rho(B)$. Then $R(\mu, B)E \subset E$. So there are $x \in D(B^2)$, $c \in \mathbb{R}$ such that $R(\mu, B)w = x + cw$. Hence $cw = R(\mu, B)w - x \in D(B)$. Thus $c = 0$. But then $R(\mu, B)w = x \in D(B^2)$. This implies $w \in D(B)$, which is a contradiction. ∎

Next we will prove statement (b) of Theorem 0.1. For that we need the following two lemmas.

Lemma 2.4. *Let $S(\cdot)$ be an exponentially bounded, $(k+1)$-times integrated semigroup on F with generator B. Then B generates a k-times integrated semigroup if and only if $S(\cdot)x \in C([0, \infty), [D(B)])$ for all $x \in F$.*

Proof. Assume that $S(\cdot)x \in C([0,\infty),[D(B)])$ for all $x \in F$. By (1.2), $T(\cdot)x := d/dtS(\cdot)x \in C([0,\infty),E)$ for all $x \in F$. Hence

$$R(\mu,B) = \mu^{k+1}\int_0^\infty e^{-\mu t}S(t)dt = \mu^k\int_0^\infty e^{-\mu t}T(t)dt$$

for μ large. By definition, B generates the k-times integrated semigroup $T(\cdot)$. The converse follows from (1.2). ∎

Lemma 2.5. *Let $S(\cdot)$ be an exponentially bounded, k-times integrated semigroup on F with generator B. Assume that $D(B) \neq F$ in the case $k = 0$ and that B does not generate a $(k-1)$-times integrated semigroup in the case $k > 0$. Then there exists a Banach space E such that $[D(B)] \hookrightarrow E \hookrightarrow F$ and such that B_E generates an exponentially bounded $(k+1)$-times integrated semigroup but not a k-times integrated semigroup on E.*

Proof. By Lemma 2.4, there is $w \in F$ such that $S(\cdot)w \notin C([0,\infty),[D(B)])$. By (1.3), $w \notin D(B)$ so that $E := [D(B)] + \mathbb{R}.w$ is a direct sum. Define $|x+cw|_E := |x|_B + |cw|_F$. Then $[D(B)] \hookrightarrow E \hookrightarrow F$. By Theorem 2.1, B_E generates a $(k+1)$-times integrated semigroup on E. Suppose that B_E generates a k-times integrated semigroup $S_E(\cdot)$ on E. Then

$$\mu^k\int_0^\infty e^{-\mu t}S_E(t)ydt = R(\mu,B_E)y = R(\mu,B)y = \mu^k\int_0^\infty e^{-\mu t}S(t)ydt$$

($y \in E$, μ large). So it follows from the uniqueness theorem for Laplace transforms that $S_E(\cdot)y = S(\cdot)y$ for all $y \in E$. Consequently $S(\cdot)y \in C([0,\infty),E)$ for all $y \in E$. In particular, there exists $S_1(\cdot)w \in C([0,\infty),[D(B)])$ and $c(\cdot) \in C([0,\infty))$ such that $S(\cdot)w = S_1(\cdot)w + c(\cdot)w$. Hence $h(\cdot) := \int_0^{\cdot}(S(s)w - c(s)w)ds \in C([0,\infty),[D(B)])$. By (1.2), $\int_0^t S(s)wds \in D(B)$ for $t \geq 0$. Hence $(\int_0^t c(s)ds)w \in D(B)$ for all $t \geq 0$. Since $w \notin D(B)$, we conclude $c(\cdot) = 0$. But then $S(\cdot)w = S_1(\cdot)w \in C[0,\infty),[D(B)])$. This contradicts Lemma 2.4. ∎

Proof of Theorem 0.1(b). Let B be the generator of a C_0-semigroup $T(\cdot)$ on a Banach space F with $D(B) \neq F$. By Lemma 2.5, there exists a Banach space E_1 such that $[D(B)] \hookrightarrow E_1 \hookrightarrow E_0 := F$ and such that the part B_1 of B in E_1 generates an exponentially bounded, 1-times integrated semigroup, but not a 0-times integrated semigroup. By Lemma 2.5, there exists a Banach space E_2 such that $[D(B_1)] \hookrightarrow E_2 \hookrightarrow E_1$ and such that the part B_2 of B_1 in E_2 generates an exponentially bounded, 2-times integrated semigroup, but not a 1-times integrated semigroup. Then B_2 is the part of B in E_2. Hence we found a Banach space E_2 such that $[D(B^2)] \hookrightarrow E_2 \hookrightarrow F$ and such that the part of B in E_2 generates an exponentially bounded, 2-times integrated semigroup, but not a 1-times integrated semigroup. Proceeding in this manner one obtains inductively Banach spaces E_k such that $[D(B^k)] \hookrightarrow E_k \hookrightarrow F$ and such that the part of B in E_k generates an exponentially bounded k-times integrated semigroup, but not a $(k-1)$-times integrated semigroup. ∎

3. Extrapolation of integrated semigroups

In this section we prove Theorem 0.2 and Corollary 0.3. Let $k \in \mathbb{N}$ and $S(\cdot)$ be a k-times integrated semigroup on E of type $w > 0$ with generator A. Let $\mu_0 > \alpha > w$ be fixed. Define F to be the completion of E with respect to the norm

(3.1) $$|x|_F := \sup_{t\geq 0}\left|e^{-\alpha t}T(t)R(\mu_0,A)^k x\right|_E,$$

where $T(\cdot)$ is given by (1.1). Since $\alpha > 0$, it follows that $|x|_F \leq \text{const} \cdot |x|_E$ $(x \in E)$. Thus $E \hookrightarrow_d F$. Next we show that

$$|(\mu - \alpha)R(\mu, A)x|_F \leq |x|_F \quad (\mu > \alpha, x \in E). \tag{3.2}$$

Let $y \in D(A^k)$ and $t \geq 0$. Then $u(s) := T(t+s)R(\mu_0, A)y$, $(s \geq 0)$ is a solution of CP(A) for the initial value $x = T(t)R(\mu_0, A)y$.
It follows from (1.5) that $R(\mu_0, A)R(\mu, A)T(t)y = R(\mu_0, A)\int_0^\infty e^{-\mu s}T(t+s)y\,ds$ $(\mu > \alpha)$. Hence

$$R(\mu, A)T(t)y = \int_0^\infty e^{-\mu s}T(t+s)y\,ds \qquad (\mu > \alpha) \tag{3.3}$$

for all $y \in D(A^k)$. Let $x \in E$ and $\mu > \alpha$. Using (3.3) one obtains

$$\begin{aligned}
\left|e^{-\alpha t}T(t)R(\mu_0, A)^k R(\mu, A)x\right|_E &= \left|e^{-\alpha t}\int_0^\infty e^{-\mu s}T(t+s)R(\mu_0, A)^k x\,ds\right|_E \\
&= \left|\int_0^\infty e^{-(\mu-\alpha)s}e^{-\alpha(t+s)}T(t+s)R(\mu_0, A)^k x\,ds\right|_E \\
&\leq \int_0^\infty e^{-(\mu-\alpha)s}ds\,|x|_F = |x|_F/(\mu - \alpha).
\end{aligned}$$

Since $t \geq 0$ is arbitrary this implies (3.2). It follows from (3.2) that $R(\mu, A)$ has a unique extension $R(\mu) \in \mathcal{L}(F)$ satisfying

$$|(\mu - \alpha)R(\mu)|_{\mathcal{L}(\mathcal{F})} \leq 1 \qquad (\mu > \alpha). \tag{3.4}$$

Then $\{R(\mu) : \mu > \alpha\}$ is a pseudo resolvent on F. We show that

$$\lim_{\mu\to\infty} |\mu R(\mu)y - y|_F = 0 \qquad (y \in F). \tag{3.5}$$

Because of (3.4) and density it sufficies to show (3.5) for $y \in E$. Let $x = R(\mu_0, A)^k y$. Then, by (3.3),

$$\begin{aligned}
|\mu R(\mu)y - y|_F &= \sup_{t\geq 0}\left|e^{-\alpha t}(\mu R(\mu, A)T(t)x - T(t)x)\right|_E \\
&= \sup_{t\geq 0}\left|e^{-\alpha t}\int_0^\infty \mu e^{-\mu s}(T(t+s)x - T(t)x)ds\right|_E.
\end{aligned}$$

Let $\varepsilon > 0$. Since $S(\cdot)$ is of type $w > 0$ there exists $M \geq 0$ such that $|T(t+s)x - T(t)x| \leq Me^{w(t+s)}$ for all $s, t \geq 0$. Since $\alpha > w$, there exists $q > 0$ such that

$$\begin{aligned}
&\sup_{t\geq q}\left|e^{-\alpha t}\int_0^\infty \mu e^{-\mu s}(T(t+s)x - T(t)x)ds\right|_E \\
&< \text{const} \cdot \mu \cdot e^{(w-\alpha)q}/(\mu - w) < \varepsilon/2
\end{aligned} \tag{3.6}$$

for all $\mu > \alpha$. Since $T(\cdot)$ is uniformly continuous on compact intervals, there exists $\delta > 0$ such that $|T(t+s)x - T(t)x|_F \leq \varepsilon/2$ for $t \in [0, q], s \in [0, \delta]$. Hence

$$\sup_{0\leq t\leq q}\left|e^{-\alpha t}\int_0^\delta \mu e^{-\mu s}(T(t+s)x - T(t)x)ds\right|_E < \varepsilon/2 \qquad (\mu > \alpha). \tag{3.7}$$

Since $S(\cdot)$ is of type w,

$$\sup_{0\le t\le q}\left|e^{-\alpha t}\int_\delta^\infty \mu e^{-\mu s}(T(t+s)x-T(t)x)ds\right|_E \le \mathrm{const}\cdot\mu\cdot e^{(w-\mu)\delta}/(\mu-w)\to 0$$

for $\mu\to\infty$. This together with (3.6), (3.7) shows that $\overline{\lim_{\mu\to\infty}}\,|\mu R(\mu)y-y|_F\le\varepsilon$. Since $\varepsilon>0$ is arbitrary, (3.5) is proved.

It follows from (3.5) that $\mathrm{Ker}R(\mu)=\{0\}$, ($\mu>0$; observe that $\mathrm{Ker}R(\mu)$ is independent of $\mu>\alpha$ because of the resolvent equation). Hence there exists an operator B on F such that $(\alpha,\infty)\subset\rho(A)$ and $R(\mu)=R(\mu,B)$ for $\mu>\alpha$ (see the proof of Theorem 1.9.3 in [Pa]). Because of (3.5) the domain of B is dense in F. It follows from (3.4) and the Hille-Yosida Theorem that B generates a C_0-semigroups of type α on F.

Since by construction $R(\mu,B)E\subset E$ and $R(\mu,B)_E=R(\mu,A)$ $(\mu>\alpha)$, it follows by Lemma 2.2 that $A=B_E$. Taking $t=0$ in (3.1) one obtains

$$|R(\mu_0,A)^k y|_E\le|y|_F\qquad(y\in E).\tag{3.8}$$

Let $x\in D(B^k)$ and $y\in F$ with $x=R(\mu_0,B)^k y$. Then there exist $y_n\in E$ such that $y_n\to y$ in F. By (3.8), $R(\mu_0,A)^k y_n$ is a Cauchy seqence in E. Hence $x=F-\lim R(\mu_0,B)^k y_n=E-\lim R(\mu_0,A)^k y_n\in E$. This implies $D(B^k)\subset E$ and $|R(\mu_0,B)^k y|_E\le|y|_F$ for all $y\in F$. Consequently $|x|_E\le|(\mu_0-B)^k x|_F$ for all $x\in D(B^k)$. This shows that $[D(B^k]\hookrightarrow E$.

We have proved part a) of Theorem 0.2. Before giving the proof of part b) we observe that $G:=[D(B^k)]$ is a Banach space with the norm $|x|_G:=|(\mu_0-B)^k x|_F$ (which is equivalent to the graph norm). Since $A=B_E$, it follows $A_G=B_k$ (the part of B in $[D(B^k)]$). Moreover, $R(\mu_0,B)^k$ is an isometric isomorphism from F onto G which coincides with $R(\mu_0,A)^k$ on E. Since E is dense in F, it follows that $D(A^k)=R(\mu_0,B)^k E$ is dense in G. For $x\in D(A^k)$ the norm is given by

$$|x|_G:=\sup_{t\ge0}|e^{-\alpha t}T(t)x|_E,\tag{3.9}$$

where $T(\cdot)$ is given by (1.1). Now we prove the maximality assertion (b). Assume that $W\hookrightarrow E$ such that A_W generates a C_0-semigroup $T_W(\cdot)$ of type α on W. Then $D(A_W^k)\subset D(A^k)$ and for $x\in D(A_W^k)$ one has (by (3.5) for $t=0$)

$$\int_0^\infty e^{-\mu s}T_W(s)xds=R(\mu,A_W)x=R(\mu,A)x=\int_0^\infty e^{-\mu s}T(s)xds\qquad(\mu>\alpha).$$

So the uniqueness of the Laplace transform implies that $T(\cdot)x=T_W(\cdot)x$. Consequently, $|x|_G:=\sup_{t\ge0}|e^{-\alpha t}T(t)x|_E\le\mathrm{const}\cdot\sup_{t\ge0}|e^{-\alpha t}T_W(t)x|_W\le\mathrm{const}\cdot|x|_W$ for all $x\in D(A_W^k)$ since $T_W(\cdot)$ is of type α. This implies that

$$W=W\text{-closure of }D(A_W^k)\quad\hookrightarrow\quad G\text{-closure of }D(A^k)=G.\qquad\blacksquare$$

Remark 3.1. a) In the situation of Theorem 0.2 one has $\rho(A)=\rho(B)$. In fact, by the construction itself it follows that $R(\mu,B)E\subset E$ for $\mu>\alpha$. Hence $\rho(B)\subset\rho(A)$ Theorem 0.1. Conversely, assume that $\mu\in\rho(A)$. Then it follows from (3.1) that $R(\mu,A)$ has a continuous extension $R(\mu)$ to F. It is easy to see that $R(\mu)=R(\mu,B)$.

b) One might define the norm $|\cdot|_G$ on $D(A^k)$ directly by formula (3.9), and then define the space G as the completion of $(D(A^k), |\cdot|_G)$. Doing so, one has to prove that G can be identified with a subspace of E. It is this point which was missed in [Ke] and [Ne1]. The proofs given there can be "repaired" if one replaces the operators $T(t)$ by their closures $(\mu_0 - A)^k T(t) R(\mu_0, A)^k$ with domain $\{x \in E : T(t)R(\mu_0, A)^k x \in D(A^k)\}$. However, these proofs a far more technical than the one given above.

Proof of Corollary 0.3. The implications (a) $\rightarrow$ (c) follow from Theorem 0.2. Choosing $G := [D(B^k)]$ in (c) one sees that (c) $\rightarrow$ (b). If (b) holds, then, for every initial value $x \in D(A^{k+1}) \subset D(A_G)$, there exists a unique solution

$$u(\cdot, x) \in C^1([0,\infty), E) \subset C^1([0,\infty), E)$$

of CP(A) with $u(t,x) \in D(A_G) \subset D(A)$ and

$$|u(t,x)| \leq \text{const} \cdot |u(t,x)|_G \leq \text{const} \cdot e^{\alpha t} |x|_G \leq \text{const} \cdot e^{\alpha t} |x|_{A^k}.$$

With this, statement (a) follows from Theorem 4.2 in [Ne1]. ∎

References

[Ar1] Arendt, W., *Vector valued Laplace transform and Cauchy problems*, Israel J. Math. **59** (1987), 327–352.

[Ar2] Arendt, W. *Resolvent positive operators*, Proc. London Math. Soc. **54** (1987), 321–349.

[A-K] Arendt, W., and H. Kellermann, *Integrated solutions of Volterra integro-differential equations and applications*, In : Integro-differential Equations, Proc. Conf. Trento 1987. G. Da Prato, M. Iannelli (eds.). Pitman. Research Notes in Mathematics **190** (1989), 21–51.

[Be] Beals, R., *On the abstract Cauchy problem*, J. Funct. Anal. **10** (1972), 281–299.

[dL] de Laubenfels, R., *Integrated semigroups, C-semigroups and the abstract Cauchy problem*, Semigroup Forum **41** (1990), 83–95.

[Do] Doetsch, G., *Handbuch der Laplace-Transformation I*, Birkhäuser Verlag, Basel 1950.

[Ka] Kantorovitz, S., *The Hille-Yosida space of an arbitrary operator*, J. Math. Anal. Appl. **136** (1988), 107–111.

[Ke] Kellermann, H., *Integrated semigroups*, Dissertation, Universität Tübingen, 1986.

[K-H] Kellermann, H., and M. Hieber, *Integrated semigroups*, J. Functional Anal. **84** (1989), 160–180.

[K-L-C] Krein, S. G., Laptev, G. I., and G.A. Cvetkova, *On Hadamard correctness of the Cauchy problem for the equation of evolution*, Soviet Math. Dokl. **11** (1970), 763–766.

[M-O-O] Miyadera, I., Oharu, S., and N. Okazawa, *Generation theorems of semi-groups of linear operators*, Publ. Res. Inst. Math. Sci., Kyoto Univ. **8** (1973), 509–555.

[Na] Nagel, R., *Sobolev spaces and semigroups*, Semesterbericht Funktionalanalysis Tübingen, Sommersemester 1984.

[Ne1] Neubrander, F., *Integrated semigroups and their applications to the abstract Cauchy problem*, Pacific J. Math. **135** (1988), 111–155.

[Ne2] Neubrander, F., *Integrated semigroups and their application to complete second order problems*, Semigroup Forum **38** (1989), 233–251.

[Ne3] Neubrander, F., *Abstract elliptic operators, analytic interpolation semigroups, and Laplace transforms of analytic functions*, Semesterbericht Funktionalanalysis, Tübingen, Wintersemester 1988/89.

[N-S] Neubrander, F., and B. Straub, *Fractional powers of operators with polynomially bounded resolvent*, Semesterbericht Funktionalysis, Tübingen, Wintersemester 1988/89.

[Oh] Oharu, S., *Semigroups of linear operators in a Banach space*, Publ. RIMS, Kyoto Univ. **7** (1971), 205–260.

[Pa] Pazy, A., "Semigroups of Linear Operators and Applications to Partial Differential Equations," Springer Verlag, New-York 1983.

[So] Sova, M., *Problème de Cauchy pour équations hyperboliques opérationnelles à coefficients constants non-bornés*, Ann. Scuola Norm. Sup. Pisa, **22** (1968), 67–100.

[T-M1] Tanaka, N., and I. Miyadera, *Some remarks on C-semigroups and integrated semigroups*, Proc. Japan Acad. **63** (1987), 139–142.

[T-M2] Tanaka, N., and I. Miyadera, *Exponentially bounded C-semigroups and integrated semigroups*, Tokyo J. Math. **12** (1989), 99–115.

[Th] Thieme, H. R., *Integrated semigroups and integrated solutions to abstract Cauchy problems*, J. Math. Analysis and Appl. **152** (1990), 416–447.

Equipe de Mathématiques
Université de Franche-Comté
F-25030 Besançon-Cedex
France

Lousiana State University
Baton Rouge, La 70803
USA

Mathematisches Institut
Universität Tübingen
D-7400 Tübingen
RFA

Received June 10, 1990
and in final form September 21, 1991

Semigroup Forum Vol. 45 (1992) 38–44

RESEARCH ARTICLE

On the structure of some group codes

Dongyang Long*

Communicated by B. M. Schein

In this paper, we study the following problem: Which characteristics does a code C possess when the syntactic monoid $syn(C^*)$ of the star closure C^* of C is a group? For a code C, if the syntactic monoid $syn(C^*)$ is a group, then we call C a group code. This definition of a group code is different from the one in [1] (see [1], 46–47). Schützenberger had characterized the structure of finite group codes and had proved that C is a finite group code if and only if C is a full uniform code (see [5], [8]). For k-prefix and k-suffix codes with $k \geq 2$, k-infix, k-outfix, p-infix, s-infix, right semaphore codes and left semaphore codes, etc., we obtain similar results. It is proved that the above mentioned codes are group codes if and only if they are uniform codes.

1. Basic notions and notation

We first introduce the necessary concepts and some notation. For additional details and definitions, see the references, in particular [1], [5], and [7].

Let X be an alphabet. Then X^* denotes the free monoid generated by X, that is, the set of all words over X, including the empty word 1, and $X^+ = X^* \backslash \{1\}$. For $w \in X^*$, by $|w|$ we denote the length of w. A language L over X is a set $L \subseteq X^*$. With any language $L \subseteq X^*$, one associates its principal congruence P_L and its syntactic monoid $syn(L) = X^*/P_L$, where

$$u \equiv v(P_L) \Longleftrightarrow ((\forall\, x, y \in X^*)\, xuy \in L \longleftrightarrow xvy \in L)\ .$$

For $w \in X^*$, by $\overline{w}$ we denote the P_L-class of w.

A language $L \subseteq X^*$ is said to be a code over X if the submonoid L^* of X^* generated by L is freely generated by L. If P is any property of languages, we call a code C a P-code if C possesses the property P. If C is a P-code and for every $u \in X^*$, $C \cup \{u\}$ is not a P-code, then C is a maximal P-code. For the sake of simplicity, in this paper we assume that X is the least alphabet for a code, that is, let C be a code over X, then $X^*aX^* \cap C \neq \emptyset$ for every $a \in X$.

Definition 1. (See [2], [3], [7].) Let X be an alphabet and k be a given positive integer. A language $C \subseteq X^*$ is

(a) a k-prefix code if for all $x_1, \ldots, x_k$, $y_1, \ldots, y_k \in X^*$, $x_1 \ldots x_k \in C$ and $x_1 y_1 x_2 \ldots x_k y_k \in C$ together imply $y_1 \ldots y_k = 1$;

(b) a k-suffix code if for all $x_1, \ldots, x_k$, $y_1, \ldots, y_k \in X^*$, $x_1 \ldots x_k \in C$ and $y_1 x_2 y_1 \ldots y_k x_k \in C$ together imply $y_1 \ldots y_k = 1$;

---

* The author would like to thank the referee who helped the author with revising the paper. Thanks are also due to Prof. Boris M. Schein for his valuable suggestions and kind help.

(c) a k-infix code if for all $x_1,\ldots,x_k$, $y_0,\ldots,y_k \in X^*$, $x_1\ldots x_k \in C$ and $y_0x_1y_1\ldots x_ky_k \in C$ together imply $y_0\ldots y_k = 1$;

(d) a k-outfix code if for all $x_0,\ldots,x_k$, $y_1,\ldots,y_k \in X^*$, $x_0\ldots x_k \in C$ and $x_0y_1x_1\ldots y_kx_k \in C$ together imply $y_1\ldots y_k = 1$;

(e) a hypercode if for any natural number n and all $x_1,\ldots,x_n$, $y_0,\ldots,y_n \in X^*$, $x_1\ldots x_k \in C$ and $y_0x_1y_1\ldots x_ny_n \in C$ together imply $y_0y_1\ldots y_k = 1$;

(f) a full uniform code if there exists some integer $m \geq 0$ such that $C = X^m$.

By $P_k(X)$, $S_k(X)$, $I_k(X)$, $O_k(X)$, $H(X)$ and $FUF(X)$ we denote the classes of k-prefix codes, k-suffix codes, k-infix codes, k-outfix codes, hypercodes and full uniform codes over X, respectively. In particular, $P(X) = P_1(X)$, $S(X) = S_1(X)$, $I(X) = I_1(X)$, $O(X) = O_1(X)$ are the classes of prefix, suffix, infix, and outfix codes, respectively.

Definition 2. (See [1], [2].) Let X be an alphabet. A language $C \subseteq X^*$ is

(a) a bifix code if C is both a prefix and a suffix code;

(b) reflective if for all $u,v \in X^*$, $uv \in C$ imply $vu \in C$;

(c) a p-infix code if for all $x,u,y \in X^*$, $xuy \in C$ and $u \in C$ together imply $y = 1$;

(d) an s-infix code if for all $x,u,y \in X^*$, $xuy \in C$ and $u \in C$ together imply $x = 1$;

(e) a right semaphore code if C is a prefix code satisfying $X^*C \subseteq CX^*$;

(f) a left semaphore code if C is a suffix code satisfying $CX^* \subseteq X^*C$.

By $B(X)$, $RE(X)$, $PI(X)$, $SI(X)$, $RSP(X)$, and $LSP(X)$ we denote the classes of bifix, reflective, p-infix, s-infix, right semaphore and left semaphore codes over X, respectively.

According to the results of [3] and Figure 1 in [4], the relations between the classes of codes defined by Definitions 1 and 2 are shown in Figures 1 and 2, respectively.

Remark 1. In [3], we proved the following:

(1) $S_{k+1}(X) \subset P_k(X)$, $P_{k+1}(X) \subset S_k(X)$, $O_{k+1}(X) \subset I_k(X)$, $I_{k+1}(X) \subset O_k(X)$.

(2) $P_{k+1}(X) \subset I_k(X)$, $P_{k+1}(X) \subset O_k(X)$, $S_{k+1}(X) \subset I_k(X)$, $S_{k+1}(X) \subset O_k(X)$.

(3) $I_k(X) \subset P_k(X)$, $O_k(X) \subset P_k(X)$, $I_k(X) \subset S_k(X)$, $O_k(X) \subset S_k(X)$.

It is easy to verify that the classes of k-prefix codes and k-suffix codes, k-infix codes and k-outfix codes, are not comparable, that is, $P_k(X) \not\subseteq S_k(X)$ and $S_k(X) \not\subseteq P_k(X)$, $I_k(X) \not\subseteq O_k(X)$ and $O_k(X) \not\subseteq I_k(X)$. In fact, let $C_1 = \{a_1a_2\ldots a_{2k-1}a_{2k},\ a_1a_2^2\ldots a_{2k}^2\}$, $C_2 = \{a_1\ldots a_{2k},\ a_1^2\ldots a_{2k}^2\}$, where $a_i \in X$, $i = 1,\ldots,2k$, $a_j \neq a_{j+1}$, $j = 1,\ldots,2k-1$. It is easy to show that $C_1 \notin P_k(X)$ and $C_1 \in S_k(X)$, $C_2 \in P_k(X)$ and $C_2 \notin S_k(X)$. Similarly, let $C_3 = \{a_1\ldots a_{2k+1},\ a_1a_2^2\ldots a_{2k+1}^2\}$, $C_4 = \{a_1\ldots a_{2k},\ a_1^2\ldots a_{2k+1}^2\}$, where $a_i \in X$, $a_j \neq a_{j+1}$, $i = 1,\ldots,2k+1$, $j = 1,\ldots,2k$. One proves that $C_3 \notin O_k(X)$ and $C_3 \in I_k(X)$. $C_4 \in O_k(X)$ and $C_4 \notin I_k(X)$.

Remark 2. $RE(X) \subset O(X)$ was not shown in Figure 1 in [2]. By Definitions 1 and 2, it is easy to see that $RE(X) \subset O(X)$.

2. The structure of finite group codes

Theorem 1. (See [5], p.213, and [8].) *Let $C \subseteq X^+$ be a finite code. Then the following conditions are equivalent:*

(1) $syn(C^*)$ is a group;

(2) $syn(C^*)$ is a cyclic group of order n for some n;

(3) C is a full uniform code, $C = X^n$.

The condition of finiteness in Theorem 1 may not be omitted. We give an example of an infinite bifix code C with $syn(C^*)$ a group. Let $X = \{a, b\}$, $C = \{a\} \cup ba^*b$, clearly, $C \cap CX^+ = \emptyset$, $C \cap X^+C = \emptyset$, that is, C is a bifix code. Also, $C^* = \{w \in X^* \mid |w|_b$ is an even number, where $|w|_b$ denotes the number of occurrences of b in w. Obviously, $syn(C^*) = \{\overline{1}, \overline{b}\}$, $\overline{b}^2 = \overline{1}$, and $syn(C^*)$ is a cyclic group of order 2.

In the sequel, a code C with $syn(C^*)$ a group is called a group code.

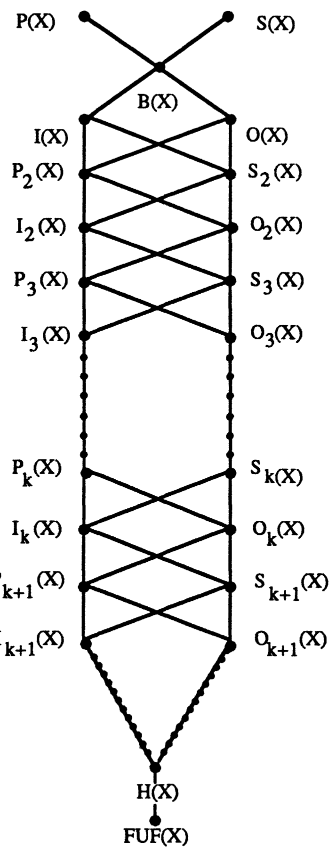


Figure 1

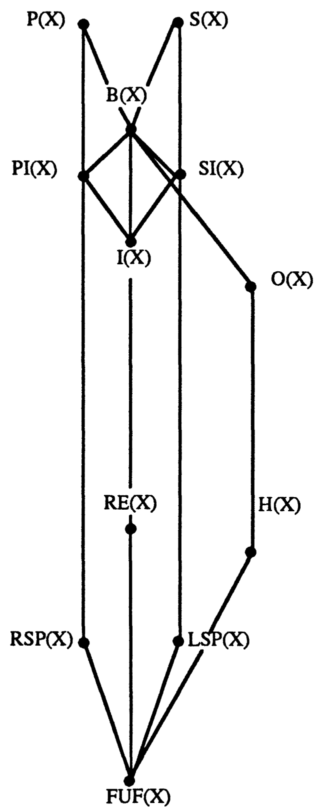


Figure 2

3. Main results

Theorem 2. *Let $C \subseteq X^+$ be a k-infix code or a k-outfix code, or a hypercode, or a reflective code. Then the following conditions are equivalent:*

(1) $syn(C^*)$ is a group;
(2) $syn(C^*)$ is a cyclic group of order n for some n;
(3) C is a full uniform code, $C = X^n$;
(4) C is a maximal code;
(5) C is a maximal prefix (suffix, bifix) code;
(6) C^* is reflective.

Theorem 3. *Let $C \subseteq X^+$ be a p-infix or a right semaphore code. Then the following conditions are equivalent:*

(1) $syn(C^*)$ is a group;
(2) $syn(C^*)$ is a cyclic group of order n for some n;
(3) C is a full uniform code, $C = X^n$;
(4) C is a maximal suffix code;
(5) C^* is reflective.

Corollary 4. *Let $C \subseteq X^+$ be an s-infix or a left semaphore code. Then the following conditions are equivalent:*

(1) $syn(C^*)$ is a group;
(2) $syn(C^*)$ is a cyclic group of order n for some n;
(3) C is a full uniform code, $C = X^n$;
(4) C is a maximal prefix code;
(5) C is reflective.

Remark 3. A p-infix code or a right semaphore code need not be a group code when it is a maximal code or a maximal prefix code. let $X = \{a, b\}$ and $C = \{b, ab, a^2\}$. Clearly, C is a p-infix code and also a right semaphore code which is a maximal code and also a maximal prefix code. But $syn(C^*)$ is not a group code. Similarly, an s-infix code or a left semaphore code need not be a group code when it is a maximal code or a maximal suffix code.

4. Proof of the theorems

We first state the following lemmas:

Lemma 1. (See [3], Theorem 2.7, and Theorem 5 in [2] or Propositions 4.10 and 4.13 in [4].) *Let $C \subseteq X^+$ be a k-infix or k-outfix code. Then the following conditions are equivalent:*

(1) C is a maximal code;
(2) C is a maximal prefix code;
(3) C is a full uniform code, $C = X^n$ for some n.

Lemma 2. (See [5].) *Let $C \subseteq X^*$ be a code. For every $w \in X^*$, if $wC^* \cap C^* \neq \emptyset$ and $C^*w \cap C^* \neq \emptyset$, then $w \in C^*$.*

Lemma 3. *Let C be a code over X. If $syn(C^*)$ is a group, then C is a maximal prefix code and a maximal suffix code.*

Proof. Assume that $u, v \in C$, and there exists $x \in X^*$ such that $uv = v$. Since $syn(C^*)$ is a group, $\overline{u}$ is an invertible element. Therefore, there exists w such that $\overline{u}\,\overline{w} = \overline{w}\,\overline{u} = \overline{1}$, where $\overline{1}$ is the identity of $syn(C^*)$. Thus $uw \equiv wu \equiv 1(P_{C^*})$. By $1 \in C^*$, $uw, wu \in C^*$, and $wC^* \cap C^* \neq \emptyset$, $C^*w \cap C^* \neq \emptyset$. According to Lemma 2, we have $w \in C^*$. Also, $ux = v$, $\overline{x} = \overline{1}\,\overline{x} = \overline{w}\,\overline{u}\,\overline{x} = \overline{w}\,\overline{v}$, hence $wv \equiv x\,(P_{C^*})$ and $x \in C^*$. Since C is a code and $ux = v$, then $x = 1$. We obtain that C is a maximal prefix code. It remains to show that C is a maximal prefix code. For any $y \in X^*$, since $syn(C^*)$ is a group, there exists $y' \in X^*$ such that $\overline{y'}\,\overline{y} = \overline{y}\,\overline{y'} = \overline{1}$. By $1 \in C^*$, $y'y$, $yy' \in C^*$. Thus y is a prefix of some word in C or some word in C is a prefix of y. Hence $C \cup \{y\}$ is not a prefix code. That is, C is a maximal prefix code. Similarly, we can prove that C is a maximal suffix code.

Lemma 4. *Let $L \subseteq X^*$ be a language such that $L = \overline{1}$ and such that $syn(L)$ is a group with zero. If X is the least alphabet for L, then $syn(L) = \{\overline{1}\}$ and $L = X^*$.*

Proof. Case (1). Let $\overline{1}$ be the zero of $syn(L)$. As $\overline{1}$ is also the identity of $syn(L)$, it follows that $syn(L)$ is trivial, hence $L = X^*$.

Case (2). Let $\overline{1}$ be not the zero of $syn(L)$. Therefore there exists $y \in X^+$ such that $\overline{y}$ is the zero of $syn(L)$. By the hypothesis, X is the least alphabet of L, then $\overline{a} \neq \overline{y}$ for any $a \in X$. Otherwise, if there exists $a \in X$ such that $\overline{a} = \overline{y}$, then there exist $u, v \in X^*$ such that $uav \in L$. But $\overline{uav} = \overline{u}\,\overline{a}\,\overline{v} = \overline{u}\,\overline{y}\,\overline{v} = \overline{y}$ and $\overline{uav} = \overline{1}$, hence $\overline{y} = \overline{1}$, a contradiction. Now, let $y = a_1 \ldots a_m$, $a_i \in X$, $i = 1, \ldots, m$, $m \geq 2$. By $\overline{a}_i \neq \overline{y}$, there exist $x_1, x_2, \ldots, x_{m-1} \in X^*$ such that $\overline{x}_1\overline{a}_1 = \overline{1}$, $\overline{x}_2\overline{a}_2 = \overline{1}, \ldots, \overline{x}_{m-1}\overline{a}_{m-1} = \overline{1}$, thus $\overline{a}_m = \overline{1}\,\overline{a}_m = \overline{x}_{m-1}\overline{a}_{m-1}\overline{a}_m = \overline{x}_{m-1}\overline{1}\,\overline{a}_{m-1}\overline{a}_m = \overline{x}_{m-1}\overline{x}_{m-2}\overline{a}_{m-2}\overline{a}_{m-1}\overline{a}_m = \overline{x}_{m-1}\overline{x}_{m-2}\ldots\overline{x}_1\overline{a}_1\overline{a}_2\ldots\overline{a}_m = \overline{x}_{m-1}\ldots\overline{x}_1\overline{y} = \overline{y}$, in contradiction with $\overline{a}_m \neq \overline{y}$. Hence $\overline{1}$ is the zero of $syn(L)$.

From Case (1) it follows that if $syn(L)$ is a group with zero, then $syn(L) = \{\overline{1}\}$ and $L = X^*$.

Corollary 5. *Let $C \subseteq X^*$ be a code. If C^* is reflective, then $syn(C^*)$ is a group.*

Proof. According to Proposition 2.1 in [6], it is known that $C^* = \overline{1}$ and $syn(C^*)$ is a group or a group with zero. As X is the least alphabet for C^*, it follows from Lemma 4 that $syn(C^*)$ is a group.

Lemma 5. (See [4], Remark 4.5.) *If C is a maximal bifix code, then C is a maximal prefix code or a maximal suffix code.*

Proof of Theorem 2. According to Figures 1 and 2 in the first section, it suffices to show the claim for k-infix or k-outfix codes.

By Lemma 3, $(1) \longrightarrow (5)$. By Lemmas 1 and 5, $(5) \longrightarrow (4)$, $(4) \longrightarrow (3)$. Clearly, $(3) \longrightarrow (2)$, $(2) \longrightarrow (1)$, namely, $(1) \longleftrightarrow (2) \longleftrightarrow (3) \longleftrightarrow (4) \longleftrightarrow (5)$. Again by Corollary 5, $(6) \longrightarrow (1)$. But $(1) \longrightarrow (3)$, $(3) \longrightarrow (6)$. Thus $(1) \longleftrightarrow (2) \longleftrightarrow (3) \longleftrightarrow (4) \longleftrightarrow (5) \longleftrightarrow (6)$.

Lemma 6. (See [6].) *Let C be a code. Then C is a full uniform code if and only if C and C^* are reflective.*

Lemma 7. *Let $C \subseteq X^+$ be a p-infix or an s-infix code. If C^* is reflective, then C is also reflective.*

Proof. Assume C is not reflective, then there exist $x, y \in X^*$, $xy \in X^+$ such that $xy \in C$ and $yx \notin C$. Since C^* is reflective $yx \in C^*$, that is, $yx = c_1c_2 \dots c_m$ for some $c_i \in C$, $i = 1, \dots, m$, $m \geq 2$. Therefore, $y = c_1 \dots c_{r-1}c'_r$, $x = c''_r c_{r+1} \dots c_m$, $c_r = c'_r c''_r$, $1 \leq r \leq m$. But $xy = c''_r c_{r+1} \dots c_m c_1 \dots c_{r-1} c'_r \in C$ and C is a p-infix code, then $c_1 = c_2 = \dots = c_{r-1} = c'_r = 1$. By $c'_r c''_r \neq 1$, $c''_r \neq 1$. Again $xy = c''_r c_{r+1} \dots c_m \in C$, $X^+C \cap C \neq \emptyset$. But according to Figure 2, we know that a p-infix code must be a bifix code, thus $X^+C \cap C = \emptyset$, a contradiction. In a similar fashion, we can prove the case for an s-infix code.

Lemma 8. *Let $C \subseteq X^+$ be a p-infix or a right semaphore code. Then C is an infix code if and only if C is a suffix code.*

Proof. This is an immediate consequence of the definitions of p-infix and right semaphore codes.

On a similar plan, we have

Lemma 9. *Let $C \subseteq X^+$ be an s-infix or a left semaphore code. Then C is an infix code if and only if C is a prefix code.*

Proof of Theorem 3. By Lemmas 3 and 8, we know if $syn(C^*)$ is a group, then C is an infix code. Thus by Theorem 2, (1) $\longleftrightarrow$ (2) $\longleftrightarrow$ (3). By Lemmas 6 and 7, (3) $\longleftrightarrow$ (5). Also by Lemma 8 and Theorem 2, it is easy to show that (3) $\longleftrightarrow$ (4). Hence (1) $\longleftrightarrow$ (2) $\longleftrightarrow$ (3) $\longleftrightarrow$ (4) $\longleftrightarrow$ (5). Similarly, Corollary 4 is proved.

5. Two remarks on group codes and maximal codes

By Theorem 1, we know that if C is a finite group code, then it is a maximal code. In general, we have

Remark 4. Let $C \subseteq X^+$ be a code. If C is a maximal code, then it need not be a group code. But a group code must be a maximal code. In fact, let $C \subseteq X^+$ be a group code. If $C^* = X^*$, then C is a maximal code. Without loss of generality, let $w \in X^* \backslash C^*$, we can prove that $C \cup \{w\}$ is not a code. Since w is an invertible element of $syn(C^*)$, there exists $w' \in X^*$ such that $\overline{w'}\,\overline{w} = \overline{w}\,\overline{w'} = \overline{1}$, $w'w \equiv ww' \equiv 1(P_{C^*})$. Thus $u = w'w$, $v = ww' \in C$, and the word $ww'w = wu = vw$ has two distinct factorizations in words in $C \cup \{w\}$. Hence $C \cup \{w\}$ is not a code and is a maximal code.

Remark 5. By Lemma 3, a group code is a maximal prefix, a maximal suffix and a maximal bifix code. Conversely, a maximal prefix, a maximal suffix, or a maximal bifix code is not necessarily a group code. For example, let $X = \{a, b\}$ and $C = \{ba, a^3, a^2b, ab^2, b^3, aba^2, abab, b^2a^2, b^2ab\}$. It is easy to verify that C is a prefix, a suffix and a bifix code. By Theorem 4.5 in [9], C is a maximal code. Thus C is a maximal prefix, a maximal suffix and a maximal bifix code. By Theorem 1, clearly, C is not a group code.

References

[1] Berstel, J., D. Perrin, *Theory of Codes*, Academic Press, New York 1985.

[2] Dongyang, Long, *k-Outfix Codes*, Chinese Annals of Math., Vol. 10, Ser. A (1989), 94–99 (in Chinese).

[3] Dongyang, Long, *k-Prefix Codes and k-Infix Codes*, Acta Mathematical Sinica, Vol. 33 (1990), 414–421 (in Chinese).

[4] Ito, M., H. Jurgensen, H. J. Shyr, G. Theirrin, *Outfix and Infix Codes and Related Classes of Languages*, J. Comput. Systems Sci., to appear.

[5] Lallement, G., *Semigroups and Combinatorial Applications*, John Wiley & Sons, New York, 1979.

[6] Reis, C. M., G. Therrin, *Reflective Star Languages and Codes*, Information and Control **42** (1979), 1–9.

[7] Shyr, H. J., *Free Monoids and Languages*, Soochow University Taipei, Taiwan, 1979.

[8] Schützenberger, M. P., *Une théorie algèbrique du codage*, C. R. Acad. Sci. Paris **242**, 862–864, Séminaire Dubriel-pisot, 1955–1956, No. 15, 1956.

[9] Salomaa, A., *Jewels of Formal Language Theory*, Computer Sci. Press, 1981.

Department of Mathematics
Zhongshan University
Guang zhou, 510275
People's Republic of china

Received June 28, 1990
and in final form November 25, 1991

Semigroup Forum Vol. 45 (1992) 45–54

RESEARCH ARTICLE

On a Wreath Product Embedding and Idempotent Pure Congruences on Inverse Semigroups

Bernd Billhardt

Communicated by Boris Schein

1. Introduction

A classical theorem of group theory says that any extension of a group A by a group G can be embedded in their standard wreath product $A \operatorname{Wr} G$. We vary the notion of wreath product to obtain a semigroup theoretic version of this theorem, and generalize a result of [1]. This gives a partial solution of problem VI. 6. 11 (ii) raised in Petrich [6]. For another approach see Houghton [2].
Then we apply our wreath product embedding to idempotent pure congruences on inverse semigroups. A congruence ρ on an inverse semigroup S is called idempotent pure if $a\rho b$ and $a \in E_S := \{e \in S \mid e^2 = e\}$ imply that $b \in E_S$. This leads, given an idempotent pure congruence ρ on S, to a representation of S as a certain subsemigroup of an inverse semigroup W, which depends on a semilattice R and S/ρ (Theorem 10). Further W can be embedded into a semidirect product of a semilatice, and and S/ρ (Corollary 13). A related representation was given by O'Carroll [4]. Both representations include the McAlister P-theorem [3]. For a succint proof of the P-theorem see Schein [7].

The notation and terminology of Petrich [6] will be used, and the basic results on inverse semigroups (particularly concerning idempotent pure congruences) contained therein will be assumed. Any order-theoretic statement made about an inverse semigroup refers to the natural order . For $A, B \subseteq S$ $AB := \{ab \mid a \in A, b \in B\}$.

2. Embedding in the Standard λ-Wreath Product

Proposition 1. *Let A and G be inverse semigroups, such that G acts on A by endomorphisms on the left, i.e., for each $g \in G$ $\alpha \mapsto g\alpha$ is an endomorphism of A, and $h(g\alpha) = (hg)\alpha$ for all $h, g \in G$, $\alpha \in A$. On $S := \{(\alpha, g) \in A \times G \mid gg^{-1}\alpha = \alpha\}$ let a multiplication be defined by $(\alpha, g)\cdot(\beta, h) := \big((gh)(gh)^{-1}\alpha \cdot g\beta, gh\big)$. Then S is an inverse semigroup with $(\alpha, g)^{-1} := (g^{-1}\alpha^{-1}, g^{-1})$. S is called a* λ**-semidirect product of A and G.**

Proof. Let $(\alpha, g), (\beta, h), (\gamma, k) \in S$. Then

$$\begin{aligned}(gh)(gh)^{-1}\left((gh)(gh)^{-1}\alpha \cdot g\beta\right) &= (gh)(gh)^{-1}\alpha \cdot ghh^{-1}g^{-1}g\beta \\ &= (gh)(gh)^{-1}\alpha \cdot ghh^{-1}\beta \\ &= (gh)(gh)^{-1}\alpha \cdot g\beta .\end{aligned}$$

Therefore the multiplication is well-defined. Also,

$$\begin{aligned}((a,g)(\beta,h))\,(\gamma,k) &= (\alpha,g)\,((\beta,h)(\gamma,k))\\ ((a,g)(\beta,h))\,(\gamma,k) &= (ghkk^{-1}h^{-1}(g^{-1}g)(hh^{-1})g^{-1}\alpha\\ &\quad \cdot g(hkk^{-1}h^{-1})(g^{-1}g)\beta\cdot gh\gamma, ghk)\\ &= (ghkk^{-1}h^{-1}g^{-1}\alpha\cdot ghkk^{-1}h^{-1}\beta\cdot gh\gamma, ghk)\\ &= (\alpha,g)\,((\beta,h)(\gamma,k)).\end{aligned}$$

Further

$$\begin{aligned}(\alpha,g)(g^{-1}\alpha^{-1},g^{-1})(\alpha,g) &= (\alpha\cdot gg^{-1}\alpha^{-1}\cdot\alpha,g)\\ &= (\alpha\cdot(gg^{-1}\alpha)^{-1}\cdot\alpha,g)\\ &= (\alpha,g),\end{aligned}$$

and

$$\begin{aligned}(g^{-1}\alpha^{-1},g^{-1})(\alpha,g)(g^{-1}\alpha^{-1},g^{-1}) &= (g^{-1}\alpha^{-1}\cdot g^{-1}\alpha\cdot g^{-1}\alpha^{-1},g^{-1}]\\ &= (g^{-1}\alpha^{-1},g^{-1}).\end{aligned}$$

Finally, since $(\delta,\ell)\in E_S \leftrightarrow \delta\in E_A, \ell\in E_G, \ell\delta=\delta$, idempotents commute. ■

An important example of the above construction is the following:

Definition 2. Let A, G be inverse semigroups, $F := A^G$. Then F is an inverse semigroup with respect to the multiplication given by $(x)(fg) := (x)f\cdot(x)g$, $x\in G$, $f,g\in F$, and $(x)f^{-1} := ((x)f)^{-1}$, $f\in F$, $x\in G$. For $t\in G$, $f\in F$, let $tf\in F$ be defined by $(x)(tf) := (xt)f$, $x\in G$, then $f\mapsto tf$ defines an action on F by endomorphisms on the left. Thus we can form the λ-semidirect product W of A^G and G by this action. W is called the **standard λ-wreath product of A and G** ($A\,\mathrm{Wr}^\lambda\, G$).

The following theorem is a generalization of Theorem 3 of [1]:

Theorem 3. *Let S be an inverse semigroup, further ρ a congruence on S, such that for each ρ-class $s\rho$, $s\in S$, $\iota((s\rho)^{-1}) := \{x^{-1}x \mid x\in s\rho\}$ contains a greatest element with respect to the natural order. Then S can be embedded in* $\operatorname{Ker}\rho\,\mathrm{Wr}^\lambda\, S/\rho$ *where* $\operatorname{Ker}\rho = \{s\in S \mid s\rho e,$ *for some* $e\in E_S\}$.

Proof. For each $\overline{s}\in S/\rho$ let $\overline{s}_r\in\overline{s}$, such that $\overline{s}_r^{-1}\overline{s}_r$ is the greatest element of $\iota(\overline{s}^{-1})$. Obviously $\operatorname{Ker}\rho$ is an inverse subsemigroup of S. Consider $\phi_r : S \longrightarrow \operatorname{Ker}\rho\,\mathrm{Wr}^\lambda\, S/\rho$, defined by $s\mapsto (f_s,\overline{s})$, with $(\overline{u})f_s := \overline{uss^{-1}}_r s\overline{us}_r^{-1}, \overline{u}\in S/\rho$. Then ϕ_r is a mapping of S into $\operatorname{Ker}\rho\,\mathrm{Wr}^\lambda\, S/\rho$: since $(\overline{u})f_s\rho us(us)^{-1}$ for each $\overline{u}\in S/\rho$, $f_s\in \operatorname{Ker}\rho^{S/\rho}$. Further $(\overline{u})\overline{ss^{-1}}f_s = (\overline{uss^{-1}})f_s = (\overline{u})f_s$ for each $\overline{u}\in S/\rho$, hence

ϕ_r maps S into $\operatorname{Ker}\rho \operatorname{Wr}^\lambda S/_\rho$. Since

$$
\begin{aligned}
(f_s, \overline{s}) &= (f_t, \overline{t}) \\
&\Rightarrow s^{-1}s \leq \overline{s}_r^{-1}\overline{s}_r = \overline{t}_r^{-1}\overline{t}_r \geq t^{-1}t\,, \\
&\quad \overline{s} = \overline{t}, \text{ and } (\overline{ss^{-1}})f_s = (\overline{tt^{-1}})f_t \\
&\Rightarrow s = \overline{ss^{-1}}_r^{-1}\overline{ss^{-1}}_r ss^{-1}ss^{-1}s\overline{s}_r^{-1}\overline{s}_r \\
&\quad = \overline{ss^{-1}}_r^{-1}\overline{ss^{-1}}_r s\overline{s}_r^{-1}\overline{s}_r \\
&\quad = \overline{ss^{-1}}_r^{-1}(\overline{ss^{-1}})f_s\overline{s}_r \\
&\quad = \overline{tt^{-1}}_r^{-1}(\overline{tt^{-1}})f_t\overline{t}_r = t\,,
\end{aligned}
$$

the mapping ϕ_r is injctive. We now prove that ϕ_r is a homomorphism:

$$
\begin{aligned}
(\overline{u})\left((stt^{-1}s^{-1}f_s)\cdot(sf_t)\right) &= \overline{ustt^{-1}s^{-1}}_r s\overline{ustt^{-1}}_r^{-1}\overline{ustt^{-1}}_r t\overline{ust}_r^{-1} \\
&= \overline{ustt^{-1}s^{-1}}_r s(\overline{ustt^{-1}s^{-1}}_r s)^{-1}\underbrace{(\overline{ustt^{-1}s^{-1}}_r s)}_{a}\overline{ustt^{-1}}_r^{-1}\overline{ustt^{-1}}_r t\overline{ust}_r^{-1} \\
&= \overline{ustt^{-1}s^{-1}}_r st\overline{ust}_r^{-1}, \text{ since } a\rho\overline{ustt^{-1}}_r \Rightarrow a^{-1}a \leq \overline{ustt^{-1}}_r^{-1}\overline{ustt^{-1}}_r \\
&= (\overline{u})f_{st}\,,
\end{aligned}
$$

and clearly $$\overline{s}\overline{t} = \overline{st}\,.$$ ■

3. An Application to Idempotent Pure Congruences

Let S be an inverse semigroup, ρ an idempotent pure congruence on S. Let $C'(S) := \{\varnothing \neq H \subseteq S \mid H \subseteq s\rho \text{ for some } s \in S,\ (ss^{-1})\rho H(s^{-1}s)\rho \subseteq H\}$.

Theorem 4. *$C'(S)$ is an inverse semigroup with operations defined by $H \circ K := (stt^{-1}s^{-1})\rho HK(t^{-1}s^{-1}st)\rho$, for $H \subseteq s\rho$, $K \subseteq t\rho$, and $H^{-1} := \{h^{-1} \in S \mid h \in H\}$.*

Proof. $H \circ K \in C'(S)$: Obviously $H \circ K$ is uniquely defined, $H \circ K \subseteq (st)\rho$, and

$$
\begin{aligned}
(stt^{-1}s^{-1})\rho H \circ K(t^{-1}s^{-1}st)\rho &= (stt^{-1}s^{-1})\rho(stt^{-1}s^{-1})\rho HK \\
&\quad \cdot (t^{-1}s^{-1}st)\rho(t^{-1}s^{-1}st)\rho \\
&\subseteq (stt^{-1}s^{-1})\rho HK(t^{-1}s^{-1}st)\rho \\
&= H \circ K\,.
\end{aligned}
$$

We next prove that $\circ$ is associative. Let $H \subseteq s\rho$, $K \subseteq t\rho$, $L \subseteq u\rho$. We get:

$$
\begin{aligned}
(H \circ K) \circ L &= ((stt^{-1}s^{-1})\rho HK(t^{-1}s^{-1}st)\rho) \circ L \\
&= (stuu^{-1}t^{-1}s^{-1})\rho((stt^{-1}s^{-1})\rho HK(t^{-1}s^{-1}st)\rho L) \\
&\quad (u^{-1}t^{-1}s^{-1}stu)\rho\,, \qquad \text{thus} \\
v \in (H \circ K) \circ L &\Rightarrow v = efhkg\ell d
\end{aligned}
$$

with

$$\begin{aligned}
&e \in (stuu^{-1}t^{-1}s^{-1})\rho\,,\\
&f \in (stt^{-1}s^{-1})\rho\,, h \in H, k \in K, \ell \in L,\\
&g \in (t^{-1}s^{-1}st)\rho\,, d \in (u^{-1}t^{-1}s^{-1}stu)\rho\\
&\quad = (ef)(hk\ell)(\ell^{-1}g\ell d) \in (stuu^{-1}t^{-1}s^{-1})\rho HKL(u^{-1}t^{-1}s^{-1}stu)\rho\\
&\quad =: A\,;
\end{aligned}$$

on the other hand for $h \in H$, $k \in K$, $\ell \in L$ we have

$$\begin{aligned}
hkl &= (hk)(hk)^{-1}hk(hk)^{-1}(hk)\ell\\
&\in (stt^{-1}s^{-1})\rho HK(t^{-1}s^{-1}st)\rho L\,,
\end{aligned}$$

implying $A \subseteq (H \circ K) \circ L$, thus $(H \circ K) \circ L = A$, and analogously follows $H \circ (K \circ L) = A$. Further $H^{-1} \in C'(S)$, since $H^{-1} \subseteq s^{-1}\rho$ and $(s^{-1}s)\rho H^{-1}(ss^{-1})\rho = ((ss^{-1})\rho H(s^{-1}s)\rho)^{-1} \subseteq H^{-1}$. Also $H \circ H^{-1} \circ H = (ss^{-1}\rho HH^{-1}H(s^{-1}s)\rho \subseteq (ss^{-1})\rho H(s^{-1}s)\rho \subseteq H \subseteq H \circ H^{-1} \circ H$. ∎

Obviously S is embedded in $C'(S)$ via $s \mapsto \hat{s} := \{esf \mid e\rho ss^{-1}, f\rho s^{-1}s\}$, since $\hat{s} \in C'(S)$ for all $s \in S$ by definition and $\hat{s} = \hat{t} \Rightarrow s \le t \le s$, with respect to the natural order on S.

We define a relation on $C'(S)$ by: $H\rho_{C'(S)}K \Leftrightarrow H$, $K \subseteq s\rho$, for some $s \in S$. $\rho_{C'(S)}$ is an idempotent pure congruence on $C'(S)$. Moreover we obtain:

Lemma 5. (i) *each $\rho_{C'(S)}$-class contains precisely one ρ-class.*

(ii) *for each $\rho_{C'(S)}$-class $H\rho_{C'(S)}$, $\iota((H\rho_{C'(S)})^{-1})$ contains a greatest element with respect to the natural order on $C'(S)$.*

Proof. (i) clear by definition of $\rho_{C'(S)}$, and the fact that $s\rho \in C'(S)$ for each $s \in S$.

(ii) $s\rho$ is the greatest element of its $\rho_{C'(S)}$-class

$$\begin{aligned}
K \in (s\rho)\rho_{C'(S)} \Rightarrow K \circ K^{-1} \circ s\rho &= (ss^{-1})\rho KK^{-1}s\rho(s^{-1}s)\rho\\
&\subseteq (ss^{-1})\rho K(s^{-1}s)\rho \subseteq K\\
&\subseteq K \circ K^{-1} \circ s\rho\,,\\
\Rightarrow K \circ K^{-1} \circ s\rho &= K\,,\\
\Rightarrow K &\le s\rho\,,
\end{aligned}$$

thus $(s\rho)^{-1} \circ (s\rho)$ is the greatest element of $\iota\left(((s\rho)\rho_{C'(S)})^{-1}\right)$. ∎

The above lemma implies $S/_{\rho} \cong C'(S)/\rho_{C'(S)}$, and by Theorem 3 $C'(S)$ is embedded in $E_{C'(S)}\,\mathrm{Wr}^{\lambda}\, S/_{\rho}$, hence we get:

Corollary 6. *Let S be an inverse semigroup, ρ be an idempotent pure congruence on S, the ρ-class of $x \in S$ being denoted by $\overline{x}$. Then*

$$\begin{aligned}
&\phi : S \longrightarrow E_{C'(S)}\,\mathrm{Wr}^{\lambda}\, S/_{\rho}\,, \text{ given by}\\
&\quad s \longmapsto (f_s, \overline{s}), \text{ with } (\overline{u})f_s := \overline{uss^{-1}} \circ \hat{s} \circ \overline{(us)^{-1}}\,,\ \overline{u} \in S/_{\rho}\\
&\qquad\qquad\qquad\qquad\qquad\quad = \overline{uss^{-1}}\hat{s}\overline{(us)^{-1}}\,,
\end{aligned}$$

is an embedding.

Proof. Note that

$$\overline{uss^{-1}} \circ \hat{s} \circ \overline{(us)^{-1}} = \overline{uss^{-1}u^{-1}}\,\overline{uss^{-1}\hat{s}(us)^{-1}}\,\overline{uss^{-1}u^{-1}} = \overline{uss^{-1}\hat{s}(us)^{-1}}. \quad \blacksquare$$

Corollary 7. *Let S be an inverse semigroup, ρ be an idempotent pure congruence on S. Then S is embeddable in a λ-semidirect product of a semilattice and $S/_\rho$.*

In what follows, for an idempotent pure congruence ρ on an inverse semigroup S, we give a representation of S as a certain subsemigroup of a λ-semidirect product of a semilattice and $S/_\rho$. For this we need the following construction:

Construction 8. Let $(R, \wedge)$ be a semilattice, and suppose that G is an inverse semigroup acting on R by endomorphisms on the left. Let T be a nonempty subset of R such that

(i) for each $\alpha \in T$ there is $e_\alpha \in E_G$ with $gg^{-1}\alpha = \alpha \Leftrightarrow gg^{-1} \geq e_\alpha$, $g \in G$.

(ii) for each $g \in G$ there is $\alpha \in T$ with $g^{-1}\alpha \in T$ and $gg^{-1} = e_\alpha$.

(iii) $\alpha, \beta \in T \Rightarrow e_\alpha\beta \wedge e_\beta\alpha \in T$, and $e_{e_\alpha\beta\wedge e_\beta\alpha} = e_\alpha e_\beta$.

(iv) $g\gamma \leq ghh^{-1}g^{-1}\alpha \Rightarrow g\gamma \in T$, for all $\alpha, \gamma \in T$, $g, h \in G$, with $gg^{-1}\alpha = \alpha$, $hh^{-1}\gamma = \gamma$, $g^{-1}g\gamma = \gamma$.

Define $\overline{T}$ by $\overline{T} := \{(\alpha, g) \mid \alpha, g^{-1}\alpha \in T, gg^{-1} = e_\alpha\}$ under the multiplication $(\alpha, g) \cdot (\beta, h) := (ge_\beta g^{-1}\alpha \wedge g\beta, gh)$.

Proposition 9. *$\overline{T}$ is an inverse semigroup. If $(\alpha, g) \in \overline{T}$, then $(\alpha, g)^{-1} = (g^{-1}\alpha, g^{-1})$. The set of idempotents of $\overline{T}$ is given by $E_{\overline{T}} = \{(\alpha, e_\alpha) \mid \alpha \in T\}$. The relation ρ which is given by $(\alpha, g)\rho(\beta, h) :\Leftrightarrow g = h$ defines an idempotent pure congruence on $\overline{T}$ with $\overline{T}/_\rho \cong G$.*

Proof. First we prove

$$(*) \qquad k^{-1}\varepsilon \in T \text{ and } kk^{-1}\varepsilon = \varepsilon \Rightarrow k^{-1}e_\varepsilon k = e_{k^{-1}\varepsilon}, \quad k \in G, \varepsilon \in T.$$

Let $k^{-1}\varepsilon \in T$ and $kk^{-1}\varepsilon = \varepsilon$. It follows $k^{-1}e_\varepsilon kk^{-1}\varepsilon = k^{-1}\varepsilon$. Further for $f \in E_G$ with $fk^{-1}\varepsilon = k^{-1}\varepsilon$ we obtain $kfk^{-1}\varepsilon = kk^{-1}\varepsilon = \varepsilon$, implying $e_\varepsilon \leq kfk^{-1}$, implying $k^{-1}e_\varepsilon k \leq k^{-1}kf \leq f$. Thus $(*)$ is proven.

Let now $(\alpha, g), (\beta, h) \in \overline{T}$. It follows $g^{-1}\alpha$, $h^{-1}\beta \in T$, $e_\alpha = gg^{-1}$, $e_\beta = hh^{-1}$, and $e_{g^{-1}\alpha} = g^{-1}g$ by $(*)$. We show that $(ge_\beta g^{-1}\alpha \wedge g\beta, gh) \in \overline{T}$.
$\gamma := ge_\beta g^{-1}\alpha \wedge g\beta = g(e_\beta g^{-1}\alpha \wedge g^{-1}g\beta) = g(e_\beta g^{-1}\alpha \wedge e_{g^{-1}\alpha}\beta) = g\delta$ with $\delta := e_\beta g^{-1}\alpha \wedge e_{g^{-1}\alpha}\beta \in T$ by (iii). Further $g\delta \leq ge_\beta g^{-1}\alpha$ with $gg^{-1}\alpha = \alpha$, $e_\beta\delta = \delta$, $g^{-1}g\delta = \delta$, which implies $g\delta \in T$ by (iv).
Moreover $(gh)^{-1}\gamma = h^{-1}g^{-1}\gamma = h^{-1}\delta \leq h^{-1}e_{g^{-1}\alpha}\beta = h^{-1}e_{g^{-1}\alpha}h(h^{-1}\beta)$ with $h^{-1}h(h^{-1}\beta) = h^{-1}\beta$, $e_{g^{-1}\alpha}\delta = \delta$, $hh^{-1}\delta = \delta$, which implies $h^{-1}\delta \in T$ by (iv).

Next we obtain

$$\begin{aligned} e_\gamma = e_{g\delta} &\overset{(*)}{=} ge_\delta g^{-1} \\ &= ge_{e_\beta g^{-1}\alpha \wedge e_{g^{-1}\alpha}\beta}g^{-1} = ge_{g^{-1}\alpha}e_\beta g^{-1}, \text{ by (iii)} \\ &= gg^{-1}ghh^{-1}g^{-1} = (gh)(gh)^{-1}. \end{aligned}$$

Finally for $(\alpha, g) \in \overline{T}$ we have $(\alpha, g)^{-1} = (g^{-1}\alpha, g^{-1}) \in \overline{T}$, since $g^{-1}g = e_{g^{-1}\alpha}$. The rest of the proof follows from the fact that $\overline{T}$ is a subsemigroup of the λ-semidirect product of R and G, defined by the action of G on R, and (ii). ■

After this preparation we prove the desired representation theorem.

Theorem 10. *Let S be an inverse semigroup, ρ be an idempotent pure congruence on S. Then S is isomorphic to a semigroup obtained by Construction 8, with $G = S/_\rho$.*

Proof. Let $\phi : S \to E_{C'(S)} \operatorname{Wr}^\lambda S/_\rho$ as in Corollary 6. We show that $\overline{T} := S\phi$ satisfies the conditions of Construction 8 with $R := E_{C'(S)}^{S/\rho}$, $G := S/_\rho$, and $T := \{f_{ss^{-1}} \in E_{C'(S)}^{S/\rho} \mid s \in S\}$. Note that from the embedding $\phi : s \mapsto (f_s, \overline{s})$ we directly obtain $f_s = f_{ss^{-1}}, s \in S$, since $(f_{ss^{-1}}, \overline{ss^{-1}}) = (ss^{-1})\phi = (s\phi)(s\phi)^{-1} = (f_s, \overline{s})(\overline{s}^{-1}f_s, \overline{s}^{-1}) = (f_s, \overline{ss^{-1}})$.

(1) : Let $e_{f_{ss^{-1}}} := \overline{ss^{-1}}$ for $f_{ss^{-1}} \in T$. Then since
$f_{ss^{-1}} = f_{tt^{-1}} \Rightarrow \widehat{ss^{-1}} = (\overline{ss^{-1}})f_{ss^{-1}} = (\overline{ss^{-1}})f_{tt^{-1}}$ and $(\overline{tt^{-1}})f_{ss^{-1}} = (\overline{tt^{-1}})f_{tt^{-1}} = \widehat{tt^{-1}} \Rightarrow tt^{-1} = ss^{-1}$,
$e_{f_{ss^{-1}}}$ is uniquely defined. Further $\overline{tt^{-1}}f_{ss^{-1}} = f_{ss^{-1}} \Rightarrow (\overline{ss^{-1}})(\overline{tt^{-1}}f_{ss^{-1}}) = (\overline{ss^{-1}})f_{ss^{-1}} \Rightarrow \overline{ss^{-1}tt^{-1}} = \overline{ss^{-1}}$, thus (i) is established.

(2) : Let $\overline{s} \in G$. Then $(f_{s^{-1}s}, \overline{s}^{-1}) = s^{-1}\phi = (s\phi)^{-1} = (\overline{s}^{-1}f_{ss^{-1}}, \overline{s}^{-1})$, thus $\overline{s}^{-1}f_{ss^{-1}} \in T$, and $e_{f_{ss^{-1}}} = \overline{ss^{-1}}$, establishing (ii).

(3) : $f_{ss^{-1}}, f_{tt^{-1}} \in T \Rightarrow (f_{ss^{-1}tt^{-1}}, \overline{ss^{-1}tt^{-1}}) = (ss^{-1}tt^{-1})\phi = (ss^{-1})\phi \cdot (tt^{-1})\phi = (f_{ss^{-1}}, \overline{ss^{-1}})(f_{tt^{-1}}, \overline{tt^{-1}}) = (\overline{tt^{-1}}f_{ss^{-1}} \cdot \overline{ss^{-1}}f_{tt^{-1}}, \overline{ss^{-1}tt^{-1}})$, establishing (iii).

(4) : We prove statement (iv).
For this, let $\overline{g}f_{ss^{-1}} \leq \overline{ghh^{-1}g^{-1}}f_{tt^{-1}}$, $\overline{gg^{-1}}f_{tt^{-1}} = f_{tt^{-1}}, \overline{hh^{-1}}f_{ss^{-1}} = f_{ss^{-1}}$, and $\overline{g^{-1}g}f_{ss^{-1}} = f_{ss^{-1}}$. It follows $\overline{gg^{-1}tt^{-1}} = \overline{tt^{-1}}$, $\overline{hh^{-1}ss^{-1}} = \overline{ss^{-1}}$ and $\overline{g^{-1}gss^{-1}} = \overline{ss^{-1}}$.
Now since $\overline{g}f_{ss^{-1}} \leq \overline{ghh^{-1}g^{-1}}f_{tt^{-1}}$,we obtain:

$$
\begin{aligned}
\widehat{ss^{-1}} = (\overline{g}^{-1})\overline{g}f_{ss^{-1}} &\leq (\overline{g}^{-1})(\overline{ghh^{-1}g^{-1}})f_{tt^{-1}} \\
&= (\overline{hh^{-1}g^{-1}})f_{tt^{-1}} \\
&= \overline{hh^{-1}g^{-1}tt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}ghh^{-1}}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \widehat{ss^{-1}} &= \widehat{ss^{-1}} \circ \overline{hh^{-1}g^{-1}tt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}ghh^{-1}} \\
&= \overline{ss^{-1}g^{-1}tt^{-1}gss^{-1}}\,\widehat{}\,\overline{hh^{-1}g^{-1}tt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}ghh^{-1}}\;\overline{ss^{-1}g^{-1}tt^{-1}g} \\
&\subseteq \overline{ss^{-1}g^{-1}(tt^{-1})(gss^{-1}hh^{-1}g^{-1})tt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}g(hh^{-1}ss^{-1})(g^{-1}tt^{-1}g)} \\
&= \overline{ss^{-1}g^{-1}(gss^{-1}hh^{-1}g^{-1})tt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}g(g^{-1}tt^{-1}g)hh^{-1}ss^{-1}} \\
&= \overline{ss^{-1}g^{-1}tt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}gss^{-1}},
\end{aligned}
$$

$\Rightarrow ss^{-1} = u^{-1}tt^{-1}ev$, for suitable $u, v \in \overline{tt^{-1}gss^{-1}}$, $e \in \overline{tt^{-1}}$.

We show that $\overline{g}f_{ss^{-1}} = f_{uss^{-1}v^{-1}}$. Note that $uss^{-1}v^{-1} \in E_S$, since $u\rho v$, and ρ is idempotent pure. We have to prove:

$$\begin{aligned} L :=\ & \overline{xgss^{-1}}\,\widehat{ss^{-1}}\,\overline{ss^{-1}g^{-1}x^{-1}} \\ =\ & \overline{xuss^{-1}v^{-1}}\,\widehat{uss^{-1}v^{-1}}\,\overline{vss^{-1}u^{-1}x^{-1}} \\ =:\ & R \text{ for all } x \in S\,. \end{aligned}$$

$$\begin{aligned} a \in L \Rightarrow a = pss^{-1}dq^{-1} & \text{ with } p,q \in \overline{xgss^{-1}}, d \in \overline{ss^{-1}} \\ & \leq (pu^{-1})(uss^{-1}v^{-1})(vdq^{-1}) \in R, \qquad \text{since} \end{aligned}$$

$\overline{pu^{-1}} = \overline{xgss^{-1}g^{-1}tt^{-1}} = \overline{xtt^{-1}gss^{-1}g^{-1}tt^{-1}} = \overline{xuss^{-1}v^{-1}}$, and

$\overline{vdq^{-1}} = \overline{tt^{-1}gss^{-1}g^{-1}x^{-1}} = \overline{tt^{-1}gss^{-1}g^{-1}tt^{-1}x^{-1}} = \overline{vss^{-1}u^{-1}x^{-1}}$. Further,

$$\begin{aligned} a \in R \Rightarrow a = (pu)(ss^{-1})(v^{-1}dq^{-1}), & \text{ with} \\ p,q \in \overline{xuss^{-1}v^{-1}},\ d \in \overline{uss^{-1}v^{-1}}. & \end{aligned}$$

Then

$$(*) \qquad \begin{aligned} \overline{gss^{-1}} &= \overline{gu^{-1}tt^{-1}ev} = \overline{(gss^{-1}g^{-1})(tt^{-1})gss^{-1}} \\ &= \overline{(tt^{-1})(gss^{-1}g^{-1})gss^{-1}} = \overline{tt^{-1}gss^{-1}}\,. \end{aligned}$$

Therefore

$$\begin{aligned} \overline{pu} &= \overline{xuss^{-1}v^{-1}u} = \overline{xtt^{-1}(gss^{-1}g^{-1})(tt^{-1})gss^{-1}} \\ &= \overline{x(tt^{-1}(gss^{-1}g^{-1})gss^{-1}} = \overline{xtt^{-1}gss^{-1}} = \overline{xgss^{-1}}\,, \end{aligned}$$

and

$$\begin{aligned} \overline{v^{-1}dq^{-1}} &= \overline{ss^{-1}g^{-1}tt^{-1}gss^{-1}g^{-1}tt^{-1}gss^{-1}g^{-1}tt^{-1}x^{-1}} \\ &= \overline{ss^{-1}g^{-1}(tt^{-1})(gss^{-1}g^{-1})tt^{-1}x^{-1}} \\ &= \overline{ss^{-1}g^{-1}(gss^{-1}g^{-1})(tt^{-1})x^{-1}} \\ &= \overline{ss^{-1}g^{-1}tt^{-1}x^{-1}} = \overline{ss^{-1}g^{-1}x^{-1}}\,, \\ & \text{since } \overline{ss^{-1}g^{-1}tt^{-1}} = \overline{ss^{-1}g^{-1}} \quad \text{by } (*)\,. \end{aligned}$$

Consequently, $a \in L$.

(5) : We shall prove that $\underline{T} := \{(f_s, \overline{s} \mid s \in S\}$ coincides with $\{(f_{ss^{-1}}, \overline{g}) \mid \overline{g}^{-1}f_{ss^{-1}} \in T,\ \overline{gg^{-1}} = \overline{ss^{-1}}\}$.

Let $\overline{g}^{-1}f_{ss^{-1}} = f_{tt^{-1}}$ and $\overline{gg^{-1}} = \overline{ss^{-1}}$. It follows

$\widehat{ss^{-1}} = (\overline{g})(\overline{g}^{-1}f_{ss^{-1}}) = (\overline{g})f_{tt^{-1}} = \overline{gtt^{-1}}\,\widehat{tt^{-1}}\,\overline{tt^{-1}g^{-1}}$

$\Rightarrow ss^{-1} = utt^{-1}ev^{-1}$ for suitable $u,v \in \overline{gtt^{-1}}$, $e \in \overline{tt^{-1}}$.

Let $p := uu^{-1}vtt^{-1}e$. It follows

$$\begin{aligned} (**) \quad \overline{ss^{-1}g} &= \overline{utt^{-1}ev^{-1}g} = \overline{gtt^{-1}g^{-1}g} = \overline{gtt^{-1}}, \text{ thus} \\ pp^{-1} &= (vtt^{-1}ev^{-1})(uu^{-1}) = (uu^{-1})(vtt^{-1}ev^{-1}) \\ &= uu^{-1}u(v^{-1}v)(tt^{-1}e)v^{-1}\,, \ (\text{since } u^{-1}v, uv^{-1} \in E_S \\ & \qquad\qquad\qquad \Rightarrow u^{-1}v = u^{-1}uv^{-1}v) \\ &= u(tt^{-1}e)(v^{-1}v)v^{-1} = utt^{-1}ev^{-1} = ss^{-1}\,, \end{aligned}$$

and

$$\begin{aligned}\overline{p} &= \overline{u(u^{-1}v)(tt^{-1}e)} = \overline{u(tt^{-1}e)(u^{-1}v)} = \overline{(utt^{-1}ev^{-1})vu^{-1}u} \\ &= \overline{ss^{-1}gtt^{-1}g^{-1}gtt^{-1}} = \overline{ss^{-1}gtt^{-1}} \\ &\overset{(**)}{=} \overline{ss^{-1}g} = \overline{g}\,, \text{ hence } (f_{ss^{-1}}, \overline{g}) \in \overline{T}\,.\end{aligned}$$ ■

At the end of this section we show that it is no restriction to assume T to be a subsemilattice of R in Construction 8, and to change (iii) by a stronger condition (iii)'.

Corollary 11. *Let $(R', \sqcap)$ be a semilattice, and suppose that G is an inverse semigroup acting on R' by endomorphisms on the left. Let T' be a subsemilattice of R' such that* (i), (ii), (iv) *of Construction 8 hold, and further* (iii)'*: $e_\alpha e_\beta = e_{\alpha\sqcap\beta}$, $\alpha, \beta \in T'$. Let $\overline{T'}$ be defined as in Construction 8. Then $\overline{T'}$ is an inverse semigroup satisfying the statements of Proposition 9.*

Proof. Clear by Proposition 9 since (iii)' implies (iii) of Construction 8:

$$e_\beta\alpha \sqcap e_\alpha\beta = e_\alpha e_\beta(\alpha \sqcap \beta) \overset{\text{(iii)'}}{=} e_{\alpha\sqcap\beta}(\alpha \sqcap \beta) = \alpha \sqcap \beta \in T'\,.$$ ■

Conversely we obtain:

Corollary 12. *Let $\overline{T}$ be as in Construction 8. Then $\overline{T}$ is isomorphic to an inverse semigroup obtained by Corollary 11.*

Proof. Let $R' := \{(u,e) \mid u \in R, e \in E_G, eu = u\}$. For $(u,e), (v,f) \in R'$ let $(u,e) \sqcap (v,f) := (fu \wedge ev, ef)$. Then $(R', \sqcap)$ is a semilattice, and $(E_{\overline{T}}, \cdot)$ is a subsemilattice of $(R', \sqcap)$.
For $g \in G$, $(u,e) \in R'$ let $g(u,e) := (gu, geg^{-1})$. We shall show that $(u,e) \mapsto g(u,e)$ defines a left action by endomorphisms on the left on $(R', \sqcap)$. Let $(u,e), (v,f) \in R'$. It follows:

$$eu = u \Rightarrow geg^{-1}gu = geu = gu \Rightarrow g(u,e) \in R'\,,$$

and

$$\begin{aligned} g((u,e) \sqcap (v,f)) = g(fu \wedge ev, ef) &= (gfu \wedge gev, gefg^{-1}) \\ &= (gfg^{-1}gu \wedge geg^{-1}gv, geg^{-1}gfg^{-1}) \\ &= (gu, geg^{-1}) \sqcap (gv, gfg^{-1}) \\ &= g(u,e) \sqcap g(v,f)\,. \end{aligned}$$

Further

$$\begin{aligned} g(h(u,e)) &= g(hu, heh^{-1}) \\ &= (ghu, gheh^{-1}g^{-1}) = (gh)(u,e)\,. \end{aligned}$$

Let $T' := E_{\overline{T}} = \{(\alpha, e_\alpha) \mid \alpha \in T\}$. We show that for R', T', and the above defined action the conditions of Corollary 11 are satisfied:

(i) : holds with $e_{(\alpha,e_\alpha)} := e_\alpha$

(ii) : holds, since $g^{-1}\alpha \in T$ and $gg^{-1} = e_\alpha \Rightarrow g^{-1}(\alpha, e_\alpha) = (g^{-1}\alpha, g^{-1}e_\alpha g) = (g^{-1}\alpha, e_{g^{-1}\alpha}) \in T'$ and $gg^{-1} = e_{(\alpha,e_\alpha)}$.

(iii)': $e_{(\alpha,e_\alpha)} \cdot e_{(\beta,e_\beta)} = e_\alpha e_\beta = e_{e_\alpha\beta\wedge e_\beta\alpha} = e_{(e_\alpha\beta\wedge e_\beta\alpha, e_\alpha e_\beta)} = e_{(\alpha,e_\alpha)\sqcap(\beta,e_\beta)}$.

(iv) : Let $\sqsubseteq$ denote the partial order of R', given by $\sqcap$; further $g(\gamma, e_\gamma) \sqsubseteq ghh^{-1}g^{-1}(\alpha, e_\alpha)$, with $gg^{-1}(\alpha, e_\alpha) = (\alpha, e_\alpha)$, $hh^{-1}(\gamma, e_\gamma) = (\gamma, e_\gamma)$, $g^{-1}g(\gamma, e_\gamma) = (\gamma, e_\gamma)$. It follows:

$$\begin{aligned}(g\gamma, ge_\gamma g^{-1}) &= (g\gamma, ge_\gamma g^{-1}) \sqcap (ghh^{-1}g^{-1}\alpha, ghh^{-1}g^{-1}e_\alpha)\\ &= (ghh^{-1}g^{-1}e_\alpha g\gamma \wedge ge_\gamma g^{-1}ghh^{-1}g^{-1}\alpha, ge_\gamma g^{-1}ghh^{-1}g^{-1}e_\alpha)\\ &= (e_\alpha g\gamma \wedge ge_\gamma g^{-1}\alpha, ge_\gamma g^{-1}e_\alpha)\\ &\Rightarrow g\gamma \le ge_\gamma g^{-1}\alpha \Rightarrow g\gamma \in T.\end{aligned}$$

Moreover, $ge_\gamma g^{-1} = e_{g\gamma}$, since $g^{-1}g\gamma = \gamma$ which implies $(g\gamma, ge_\gamma g^{-1}) \in T'$.

Hence we can form $\overline{T'}$ by Corollary 11. We shall show that $\psi : \overline{T} \to \overline{T'}$, defined by $(\alpha, g) \mapsto ((\alpha, e_\alpha), g)$ is an isomorphism. Obviously ψ is a bijection. Also ψ is a homomorphism since for $(\alpha, g), (\beta, h) \in \overline{T}$

$$\begin{aligned}(\alpha, g)\psi \cdot (\beta, h)\psi &= ((\alpha, e_\alpha), g)((\beta, e_\beta), h)\\ &= (ge_\beta g^{-1}(\alpha, e_\alpha) \sqcap g(\beta, e_\beta), gh)\\ &= ((ge_\beta g^{-1}\alpha \wedge ge_\beta g^{-1}e_\alpha g\beta, ge_\beta g^{-1}e_\alpha), gh)\\ &= ((ge_\beta g^{-1}\alpha \wedge g\beta, ge_\beta g^{-1}), gh)\\ &= ((\alpha, g)(\beta, h))\psi \qquad \blacksquare\end{aligned}$$

Note that for $(\alpha', g), (\beta', h) \in \overline{T'}$ we have

$$\begin{aligned}(\alpha', g)(\beta', h) &= (ge_{\beta'}g^{-1}\alpha' \sqcap g\beta', gh)\\ &= (ge_{\beta'}g^{-1}g(g^{-1}\alpha' \sqcap \beta'), gh)\\ &= (ge_{\beta'}e_{g^{-1}\alpha'}(g^{-1}\alpha' \sqcap \beta'), gh)\\ &= (g(g^{-1}\alpha' \sqcap \beta'), gh) \qquad \text{(by (iii)')}\\ &= (\alpha' \sqcap g\beta', gh).\end{aligned}$$

Corollary 13. *$\overline{T'}$ is an inverse subsemigroup of a semidirect product of a semilattice and G.*

4. Concluding Remarks.

If ρ is an idempotent pure group congruence on S, then clearly S is E-unitary, i.e. $ae, e \in E_S \Rightarrow a \in E_S$, and ρ is the least group congruence on S. In this case Proposition 9, respectively Theorem 10 produce the McAlister$P$-theorem.

Recall that an inverse semigroup S is called (strongly) E-reflexive, iff $aeb \in E_S \Rightarrow bea \in E_S$ for all $a, b \in S$, $e \in E_S$. O'Carroll has proven in [5] that S is (strongly) E-reflexive, iff there is an idempotent pure Clifford congruence on S. Recall further that if we let A be a semilattice and G be a group in Proposition 1, we obtain the semidirect product of a semilattice and a group. In connection with this, we note:

Corollary 14. *Let S be a γ-semidirect product of a semilattice and a Clifford semigroup. Then S is a strong semilattice of semigroups, each of which is a semidirect product of a semilattice and a group.*

Proof. For (α, g), $(\beta, h) \in S$ let $(\alpha, g)\eta(\beta, h) :\Leftrightarrow gg^{-1} = hh^{-1}$. Obviously η is a semilattice congruence whose classes are semidirect products of semilattices and groups. Let structure homomorphisms $\varphi_{gg^{-1},hh^{-1}} : (\alpha,g)\eta \to (\beta,h)\eta$, $gg^{-1} \geq hh^{-1}$ be defined by $(\gamma,k) \mapsto (hh^{-1}\gamma, hh^{-1}k)$. Then for arbitrary (γ,k), $(\delta,\ell) \in S$:

$$\begin{aligned}(\gamma,k)(\delta,\ell) &= (\ell\ell^{-1}\gamma \wedge k\delta, k\ell)\\ &= (\ell\ell^{-1}\gamma, \ell\ell^{-1}k)\cdot(kk^{-1}\delta, kk^{-1}\ell)\\ &= (\gamma,k)\varphi_{kk^{-1},kk^{-1}\ell\ell^{-1}}\\ &\quad \cdot(\delta,\ell)\varphi_{\ell\ell^{-1},kk^{-1}\ell\ell^{-1}},\end{aligned}$$

proving the Corollary. ∎

Now given a (strongly) E-reflexive inverse semigroup S, we can apply Corollary 7 to S using an idempotent pure Clifford congruence, and obtain O'Coarroll's result of [5]: any (strongly) E-reflexive inverse semigroup is embedded in a strong semilattice of inverse semigroups, each of which is a semidirect product of a semilattice and a group.

References

[1] Billhardt, B., *On inverse semigroups whose closure of idempotents is a Clifford semigroup*, submitted to Semigroup Forum.

[2] Houghton, C.H., *Embedding inverse semigroups in wreath products*, Glasgow Math. J. **17** (1976), 77–82.

[3] McAlister, D.B., *Semilattices and inverse semigroups*, II, Trans. Amer. Math. Soc. **196** (1974), 351–370.

[4] O'Carroll, L., *Inverse semigroups as extensions of semilattices*, Glasgow Math. J. **16** (1975), 12–21.

[5] O'Carroll, L., *Strongly E-reflexive inverse semigroups*, Proc. Edinburgh Math. Soc. **2** (1976/77), 339–354.

[6] Petrich, M., "Inverse semigroups," John Wiley & Sons, Inc., New York, 1984.

[7] Schein, B.M., *A new proof for the McAlister P-theorem*, Semigroup Forum **10** (1975), 185–188.

Fachbereich Mathematik
Gesamthochschule Kassel
Nora-Platiel-Str. 1
D-3500 Kassel, FR Germany

Received July 17, 1990
and in final form September 19, 1991

Semigroup Forum Vol. 45 (1992) 55–62

RESEARCH ARTICLE

Perfect elements in Dubreil-Jacotin regular semigroups

T. S. Blyth* and Emília Giraldes

Communicated by R. McFadden

Abstract. If S is a strong Dubreil-Jacotin regular semigroup then $x \in S$ is said to be *perfect* if $x = x(\xi : x)x$ where ξ is the bimaximum element of S. It is shown that the set $P(S)$ of perfect elements is an ideal of S, and is also a strong Dubreil-Jacotin subsemigroup. It is then proved that every element of S is perfect if and only if S is naturally ordered. Finally, necessary and sufficient conditions for $P(S)$ to be orthodox are determined.

1. Introduction

In the theory of ordered sets an important rôle is played by the *residuated mappings*. If A, B are ordered sets then a mapping $f : A \to B$ is said to be *residuated* if the pre-image of every principal order ideal of B is a principal order ideal of A. Placing some algebraic structure on the ordered sets A, B we are led to the notion of a *strong Dubreil-Jacotin semigroup* (see, for example, [1]). This is an ordered semigroup S for which there is an ordered group G and a residuated epimorphism $f : S \to G$. The pre-image under f of the negative cone $N(G) = \{x \in G \mid x \leq 1\}$ of G is then a principal order ideal $\xi^{\downarrow} = \{x \in S \mid x \leq \xi\}$ of S, the so-called *bimaximum element* ξ being *equiresidual* in the sense that for every $x \in S$ the order ideals $\{y \in S \mid xy \leq \xi\}$ and $\{y \in S \mid yx \leq \xi\}$ coincide and have a greatest element, this being denoted by $\xi : x$. It turns out (see [1, Theorem 25.7]) that the ordered group G in question is unique up to isomorphism and is given by S/A_ξ where A_ξ is the closure equivalence given by

$$(x, y) \in A_\xi \iff \xi : x = \xi : y\ .$$

Here we shall be concerned uniquely with the case where the semigroup in question is regular. Thus, *in what follows S will always denote a strong Dubreil-Jacotin regular semigroup.* It is important to note that, in such a semigroup, ξ is an idempotent. In fact, ξ is the greatest element in the identity class of the group S/A_ξ, so if $\xi' \in V(\xi)$ then, since ξ^2 and ξ' are also in this class, we have $\xi^2 \leq \xi$ and $\xi' \leq \xi$ whence $\xi = \xi\xi'\xi \leq \xi^3 \leq \xi^2 \leq \xi$ and so $\xi^2 = \xi$. Furthermore, every idempotent e belongs to the identity class of the group S/A_ξ and therefore $\xi : e = \xi : \xi = \xi$.

In [2] a characterization was given of Dubreil-Jacotin semigroups that are perfect. Here we introduce the notion of a perfect element. We show that the set of perfect elements forms an ideal which is a strong Dubreil-Jacotin regular subsemigroup, and use this to prove that a strong Dubreil-Jacotin regular semigroup is perfect if and only if it is naturally ordered. Finally, we determine necessary and sufficient conditions for the subsemigroup of perfect elements to be orthodox.

---

* Support from the Junta Nacional de Investigação Científica e Tecnológica of Portugal is gratefully acknowledged.

2. Perfect elements

We begin by establishing a basic inequality that leads in a natural way to the type of element that we shall consider.

Theorem 1. $(\forall\, x \in S) \qquad x \leq x(\xi : x)x$.

Proof. Given $x \in S$, let $x' \in V(x)$. Since ξ is the biggest idempotent in S we have $xx' \leq \xi$ and so $x' \leq \xi : x$, whence

$$x = xx'x \leq x(\xi : x)x \ . \qquad \blacksquare$$

Definition. We shall say that $x \in S$ is *perfect* if $x = x(\xi : x)x$.

In what follows, we shall denote by $P(S)$ the set of perfect elements of S. Note that $P(S) \neq \emptyset$ since it contains ξ; for, ξ is idempotent and $\xi : \xi = \xi$.

Theorem 2. *$P(S)$ is an ideal of S.*

Proof. Suppose that $x \in P(S)$. Since $xy(\xi : xy) \leq \xi$ we see that $y(\xi : xy)$ belongs to the order ideal $\{z \in S \mid xz \leq \xi\}$ and so $y(\xi : xy) \leq \xi : x$. Thus

$$xy(\xi : xy)xy \leq x(\xi : x)xy = xy \ .$$

It follows by Theorem 1 that $xy \in P(S)$. Hence $P(S)$ is a right ideal of S; and similarly it is also a left ideal. $\blacksquare$

To investigate the ideal of perfect elements, we require the following result.

Theorem 3. $(\forall\, x \in S) \qquad (\xi : x)\xi = \xi : x = \xi(\xi : x)$.

Proof. Since $x(\xi : x)\xi \leq \xi^2 = \xi$ we see that $(\xi : x)\xi$ belongs to the order ideal $\{z \in S \mid xz \leq \xi\}$ and so $(\xi : x)\xi \leq \xi : x$; and similarly $\xi(\xi : x) \leq \xi : x$. But by Theorem 1 we have $x \leq \xi x$ and $x \leq x\xi$. Writing $\xi : x$ for x in these, we obtain $\xi : x \leq \xi(\xi : x)$ and $\xi : x \leq (\xi : x)\xi$ whence the required equalities follow. $\blacksquare$

Corollary 1. $(\forall\, x \in S) \qquad \xi : x \in P(S)$.

Proof. Since $\xi \in P(S)$, this follows from Theorems 2 and 3. $\blacksquare$

Corollary 2. *$P(S)$ is a regular subsemigroup of S.*

Proof. For every $x \in P(S)$ we have $x = x(\xi : x)x$ with $\xi : x \in P(S)$. $\blacksquare$

Corollary 3. *$x(\xi : x)$ and $(\xi : x)x$ are perfect idempotents.*

Proof. That $x(\xi : x) \in P(S)$ and $(\xi : x)x \in P(S)$ follow by Corollary 1 and Theorem 2. Observe now by Theorem 1 that

$$x(\xi : x) \leq x(\xi : x)x(\xi : x) \ .$$

But, using Theorem 3, we also have

$$x(\xi : x)x(\xi : x) \leq x(\xi : x)\xi = x(\xi : x) \ .$$

Hence $x(\xi : x)$ is idempotent; and similarly so is $(\xi : x)x$. $\blacksquare$

Example 1. Let V be a $\vee$-semilattice with greatest element 1 and let G be an ordered group. Endow the set $S = V \times V \times G$ with the cartesian order and the multiplication

$$(x, y, g)(x', y', g') = (x \vee x', y, gg') \ .$$

Then S is an ordered semigroup. Moreover, S is regular; for example, we have

$$(x, y, g)(x, y, g^{-1})(x, y, g) = (x, y, g) \ .$$

Now $\vartheta : S \to G$ given by $\vartheta(x, y, g) = g$ is readily seen to be a residuated epimorphism, so S is a strong Dubreil-Jacotin semigroup. Here we have $\xi = (1, 1, 1_G)$, and the residuals of ξ are given by

$$\xi : (x, y, g) = (1, 1, g^{-1}) \ .$$

The idempotents of S are the elements of the form $(x, y, 1_G)$. These do not commute, so S is not an inverse semigroup. Since

$$(x, y, g)[\xi : (x, y, g)](x, y, g) = (1, y, g)$$

it follows that the set of perfect elements of S is

$$P(S) = \{(1, y, g) \mid y \in V, g \in G\} \ .$$

Recall that the *natural order* $\preceq$ on the set E of idempotents is defined by

$$e \preceq f \iff e = ef = fe \ .$$

We say that S is *naturally ordered* if the ordering $\leq$ on S extends the natural order $\preceq$ on E, in the sense that

$$e \preceq f \Longrightarrow e \leq f \ .$$

Theorem 4. *$P(S)$ is naturally ordered.*

Proof. Let $e, f \in E \cap P(S)$ with $e \preceq f$. Then

$$e = e(\xi : e)e = e\xi e = fe\xi ef = fef \leq f\xi f = f(\xi : f)f = f \ . \qquad \blacksquare$$

Definition. Let S be a regular semigroup and let $\overline{E}$ be the (regular) subsemigroup generated by the set E of idempotents of S. Then an idempotent u of S is said to be *medial* if

$$(\forall\, \overline{e} \in \overline{E}) \qquad \overline{e} = \overline{e}u\overline{e} \ .$$

Definition. We shall say that an idempotent u is a *weak middle unit* if

$$(\forall\, x \in S)(\forall\, x' \in V(x)) \qquad xux' = xx' \ .$$

Definition. We shall say that S is *perfect* if every element of S is perfect.

We now show that S is perfect if and only if it is naturally ordered.

Theorem 5. *The following statements are equivalent:*

(1) *S is perfect;*

(2) *every idempotent of S is perfect;*

(3) *ξ is a medial idempotent;*

(4) *ξ is a weak middle unit;*

(5) *S is naturally ordered.*

Proof. (1) $\Longrightarrow$ (2) is clear.

(2) $\Longrightarrow$ (3): Every $\overline{e} \in \overline{E}$ is of the form

$$\overline{e} = e_1 e_2 \dots e_n$$

where each $e_i \in E$. It follows that $\xi : \overline{e} = \xi$. Now if (2) holds then every $e_i \in P(S)$ and so, by Theorem 2, $\overline{e} \in P(S)$. Consequently,

$$\overline{e} = \overline{e}(\xi : \overline{e})\overline{e} = \overline{e}\xi\overline{e}$$

and so ξ is a medial idempotent.

(3) $\Longrightarrow$ (4): This is clear.

(4) $\Longrightarrow$ (1): Given $x \in S$ let $x' \in V(x)$. Then if (4) holds we have

$$x'(\xi : x')x' = x'xx'(\xi : x')x' \leq x'x\xi x' = x'xx' = x'$$

and so, by Theorem 1, $x' \in P(S)$. It follows by Theorem 2 that $x = xx'x \in P(S)$, whence S is perfect.

(1) $\Longrightarrow$ (5): This follows by Theorem 4.

(5) $\Longrightarrow$ (2): By Corollary 3 of Theorem 3 we have $e\xi e = e(\xi : e)e \in E$ for every $e \in E$. But $e\xi e \preceq e$ so, by (5), $e\xi e \leq e$. Consequently, $e = e\xi e \in P(S)$. ∎

Our objective now is to describe the perfect elements of S. For this purpose, for every $x \in S$ define x° by

$$x^\circ = (\xi : x)x(\xi : x)\ .$$

Note that, by Theorem 2 and Corollary 1 of Theorem 3, we have $x^\circ \in P(S)$ for every $x \in S$. In order to investigate the elements of this form, we require the following result.

Theorem 6. *For every $x \in S$,*

$$x(\xi : x)[\xi : (\xi : x)] = x\xi \quad \textit{and} \quad [\xi : (\xi : x)](\xi : x)x = \xi x\ .$$

Proof. By Theorem 1 we have

$$x \leq x(\xi : x)x \leq x(\xi : x)[\xi : (\xi : x)] \leq x\xi\ .$$

Multiplying throughout on the right by the idempotent ξ and using Theorem 3, we obtain

$$x\xi = x(\xi : x)[\xi : (\xi : x)]\ .$$

The other identity is established similarly. ∎

Theorem 7. $(\forall\, x \in S) \quad x \leq x^{\circ\circ} = \xi x \xi$.

Proof. Observe first that for every idempotent $e \in S$ we have $\xi : e = \xi$ and therefore, by Corollary 3 of Theorem 3,

$$\xi : x^{\circ} = \xi : (\xi : x)x(\xi : x) = \xi : (\xi : x)\ .$$

Consequently, we have

$$\begin{aligned} x^{\circ\circ} &= (\xi : x^{\circ})x^{\circ}(\xi : x^{\circ}) \\ &= [\xi : (\xi : x)](\xi : x)x(\xi : x)[\xi : (\xi : x)] \\ &= \xi x \xi \qquad \text{[by Theorem 6]} \\ &\geq x(\xi : x)x(\xi : x)x \\ &= x(\xi : x)x \quad \text{[by Corollary 3 of Theorem 3]} \\ &\geq x \qquad \text{[by Theorem 1]}\ . \end{aligned}$$

■

Theorem 8. $(\forall\, x \in S) \quad (x^{\circ})^{\circ\circ} = x^{\circ} = (x^{\circ\circ})^{\circ}$.

Proof. By Theorem 3, we have $\xi x^{\circ} = x^{\circ} = x^{\circ}\xi$ and so, by Theorem 7,

$$(x^{\circ})^{\circ\circ} = \xi x^{\circ} \xi = x^{\circ}\ .$$

Also, observing that

$$\xi : x^{\circ\circ} = \xi : (\xi : x^{\circ}) = \xi : x\ ,$$

we have

$$\begin{aligned} (x^{\circ\circ})^{\circ} &= (\xi : x^{\circ\circ})x^{\circ\circ}(\xi : x^{\circ\circ}) \\ &= (\xi : x)\xi x \xi(\xi : x) \quad \text{[by Theorem 7]} \\ &= (\xi : x)x(\xi : x) \qquad \text{[by Theorem 3]} \\ &= x^{\circ}\ . \end{aligned}$$

■

Note that by Theorem 8 we can write $x^{\circ\circ\circ}$ unambiguously, and that $x^{\circ\circ\circ} = x^{\circ}$. Also, by Theorem 7, $x \mapsto x^{\circ\circ}$ is a closure.

Theorem 9. *The perfect elements of S are precisely those elements x for which x° is an inverse of x, in which case it is the greatest inverse.*

Proof. Using Corollary 3 of Theorem 3 and the definition of x°, it is readily seen that for every $x \in S$ we have

$$x^{\circ}xx^{\circ} = x^{\circ} \quad \text{and} \quad xx^{\circ}x = x(\xi : x)x\ .$$

It follows that if $x \in P(S)$ then $x^{\circ} \in V(x)$. In this case, it follows by [4, Corollary 2 of Theorem 1.2] that the greatest inverse of x is $\xi x^{\circ}\xi = x^{\circ}$. Conversely, if x° is an inverse of x then it is clear from the above that $x = x(\xi : x)x$ and so $x \in P(S)$. ■

3. When is $P(S)$ orthodox?

We have seen above that $P(S)$ is a regular subsemigroup of S that contains ξ and $\xi : x$ for every $x \in S$. It follows that $P(S)$ is also strong Dubreil-Jacotin; for, if $f : S \to G$ is a residuated epimorphism then since f carries idempotents to 1_G, and in particular the perfect idempotents of the form $(\xi : x)x$, we have $f[x(\xi : x)x] = f(x)$ and so $f|_{P(S)} : P(S) \to G$ is also a residuated epimorphism. Our objective now is to determine precisely when $P(S)$ is orthodox. This is the case in Example 1 above , but need not be so in general.

Example 2. The semigroup N_5 which is described in [4, Theorem 3.2] is a non-orthodox Dubreil-Jacotin regular semigroup. It is naturally ordered and so, by Theorem 5, is perfect. Thus $P(N_5) = N_5$ is non-orthodox.

For our present purpose we define, for every $x \in S$, the element x_ϑ by

$$x_\vartheta = x(\xi : x)x \ .$$

Note that $x_\vartheta \in P(S)$, and that $x \in P(S)$ if and only if $x = x_\vartheta$. We shall require the following properties of x_ϑ.

Theorem 10. $(\forall\ x \in S) \qquad \xi x_\vartheta = \xi x, \quad x_\vartheta \xi = x\xi.$

Proof. Since $\xi \in P(S)$ we have $\xi x \in P(S)$ and so

$$\begin{aligned} \xi x = (\xi x)_\vartheta &= \xi x(\xi : \xi x)\xi x \\ &= \xi x(\xi : x)\xi x && [\text{since } \xi : \xi = \xi] \\ &= \xi x(\xi : x)x && [\text{by Theorem 3}] \\ &= \xi x_\vartheta \ . \end{aligned}$$

Similarly we can show that $x_\vartheta \xi = x\xi$. ■

Theorem 11. $(\forall\ x \in S) \qquad (x^\circ)_\vartheta = x^\circ = (x_\vartheta)^\circ.$

Proof. The first equality follows from the fact that $x^\circ \in P(S)$. As for the second, using Theorems 7 and 10 we have

$$(x_\vartheta)^{\circ\circ} = \xi x_\vartheta \xi = \xi x \xi = x^{\circ\circ}$$

and hence, by Theorem 8,

$$(x_\vartheta)^\circ = (x_\vartheta)^{\circ\circ\circ} = x^{\circ\circ\circ} = x^\circ \ .$$ ■

Definition. An element u of a semigroup S is said to be a *middle unit* if

$$(\forall\ x, y \in S) \quad xuy = xy \ .$$

As shown in [5, Proposition 1.9], an ordered regular semigroup with a biggest idempotent u is naturally ordered and orthodox if and only if u is a middle unit. Applying this to the present context we can assert, by Theorem 5, that a perfect Dubreil-Jacotin regular semigroup is orthodox if and only if ξ is a middle unit.

Also, as shown in [3, Lemma 5.1], if S is an ordered regular semigroup with a greatest element ξ then the subsemigroup $\xi S \xi$ is a semilattice. In the present context, we can therefore assert that if $B = [\xi]_{A_\xi}$ is the identity class of the ordered group S/A_ξ then $\xi B \xi$ is a semilattice.

We shall use both of these observations in what follows.

Theorem 12. *The following statements are equivalent:*

(1) *$P(S)$ is orthodox;*

(2) $(\forall\, x, y \in S) \quad x_\vartheta y_\vartheta = x\xi y$;

(3) $(\forall\, x, y \in S)(\forall\, e, f \in E) \quad x_\vartheta efy_\vartheta = x_\vartheta fey_\vartheta$.

Proof. (1) $\Longrightarrow$ (2): If (1) holds then ξ is a middle unit for $P(S)$ and so, by Theorem 10, $x_\vartheta y_\vartheta = x_\vartheta \xi y_\vartheta = x\xi y$.

(2) $\Longrightarrow$ (1): Let $x, y \in P(S)$, so that $x = x_\vartheta$ and $y = y_\vartheta$. Then if (2) holds we have $xy = x_\vartheta y_\vartheta = x\xi y$, whence ξ is a middle unit for $P(S)$ and so $P(S)$ is orthodox.

(1) $\Longrightarrow$ (3): Since $P(S)$ is an ideal of S we have $x_\vartheta e \in P(S)$ and $fy_\vartheta \in P(S)$ for all $x, y \in S$ and all $e, f \in E$. It follows that if (1) holds then

$$\begin{aligned} x_\vartheta efy_\vartheta &= x_\vartheta e\xi fy_\vartheta \\ &= x_\vartheta e_\vartheta \xi f_\vartheta y_\vartheta && \text{[by Theorem 10]} \\ &= x_\vartheta \cdot \xi e_\vartheta \xi \cdot \xi f_\vartheta \xi \cdot y_\vartheta && \text{[by (1)]} \\ &= x_\vartheta \cdot \xi f_\vartheta \xi \cdot \xi e_\vartheta \xi \cdot y_\vartheta && [\xi B \xi \text{ a semilattice}] \end{aligned}$$

and so, reversing the steps, we see that $x_\vartheta efy_\vartheta = x_\vartheta fey_\vartheta$.

(3) $\Longrightarrow$ (1): Taking $x = e = e_\vartheta$ and $y = f = f_\vartheta$ in (3), we obtain $e_\vartheta f_\vartheta = e_\vartheta f_\vartheta e_\vartheta f_\vartheta$, i.e. the product of two perfect idempotents is idempotent. Hence $P(S)$ is orthodox. ■

Using Theorem 12 we now establish a more satisfying criterion for $P(S)$ to be orthodox.

Theorem 13. *The following statements are equivalent:*

(1) *$P(S)$ is orthodox;*

(2) $(\forall\, x, y \in S) \qquad (xy)^\circ = y^\circ x^\circ$.

Proof. (1) $\Longrightarrow$ (2): Observe first that if $a, b \in P(S)$ then by Theorem 9 we have $a^\circ \in V(a)$ and $b^\circ \in V(b)$, so if $P(S)$ is orthodox it follows that $b^\circ a^\circ \in V(ab)$. Consequently,

$$abb^\circ a^\circ ab(ab)^\circ = ab(ab)^\circ \quad \text{and} \quad ab(ab)^\circ abb^\circ a^\circ = abb^\circ a^\circ,$$

from which we see that the idempotents $abb^\circ a^\circ$ and $ab(ab)^\circ$ are in the same $\mathcal{R}$-class. But if $e, f \in E$ are $\mathcal{R}$-equivalent then from $e = fe$ we have $e \leq f\xi$ whence $e\xi \leq f\xi$, and similarly $f\xi \leq e\xi$, so that $e\xi = f\xi$. Hence, using Theorem 3, we see that

$$ab(ab)^\circ = ab(ab)^\circ \xi = abb^\circ a^\circ \xi = abb^\circ a^\circ$$

and similarly $(ab)^\circ ab = b^\circ a^\circ ab$. These identities give, for all $a, b \in P(S)$,

$$(ab)^\circ = (ab)^\circ ab(ab)^\circ = b^\circ a^\circ ab(ab)^\circ = b^\circ a^\circ abb^\circ a^\circ = b^\circ a^\circ \ .$$

Suppose now that (1) holds. Then by Theorem 11, the above observation, and Theorem 12(2), we have

$$(\forall\, x, y \in S) \qquad y^\circ x^\circ = y^\circ_\vartheta x^\circ_\vartheta = (x_\vartheta y_\vartheta)^\circ = (x\xi y)^\circ \ .$$

But we also have

$$\begin{aligned} y^\circ x^\circ &= (\xi : y)y(\xi : y)(\xi : x)x(\xi : x) \\ &= (\xi : y)(\xi : x)xy(\xi : y)(\xi : x) && \text{[by Theorem 12(3)]} \\ &\leq (\xi : xy)xy(\xi : xy) = (xy)^\circ \\ &\leq (\xi : xy)x\xi y(\xi : xy) \\ &= (\xi : x\xi y)x\xi y(\xi : x\xi y) \\ &= (x\xi y)^\circ \ , \end{aligned}$$

from which (2) now follows.

(2) $\Longrightarrow$ (1): Let $e, f \in E \cap P(S)$. By (2), we have

$$\xi ef\xi = (\xi : ef)ef(\xi : ef) = (ef)^\circ = f^\circ e^\circ = \xi f\xi e\xi .$$

Pre-multiplying by e and post-multiplying by f, and using the fact that $e = e\xi e$ and $f = f\xi f$, we obtain

$$ef = e\xi f \cdot \xi \cdot e\xi f = (e\xi f)_\vartheta = e\xi f .$$

Now let $x, y \in P(S)$. Then by the above we have

$$\begin{aligned} xy &= x(\xi : x)xy(\xi : y)y \\ &= x(\xi : x)x\xi y(\xi : y)y \\ &= x\xi y . \end{aligned}$$

Thus ξ is a middle unit for $P(S)$ and so $P(S)$ is orthodox. ∎

References

[1] Blyth, T. S. and M. F. Janowitz, *Residuation Theory*, Pergamon Press, 1972.

[2] Blyth, T. S., *Perfect Dubreil-Jacotin semigroups*, Proc. Roy. Soc. Edinburgh **78A** (1977), 101–104.

[3] Blyth, T. S. and D. B. McAlister, *Split orthodox semigroups*, Journal of Algebra **51** (1978), 491–525.

[4] Blyth, T. S. and R. McFadden, *Naturally ordered regular semigroups with a greatest idempotent*, Proc. Roy. Soc. Edinburgh **91A** (1981), 107–122.

[5] McAlister, D. B., *Regular Rees matrix semigroups and regular Dubreil-Jacotin semigroups*, J. Australian Math. Soc. **31** (1981), 325–336.

Mathematical Institute
University of St. Andrews
Scotland

Departamento de Matemática
Universidade Nova de Lisboa
Portugal

Received July 17, 1990
and in final form January 16, 1991

Semigroup Forum Vol. 45 (1992) 63–70

RESEARCH ARTICLE

Extension properties of WS-groups

G. Hansel and J. P. Troallic

Communicated by J. Pym

1. Introduction

In a previous work [8], we have established several properties of the class of "strictly" weakly almost periodic functions (see Definition 2.2) on a topological group. In particular, we gave there examples of WS-groups, namely topological groups for which any weakly almost periodic function has also the corresponding "strict" property.

In the present work, we carry on with this investigation. In Section 3, we show that under a wide hypothesis, any WS-group is a SIN-group, i.e. admits a basis of neighbourhoods of the identity which is invariant under all inner automorphisms. In Sections 5 and 6, we give conditions which insure that a group obtained as an "extension" of a WS-group is itself a WS-group. In Section 5, we consider compact extensions and Section 6 is dedicated to the particular case of semidirect products.

Note that unlike [8], we do not assume anymore (unless explicitly stated) that the groups which we consider are locally compact.

2. Preliminaries

All topological spaces considered in this work are assumed to be Hausdorff. Let us first recall an important criterion of weak relative compactness in function spaces due to A. Grothendieck [7].

Lemma 2.1 [7]. *Let X be a topological space and $C(X)$ the complex Banach space of bounded continuous complex valued functions on X. Let A be a subset of $C(X)$. The following conditions are equivalent:*

1) A is a weakly relatively compact subset of $C(X)$.

2) A is bounded and for every sequence $(x_n)_{n\in\mathbb{N}}$ of points in X and every sequence $(f_p)_{p\in\mathbb{N}}$ of points in A,

$$\lim_n \lim_p f_p(x_n) = \lim_p \lim_n f_p(x_n),$$

as soon as the limits in the above equation exist.

Let us recall some basic definitions about weak almost periodicity in topological groups (see [2] or [3]). Let G be a locally compact topological group. We denote by $C(G)$ the set of bounded continuous complex valued functions on G. It is supposed to be equipped with its usual commutative C^*-algebra structure. For all $f \in C(G)$ and all $g \in G$, we denote by f_g (respectively f^g) the *left translate* (respectively *right translate*) of f by g defined by the formula $f_g(x) = f(gx)$ (respectively $f^g(x) = f(xg)$) for all $x \in G$. For all $g, h \in G$, we put $f_g^h = (f_g)^h (= (f^h)_g)$.

A function $f \in C(G)$ is *weakly almost periodic* if $\{f_g \mid g \in G\}$ is a weakly relatively compact subset of $C(G)$. We denote by $W(G)$ the set of these functions. As an easy consequence of Lemma 2.1, the set $\{ f_g \mid g \in G \}$ is weakly relatively compact iff the set $\{ f^g \mid g \in G \}$ is. Thus weakly almost periodic functions on G could have been defined in terms of right translation instead of left one. The set $W(G)$ is a sub-C^*-algebra of $C(G)$ which contains the constant functions; $W(G)$ is *translation invariant*, namely for all $f \in W(G)$ and all $g \in G$, f_g and f^g belong to $W(G)$.

Now we introduce a particular class of weakly almost periodic functions.

Definition 2.2. Let G be a topological group and let $f \in C(G)$. We shall say that the function f is *strictly weakly almost periodic* if $\{ f_g^h \mid g, h \in G \}$ is a weakly relatively compact subset of $C(G)$. We denote by $WS(G)$ the set of these functions. The set $WS(G)$ is a translation invariant sub-C^*-algebra of $W(G)$ and contains the constant functions. We denote by [WS] the class of all topological groups G which have the property that $WS(G) = W(G)$. If a group belongs to [WS], we say it is a WS-group.

Finally, let us recall the definition of some usual classes of topological groups.

A topological group G is called *central* if the quotient G/Z of G by its center Z is compact. It is called *almost connected* if the quotient G/G_0, where G_0 is the connected component of the identity e of G, is compact.

Let G be a topological group and $\mathcal{A}(G)$ the group of its (topological) automorphisms. A subset $\mathcal{B}$ of $\mathcal{A}(G)$ is called *equicontinuous* if for every neighbourhood V of e in G, the set $\bigcap_{\phi \in \mathcal{B}} \phi^{-1}(V)$ is a neighbourhood of e. It is equivalent to say that $\mathcal{B}$ is equicontinuous when G is equipped with one of its usual uniform structures. The topological group G is said to be a *SIN-group* if the group $\mathcal{I}(G)$ of its inner automorphisms is equicontinuous, i.e. if for any neighbourhood V of e, the set $\bigcap_{g \in G} g^{-1}Vg$ is still a neighbourhood of e. This amounts to saying that its left and right uniform structures are equal [9]. The class of all SIN-groups is denoted by [SIN].

3. WS-groups and SIN-groups

In this section, we prove that within a wide class of topological groups, if a group belongs to [WS], it also necessarily belongs to [SIN]. This class includes all locally compact topological groups.

Theorem 3.1. *Let G be a topological group which belongs to* [WS]. *Let us suppose that the topology on G is induced by the functions of $W(G)$. Then G belongs to* [SIN].

Proof. Let $(x_\alpha)_{\alpha \in D}$ and $(y_\alpha)_{\alpha \in D}$ be two nets of elements of G. We are going to establish that G is a SIN-group by proving that if (x_α) converges to e, then $(y_\alpha x_\alpha y_\alpha^{-1})$ also converges to e. Since $W(G)$ induces the topology of G, it is sufficient to show that for all $f \in W(G)$, the net $\big(f(y_\alpha x_\alpha y_\alpha^{-1})\big)$ converges to $f(e)$.

Let $\varepsilon > 0$ and let us show that $|f(y_\alpha x_\alpha y_\alpha^{-1}) - f(e)| < \varepsilon$ if α is big enough. Let G_1 be the opposite group of G (as a topological space $G_1 = G$ and the product on G_1 is defined by $(x, y) \mapsto yx$) and let $G_1 \times G$ be the direct product of G_1 and G. Let $F : G_1 \times G \to \mathbb{C}$ be the bounded continuous function defined by $F(x_1, x) = f(x_1 x)$, $(x_1, x) \in G_1 \times G$. Then by Lemma 2.1, F belongs to $W(G_1 \times G)$. Indeed, let (g_{1n}, g_n) and (h_{1p}, h_p) be two sequences of elements of $G_1 \times G$ such that

$$l_1 = \lim_n \lim_p F\big((g_{1n}, g_n)(h_{1p}, h_p)\big) \quad \text{and} \quad l_2 = \lim_p \lim_n F\big((g_{1n}, g_n)(h_{1p}, h_p)\big)$$

exist; we have

$$l_1 = \lim_n \lim_p f(h_{1p} g_{1n} g_n h_p) \quad \text{and} \quad l_2 = \lim_p \lim_n f(h_{1p} g_{1n} g_n h_p).$$

Since $f \in WS(G)$, using Lemma 2.1, we get that $l_1 = l_2$.

The function F being weakly almost periodic, it is uniformly continuous when $G_1 \times G$ is equipped with its right (for instance) uniform structure [3]. Let V be a neighbourhood of e in G such that for all $(g_1, g), (h_1, h) \in G_1 \times G$, the condition $(g_1, g)(h_1, h)^{-1} \in V \times V$ implies $|F(g_1, g) - F(h_1, h)| < \varepsilon$ or equivalently,

$$h_1 g_1 \in V,\ gh \in V \Rightarrow |f(g_1 g) - f(h_1 h)| < \varepsilon. \tag{1}$$

Let $\alpha_0 \in D$ such that $x_\alpha \in V$ for $\alpha \geq \alpha_0$. Let $\alpha \geq \alpha_0$; let us put $h_1 = y_\alpha^{-1}$, $g_1 = y_\alpha x_\alpha$, $g = y_\alpha^{-1}$, and $h = y_\alpha$; then it follows from (1) that

$$|f(y_\alpha x_\alpha y_\alpha^{-1}) - f(e)| < \varepsilon. \qquad \blacksquare$$

Corollary 3.2[8]. *Let G be a locally compact group. If G belongs to* [WS], *then it belongs to* [SIN].

Proof. The functions $f \in C(G)$ vanishing at infinity induce the topology of G and are well known to belong to $W(G)$. $\blacksquare$

Remark 3.3. Let G be a topological group and (G^w, ϕ) its weakly almost periodic compactification. If ϕ is a homeomorphic imbedding, then the topology on G is induced by the functions of $W(G)$ [13].

4. Two basic lemmas

The following lemma states an important technical condition which insures that a group belongs to [WS].

Lemma 4.1. *Let G be a topological group and H a subgroup of G. Let us assume that H belongs to* [WS] *and that for every neighbourhood V of e in G, the following condition (C) is satisfied: there exist a subset A of V invariant under all inner automorphisms of G and a finite subset F of G such that $G = AFH = HFA$. Then G belongs to* [WS].

Proof. Let $f \in W(G)$ and let $l_1, l_2 \in \mathbb{C}$. Let (g_n), (x_p) and (h_n) be sequences of elements of G such that

$$\lim_n \lim_p f(g_n x_p h_n) = l_1 \quad \text{and} \quad \lim_p \lim_n f(g_n x_p h_n) = l_2.$$

According to Lemma 2.1, we only have to show that $l_1 = l_2$ and it will follow that $f \in WS(G)$.

Let $\varepsilon > 0$. Let us equip G with its right uniform structure; then every element of $W(G)$ is uniformly continuous [3]. Hence there is a neighbourhood W of e in G such that

$$\forall (x, y) \in G \times G : xy^{-1} \in W \Rightarrow |f(x) - f(y)| \leq \varepsilon/2. \tag{1}$$

Let V be a neighbourhood of e in G such that $V^5 \subset W$. Let A and F be sets satisfying Condition (C) with respect to V. For all $n, p \in \mathbb{N}$ we can write, using the equality $G = AFH$,

$$g_n = g'_n a_n g''_n,\ x_p = x'_p r_p x''_p,\ h_n = h'_n b_n h''_n$$

with $g'_n, x'_p, h'_n \in A$, $a_n, r_p, b_n \in F$, $g''_n, x''_p, h''_n \in H$. Since F is finite, we can suppose, by taking -if necessary- subsequences of (g_n), (x_p) and (h_n), that there exist $a, r, b \in F$ such that for all $n \in \mathbb{N}$, $a_n = a$, $r_n = r$ and $b_n = b$. Then for all $n \in \mathbb{N}$ we can write, using the equalities $G = AFH = HFA$,

$$g''_n r = u'_n s_n u''_n, \quad bh''_n = v'' c_n v'_n$$

with $u'_n, v'_n \in A$, $s_n, c_n \in F$, $u''_n, v''_n \in H$. Again, we can suppose that there exist $s, c \in F$ such that for all $n \in \mathbb{N}$, $s_n = s$ and $c_n = c$.

For all $n, p \in \mathbb{N}$, we get the equalities

$$(g_n x_p h_n)(ag''_n r x''_p b h''_n)^{-1} = g'_n\Big((ag''_n)x'_p(ag''_n)^{-1}\Big)\Big((ag''_n r x''_p)h'_n(ag''_n r x''_p)^{-1}\Big)$$

and

$$(ag''_n r x''_p b h''_n)(asu''_n x''_p v''_n c)^{-1} = (au'_n a^{-1})\Big((asu''_n x''_p v''_n c)v'_n(asu''_n x''_p v''_n c)^{-1}\Big).$$

Since $g'_n, x'_p, h'_n, u'_n, v'_n$ belong to A which is invariant under all inner automorphisms, we deduce from these equalities that $(g_n x_p h_n)(asu''_n x''_p v''_n c)^{-1}$ belongs to A^5, hence to V^5 and hence to W. It follows by (1) that for all $n, p \in \mathbb{N}$, we have

$$|f(g_n x_p h_n) - f^c_{as}(u''_n x''_p v''_n)| \leq \varepsilon/2. \tag{2}$$

By taking, if necessary, subsequences of (g_n), (x_p) and (h_n), we can suppose that the limits

$$l'_1 = \lim_n \lim_p f^c_{as}(u''_n x''_p v''_n) \quad \text{and} \quad l'_2 = \lim_p \lim_n f^c_{as}(u''_n x''_p v''_n)$$

exist. Since the function f belongs to $W(G)$, the function f^c_{as} also belongs to $W(G)$ and its restriction to H belongs to $W(H)$. As the group H is supposed to be in [WS], we get by Lemma 2.1 that $l'_1 = l'_2 (= l')$. Then it follows from (2) that $|l_1 - l'| \leq \varepsilon/2$ and $|l_2 - l'| \leq \varepsilon/2$. Hence $|l_1 - l_2| \leq \varepsilon$. But ε can be fixed arbitrarily small and thus $l_1 = l_2$. ∎

When H is a normal subgroup of a topological group G, each inner automorphism of G induces an automorphism of H. The next lemma gives a condition which insures that this set of induced automorphisms is equicontinuous.

Lemma 4.2. *Let G be a topological group, H a normal subgroup of G, and let G' be a subgroup of G such that the homogeneous space G/G' is compact. For all $g \in G$, let θ_g be the automorphism $h \mapsto ghg^{-1}$ of H. Let us suppose that the subgroup $\{\theta_g \mid g \in G'\}$ of $\mathcal{A}(H)$ is equicontinuous. Then the subgroup $\{\theta_g \mid g \in G\}$ of $\mathcal{A}(H)$ is also equicontinuous.*

Proof. Let V be a neighbourhood of e in H. Let V' be a neighbourhood of e in G such that $V' \cap H = V$. Let W be an open symmetrical neighbourhood of e in G such that $W^3 \subset V'$. Let ϕ be the canonical function from G onto G/G'. Then the family $\big(\phi(Wg)_{g\in G}\big)$ is an open cover of the compact space G/G'. Let F be a finite subset of G such that $\bigcup_{g\in F} \phi(Wg) = G/G'$ or equivalently such that

$$G = WFG'. \tag{1}$$

By hypothesis, since F is finite, the set $\{\theta_x \mid x \in FG'\}$ is equicontinuous. Consequently, there exists a neighbourhood U of e in H such that

$$\forall x \in FG', \ \forall u \in U : xux^{-1} \in W. \tag{2}$$

Let $g \in G$. By (1) there exist $w \in W$ and $x \in FG'$ such that $g = wx$. Let $u \in U$. Making use of (2), we get

$$gug^{-1} = w(xux^{-1})w^{-1} \in W^3 \cap H \subset V' \cap H = V.$$

Thus $\bigcup_{g\in G} gUg^{-1} \subset V$ and $\{\theta_g \mid g \in G\}$ is equicontinuous. ∎

5. Compact extensions of WS-groups

Theorem 5.1. *Let G be a topological group and H a subgroup of G. Let us suppose that G belongs to* [SIN], *that H belongs to* [WS] *and that the homogeneous space G/H (made up by the left cosets of G with respect to H) is compact. Then G belongs to* [WS].

Proof. Let V be a neighbourhood of e in G and let $A \subset V$ be an open symmetrical neighbourhood of e which is invariant under all inner automorphisms of G. To conclude, it suffices to show that there exists a finite subset F of G such that A and F satisfy Condition (C) of Lemma 4.1 with respect to V.

Let ϕ be the canonical function from G onto G/H. Then the family $\left(\phi(Ag)_{g\in G}\right)$ is an open cover of the compact space G/H. Let F be a symmetrical finite subset of G such that $\bigcup_{g\in F}\phi(Ag) = G/H$. We have $G = AFH$ and $G = G^{-1} = (AFH)^{-1} = H^{-1}F^{-1}A^{-1} = HFA$. ■

Corollary 5.2. *Let G be a topological group and let H be an abelian subgroup of G. If G belongs to* [SIN] *and G/H is compact, then G belongs to* [WS].

Corollary 5.3. *A central topological group G belongs to* [WS].

Proof. By 5.2, we only have to show that G belongs to [SIN]. This follows immediately from Lemma 4.2 (applied with $G' = Z$ and $H = G$). ■

Corollary 5.4[8]. *Let G be a locally compact topological group. If G belongs to* [SIN] *and is almost connected, then G belongs to* [WS].

Proof. The group G_0 is locally compact, connected, and belongs to [SIN]. Hence, by Freudenthal-Weil theorem (see 16.4.6 in [4] or Theorem 4.3 in [5]), the group G_0 is central. Therefore, according to 5.3, G_0 belongs to [WS]. Then, by applying 5.1 with $H = G_0$, we conclude that G belongs to [WS]. ■

Corollary 5.5. *Let G be a locally compact topological group. Let us suppose that $\mathbb{R}$ is a normal subgroup of G and that the quotient $G/\mathbb{R}$ is compact. Then G belongs to* [WS].

Proof. By Theorem 7 of [12], G is a SIN-group. Hence, by 5.2, G is a WS-group. ■

Remark 5.6. It is not possible to replace $\mathbb{R}$ by $\mathbb{C}$ in the above corollary (see Example 6.3 below).

6. Semidirect products

In this section, we show that in several situations, a semidirect product of groups is in [WS], as soon as one of its factors is.

Notations 6.1. Let H and L be topological groups, $\mathcal{A}(H)$ the group of automorphisms of H. Let $\eta : L \to \mathcal{A}(H)$ be an homomorphism such that the function $(h,l) \mapsto \eta(l)h$ is continuous from the product space $H \times L$ to H. The *semidirect product* $H \times_\eta L$ is the topological group obtained by equipping the space $H \times L$ with the multiplication

$$(h_1,l_1)(h_2,l_2) = \left(h_1\left(\eta(l_1)h_2\right), l_1l_2\right).$$

In such a situation, let us put $H' = \{(h,e) \mid h \in H\}$ and $L' = \{(e,l) \mid l \in L\}$. Then H' is a normal subgroup of $H \times_\eta L$; H' (respectively L') is a closed subgroup of $H \times_\eta L$ and the function $h \mapsto (h,e)$ (respectively $l \mapsto (e,l)$) is an isomorphism from H onto H' (respectively L onto L'). Moreover, $H \times_\eta L = H'L'$, $H' \cap L' = \{(e,e)\}$ and for all $h \in H$, $l \in L$,

$$(e,l)(h,e)(e,l)^{-1} = (\eta(l)h, e).$$

Theorem 6.2. *Let $H \times_\eta L$ be a semidirect product of topological groups H and L. If $H \times_\eta L$ belongs to [SIN], then it belongs to [WS] in the following two situations:*
1) H belongs to [WS] and L is compact;
2) H is compact and L belongs to [WS].

Proof. A straightforward consequence of Theorem 5.1.

Example 6.3. Even when H is abelian and L is compact, a semidirect product $H \times_\eta L$ does not necessarily belong to [WS]. For instance, let us equip $\mathbb{C}$ with the usual addition and the torus $\mathbb{T}$ ($\mathbb{T} = \{z \mid z \in \mathbb{C}, |z| = 1\}$) with its usual multiplication; let η be the homomorphism from $\mathbb{T}$ into the group $\mathcal{A}(\mathbb{C})$ of automorphisms of $\mathbb{C}$ defined by $\eta(t)z = tz$, $t, z \in \mathbb{C}$. Then, since $\eta(\mathbb{T})$ is infinite, the semidirect product $\mathbb{C} \times_\eta \mathbb{T}$ does not belong to [WS]: this is a consequence of the following proposition.

Proposition 6.4. *Let $\mathbb{R}^n \times_\eta K$ be a semidirect product, where K is a compact topological group. Then the following conditions are equivalent:*
1) *$\mathbb{R}^n \times_\eta K$ belongs to [WS].*
2) *$\mathbb{R}^n \times_\eta K$ belongs to [SIN].*
3) *$\eta(K)$ is finite.*

Proof. By Corollary 3.2, 1) implies 2) and by Theorem 6.2, 2) implies 1). Equivalence between 2) and 3) follows from Theorem 2.9 of [6]. ∎

Theorem 6.5. *Let G be a topological group. Let us suppose that $G = HL$ with H a normal compact subgroup of G and L a subgroup of G which belongs to [WS]. For all $g \in G$, let us denote by θ_g the automorphism $h \mapsto ghg^{-1}$ of H. Then if $\{\theta_l \mid l \in L\}$ is equicontinuous, G belongs to [WS].*

Proof. First note that $\{\theta_g \mid g \in G\}$ is an equicontinuous group of automorphisms of H. Indeed, $\{\theta_h \mid h \in H\}$ is equicontinuous (since H is compact), $\{\theta_l \mid l \in L\}$ is equicontinuous by hypothesis, and $\{\theta_g \mid g \in G\} = \{\theta_h \circ \theta_l \mid h \in H, l \in L\}$ (since $G = HL$).

To show that G is a WS-group, we shall make use of Lemma 4.1. Let V be a neighbourhood of e in G. It follows from what we have noted above that there exists a symmetrical neighbourhood $A \subset V \cap H$ of e in H which is invariant under all inner automorphisms of G. Since H is compact, there exists a finite and symmetrical subset F of H such that $H = AF$. Then we have $G = HL = AFL$ and $G = G^{-1} = (AFL)^{-1} = L^{-1}F^{-1}A^{-1} = LFA$. By applying Lemma 4.1 (with the group H of the lemma replaced by our group L), we can conclude that G belongs to [WS]. ∎

Corollary 6.6. *Let $H \times_\eta L$ be a semidirect product. Let us suppose that H is compact, that L belongs to [WS] and that $\eta(L)$ is equicontinuous. Then $H \times_\eta L$ belongs to [WS].*

Theorem 6.7. *Let G be an almost connected topological group. Let us suppose that $G = HL$, with H a normal compact subgroup of G and L a subgroup of G which belongs to [WS]. Then G belongs to [WS].*

Proof. Making use of 6.5, we only have to prove that $\{\theta_g \mid g \in L\}$ is an equicontinuous set of automorphisms of H.

Let us equip $\mathcal{A}(H)$ with the uniform convergence topology. Then $\mathcal{A}(H)$ is a topological group and the function $g \mapsto \theta_g$ is continuous from G to $\mathcal{A}(H)$ (see for instance [10]). Hence $\{\theta_g \mid g \in G_0\}$ is a connected subgroup of $\mathcal{A}(H)$. But according to a theorem of K. Iwasawa ([11] Theorem 1), the connected component of e in $\mathcal{A}(H)$ is contained in the group $\mathcal{I}(H)$ of all inner automorphisms of H. Then by applying Lemma 4.2 (with $G' = G_0$), we get that $\{\theta_g \mid g \in G\}$ is equicontinuous. ∎

Corollary 6.8. *Let H be a compact topological group and let L be an almost connected topological group which belongs to* [WS]. *Then every semidirect product $H \times_\eta L$ belongs to* [WS].

Proof. Let us put $G = H \times_\eta L$. By hypothesis, the quotients G/L' and L'/L'_0 are compact. Hence G/L'_0 is compact and a fortiori G/G_0 is compact. The conditions of Theorem 6.7 are satisfied. ∎

Corollary 6.9. *Let H be a compact topological group and let L be a locally compact almost connected group which belongs to* [SIN]. *Then any semidirect product $H \times_\eta L$ belongs to* [WS].

Proof. By Corollary 5.4, group L belongs to [WS]. Then the conclusion follows from 6.8. ∎

Remark 6.10. In [1], R. W. Bagley and J. S. Yang prove, under the hypotheses of 6.9, that $H \times_\eta L$ belongs to [SIN]. This result can be recovered as a consequence of 6.9 and 3.1.

Example 6.11 Even when H is compact and L abelian, a semidirect product $H \times_\eta L$ does not necessarily belong to [WS]. Let $H = \{-1, +1\}^{\mathbf{Z}}$ be the infinite direct product of copies of the multiplicative group $\{-1, +1\}$. Let $L = \mathbb{Z}$ equipped with the discrete topology. For each $p \in \mathbb{Z}$, let us define the automorphism $\eta(p)$ of H by $\eta(p)\big((x_n)_{n\in\mathbf{Z}}\big) = (x_{n+p})_{n\in\mathbf{Z}}$. A simple calculation shows that in the semidirect product $H \times_\eta L$ there can be no invariant neighbourhood of the identity smaller than $H \times \{0\}$. Thus $H \times_\eta L$ is not a SIN-group. Since it is locally compact, it does not belong to [WS] (by 3.1).

References

[1] Bagley, R. W., and J. S. Yang, *Locally invariant topological and semidirect products,* Proc. Amer. Math. Soc. (1) **93** (1985), 139–144.

[2] Berglund, J. F., and K. H. Hofmann, *Compact semitopological semigroups and weakly almost periodic functions,* Lecture Notes in Mathematics **42**, Springer-Verlag, Berlin, Heidelberg, New-York (1967).

[3] Berglund, J. F., H. D. Junghenn and P. Milnes, *Analysis on Semigroups,* Canadian Math. Soc., Wiley-Interscience Publication (1989).

[4] Dixmier, J., *Les C*-algèbres et leurs Représentations,* Gauthier-Villars, Paris (1969).

[5] Grosser, S., and M. Moskowitz, *On central topological groups,* Trans. Amer. Math. Soc. **127** (1967), 317–340.

[6] Grosser, S., and M. Moskowitz, *Compactness conditions in topological groups,* J. Reine Angew. Math. **246** (1971), 1–40.

[7] Grothendieck, A., *Critères de compacité dans les espaces fonctionnels généraux,* Amer. J. Math. **74** (1952), 168–186.

[8] Hansel, G., and J. P. Troallic, *On a class of weakly almost periodic mappings,* to appear in Semigroup Forum.

[9] Hewitt, E., and K. A. Ross, *Abstract Harmonic Analysis I,* Springer-Verlag, New-York (1963).

[10] Hochschild, G., *La structure des groupes de Lie,* Dunod, Editeur, Paris (1968).

[11] Iwasawa, K., *Some types of topological groups,* Annals of Math. (2) **50** (1948), 507–558.

[12] Ramsay, R. T., *Groups with equal uniformities,* Canad. J. Math. **21** (1969), 655–659.

[13] Ruppert, W., *Compact Semitopological Semigroups: An Intrinsic Theory,* Lecture Notes in Mathematics **1079**, Springer-Verlag, Berlin, Heidelberg, New-York (1984).

Faculté des Sciences
Université de Rouen
F-76134 Mont-Saint-Aignan
France
GH@FRUNIP11

Faculté des Sciences et des Techniques
Université du Havre
25, rue Philippe Lebon
F-76600 Le Havre
France

Received 20 July 1990
and in final form 2 October 1990

Semigroup Forum Vol. 45 (1992) 71–76

RESEARCH ARTICLE

On the Linear Orderability of Two Classes of Finite Semigroups*

Kalčo Todorov

Communicated by J. S. Ponizovskǐ

1. Introduction and Basic Remarks

As it is known, any finite non-trivial group is not linearly (totally) orderable. At the same time as shown by Gabovič in [2] and by Zibina in [3] (see [1], p. 150) there are (also finite) semigroups, which are orderable with respect to every admissible order in them. The question naturally arising in this situation "when a (finite) semigroup is totally orderable", was formulated for the first time by Schein in [8]. The answer to the similar more general problem for groupoids, namely, when a groupoid is linearly orderable, was given after some virtuous investigations by Gabovič in [6]. Using the methods of proving of those namely investigations Krasimir Iordjev and the present author succeeded to show in [4] that any finite, or, more generally, any periodic, semigroups S is linearly orderable iff there exist linear orders $<$ and $<_1$, respectively in the set E of its idempotents and in every torsion class K_e, $e \in E$ such that

$$L(x)L(y) \subseteq L(xy) ,$$

for any $x, y \in S$, where for $x \in K_e$

$$L(x) \stackrel{d}{=} \left\{ \bigcup_{f<e} K_f \right\} \cup \{ y \in K_e \mid y \leq x \} .$$

The above described total test of linear orderability of semigroups does not have any information about construction of the orderable semigroup, for example about its generating elements and its defining relations, i.e. for the presentations of the torsion classes of linearly orderable semigroups, a problem studied at an early stage in the history of the theory of semigropus (see [5], 1.4).

Our investigations are also directed to this effect. It is shown – Theorem 1: any (finitely generated) periodic semigroup with defining relations (1) and (2) is right linear orderable, but it is not linear orderable. There exists a right order about which the semigroup S is left positively orderable. Theorem 2: any finite semigroup with generating relations (1) and (3) is positively linearly orderable.

The similar problem for bands is comprehensively studied in detail by Merlier in [7].

With respect to the terminology and the notion, we have been guided by the excellent (and still unique in this field) review of the contemporary state of the Ja. Gabovič in [1] which has noticeably influenced the level of our knowledge in this field. Following the terminology of Gabovič [1], an element x of a linear ordered semigroup S is called *right (left) positive* if $a \leq ax$ $(a \leq xa)$ for every element $a \in S$. A semigroup S is called *positively ordered*, if each of its elements is simultaneously right and left positive.

---

* This research was partially supported by the Committee of Science at the Council of Ministers of Bulgaria under Contract No. II.28/1989.

2. Basic Results

The fact that every finite non-trivial group is not linearly orderable shows that the kernel of every linearly orderable finite (or even periodic) semigroup S is trivial. Therefore, every linearly orderable periodic semigroup S is combinatorial, i.e.

$$(\forall\ x \in S)(\exists\ m(x) \in \boldsymbol{N})(x^{m(x)+1} = x^{m(x)})\ .$$

The converse is not true – not every finite combinatorial semigroup is linearly orderable.

Even more, since every band is combinatorial, there exist (Theorem 1-2 [7]) an infinite set of linearly unorderable combinatorial semigroups.

In this direction of interest is the following:

Theorem 1. *Any (finitely generated) periodic semigroup $S = \langle x_1, x_2, \ldots, x_s\rangle$ for $s > 1$, with defining relations*

(1) $x_i = x_j x_i$ *for* $1 \le i < j \le s$,

(2) $x_i^{m_i+1} = x_i^{m_i}$ *for* $m_i \in \boldsymbol{N}$

is right linearly orderable but it is not linearly orderable. There exists a right order about which the semigroup S is left positively orderable.

Proof. We shall prove that the semigroup S could be not linearly orderable. For this aim, it is sufficient to show that there exists (two-generated) its subsemigroup, which could be not linearly orderable. Suppose, by way of contradiction, that the semigroup S is linearly ordered relatively to its order $<$ and that in the subsemigroup $T = \langle x_1, x_2\rangle$ of S –

$$x_2 < x_1\ . \tag*{/1/}$$

Then the inequality /1/ multiplied (for $i, j, k, l \in \boldsymbol{N}$),

$$\left.\begin{array}{llll} \text{on the right by} & x_1^i & \text{implies} & x_1^i \le x_1^{i+1}\ ; \\ \text{on the right by} & x_2^k & \text{implies} & x_2^{k+1} \le x_1 x_2^k\ ; \\ \text{on the left by} & x_2^j & \text{implies} & x_2^{j+1} \le x_1\ ; \\ \text{on the left by} & x_1^l & \text{implies} & x_1^l x_2 \le x_1^{l+1}\ . \end{array}\right. \tag*{/2/}$$

The last inequalities and the assumption that the semigroup $S(<)$ is linearly ordered show that either

$$1)\qquad x_2 < x_2^2 < \ldots < x_2^{m_2} < x_1 < x_1^2 < \ldots < x_1^{m_1}\ ,$$

or

$$2)\qquad x_2^{m_2} < x_2^{m_2-1} < \ldots < x_2 < x_1 < x_1^2 < \ldots < x_1^{m_1}\ .$$

In the case 1), on the one hand, the inequality $x_2^u < x_2^{u+1}$ multiplied on the left by $x_1^{m_1}$ implies the inequalities

$$x_1^{m_1} x_2^u \le x_1^{m_1} x_2^{u+1} \qquad (u = 1, 2, \ldots, m_2)\ ,$$

and, on the other hand, the last inequality of /2/ multiplied on the right by x_2^u for $1 = m_1$ implies

$$x_1^{m_1} x_2^{u+1} \le x_1^{m_1} x_2^u\ .$$

Consequently, since $x_1^{m_1}x_2^{u+1} \neq x_1^{m_1}x_2^u$ for $u \geq 1$, the linear order $<$ of the semigroup S has as its corollaries two incompatible inequalities.

Case 2) is reduced to the following two subcases:

$$\text{i)} \quad x_1 < x_1\, x_2^{m_2} \quad \text{and} \quad \text{ii)} \quad x_1\, x_2^{m_2} < x_1\ .$$

Let us consider each one of these two cases separately.

i) The inequality $x_1 < x_1x_2^{m_2}$ implies the inequalities $x_1x_2^u \leq x_1x_2^{m_2}$ for $u = 1, 2, \ldots, m_2$. At the same time, from $x_2^{m_2} \leq x_2^u$, it follows that $x_1x_2^{m_2} \leq x_1x_2^u$.

Consequently, the linear order $<$ of the semigroup S has as its corollaries two incompatible inequalities.

ii) Note, that since the inequality $x_1^2x_2^{m_2} < x_1$ multiplied on the right by x_1 implies the incompatible with the considered order inequality $x_1^3 \leq x_1^2$, the inequality $x_1 < x_1^2\, x_2^{m_2}$ remains to hold.

The last inequality multiplied on the left by $x_1^{m_1-1}$ implies the inequality $x_1^{m_1} \leq x_1^{m_1}x_2^{m_2}$.

At the same time, the inequality $x_1x_2^{m_2} < x_1$ multiplied on the left by $x_1^{m_1-1}$ implies $x_1^{m_1}x_2^{m_2} \leq x_1^{m_1}$.

Consequently, in general we reach to the incompatible inequalities for both cases.

Hence, the semigroup S is not linearly orderable.

Before proceeding to the proof of the second part of the statement, we note that every element $a \in S$ is expressible by the generating relations (1) and (2) in the form

$$a = x_1^{u_1}x_2^{u_2}\ \ldots\ x_s^{u_s}\ , \qquad /3/$$

where $0 \leq u_1 \leq m_1$ for $i = 1, 2, \ldots, s$.

Besides, if the element a is such that $u_1 = \ldots = u_{p-1} = 0$ and $u_p \neq 0$, it can be seen directly that

$$a^2 = x_p^{2u_p}x_{p+1}^{u_{p+1}}\ \ldots\ x_s^{u_s}\ , \qquad /4/$$

and

$$a^{m_p+1} = a^{m_p}\ . \qquad /5/$$

From the equality /5/ it follows that for $s < \infty$ the semigroup S will be finite.

Further, we shall denote throughout by $<$ the (lexicographic) order on the semigroup S, defind by the rule

$$a = x_1^{u_1}\ x_2^{u_2}\ \ldots\ x_s^{u_s} < x_1^{v_1}\ x_2^{v_2}\ \ldots\ x_s^{v_s} = b\ , \qquad /6/$$

if either $u_1 < v_1$ or there exists a natural number $1 : 1 < 1 \leq n$, such that

$$u_1 = v_1, \ldots, u_{l-1} = v_{l-1}, \quad \text{and} \quad u_1 = v_1\ .$$

It is known that the introduced order is linear. We shall show that with respect to it, the semigroup S is right ordered, i.e. that $ac \leq bc$ for every element $c \in S$.

Indeed, let $c = x_t$ and let a be in the form of /3/. Then either $t < l$ and

$$ac = bc = x_1^{u_1} x_2^{u_2} \ \ldots \ x_{t-1}^{u_{t-1}} x_t^{u_t+1} \ ,$$

or $t = l$ and

$$ac = x_1^{u_1} \ \ldots \ x_l^{u_l+1} \le x_1^{v_1} \ \ldots \ x_l^{v_l+1} = bc \ ,$$

or $t > l$ and

$$ac = x_1^{u_1} \ \ldots \ x_l^{u_l+1} \ \ldots \ x_t^{u_t+1} < x_1^{v_1} \ \ldots \ x_l^{v_l+1} = bc \ .$$

Thus, we proved that the introduced order is right compatible with respect to every one of the generalizing elements of the semigroup S. Therefore, it will be compatible with respect to every element of the semigroup S.

We shall prove now the validity of the inequality $xa \ge a$ for any one of the generating elements x_t of the semigroup S and for every element a from /3/, where $u_1 = \ldots = u_{p-1} = 0$ and $u_p \ne 0$. Indeed, setting $x = x_t$, we have

(i) $t < p \Rightarrow x_t a > a$;

(ii) $p = t \Rightarrow x_p^{u_p+1} x_{p+1}^{u_{p+1}} \ \ldots \ x_s^{u_s} \ge a$;

and $xa = a$ in case that $x_p^{u_p+1} = x_p^{u_p}$;

(iii) $p < t \Rightarrow x_t a = a$, i.e. $x_t a \ge a$. ■

The so considered order is not right positive since for the element $a = x_1^{m_1} x_2^{m_2}$ we have

$$ax_1 = x_1^{m_1} < x_1^{m_1} x_2^{m_2} = a \ ,$$

whereas in every right positive order by definition $ax \ge a$ for every element $x \in S$. ■

Theorem 2. *Any finite semigroup $S = \langle x_1, \ldots x_s \rangle$ for $s > 1$, with generating relations*

(1) $x_i = x_j x_i$ *for* $1 \le i < j \le s$;

(3) $x_i^{m_i} x_j = x_i^{m_i}$ *for* $1 \le i \le j \le s$ *and* $m_i \in \boldsymbol{N}$,

is positively linearly orderable.

Proof. By relations (1) and (3), every elemen $a \in S$ is expressible in the form (see /3/)

$$a = x_1^{u_1} x_2^{u_2} \ \ldots \ x_q^{u_q} \ , \qquad /7/$$

where $0 \le u_i < m_i$ for $i = 1, \ldots, q-1$ and $0 \le u_q \le m_q$.

Further, from condition (3) it follows that

$$x_i^{m_i} x_j^k = x_i^{m_i} \quad \text{for} \quad k = 1, 2, \ldots ,$$

and consequently by $i = j$ we reach condition (2) of Theorem 1. With some elementary additional reasoning to Theorem 1, it can be seen that the semigroup S is right orderable with respect to its lexicographic order /6/ too.

With respect to the order introduced in this way, the semigroup S is also left ordered, i.e. for every element $c \in S$ and for every inequality of the kind /6/, it follows that $ca \le cb$. To this purpose it is enough to show that the proposition holds for every one of the generating elements of the semigroup S.

Let $c = x_t$ and let in /6/ $u_1 = v_1 = 0, \ldots, u_{p-1} = v_{p-1} = 0$ and $0 < u_p = v_p$ for some $p \le l$.

Then either

$$ca = a < b = cb \ , \quad \text{if} \quad 1 \le p < t \ ,$$

or

$$ca = x_t x_p^{u_p} \ \dots \ x_s^{u_s} \leq x_t x_p^{v_p} \ \dots \ x_s^{v_s} = cb \ , \quad \text{if} \quad t \leq p \ .$$

In particular,

$$ca = x_t x_p^{m_p - 1} \ \dots \ x_s^{u_s} = x_p^{m_p} \ \dots \ x_s^{u_s} = x_p^{m_p} = cb$$

if either $t = p < l$ and $u_p = v_p = m_p - 1$ or $t = p = 1$ and $u_p = m_l - 1 < v_l = m_l$.

Also, the semigroup S is a linear ordered semigroup. Now, we shall show that the semigroup S with respect to the above introduced order is positively orderable.

Indeed, by Theorem 1, the semigroup S is left positively orderable. It is also right positively orderable, since for every element x_t by the expression /7/, we obtain

$$ax_t > a \ , \quad \text{if} \ \ t < q$$

and

$$ax_t \geq a \ , \quad \text{if} \ \ q \leq t \ ,$$

whereupon

$$ax_t = a \ , \quad \text{for} \ \ u_q = m_q \ .$$

This completes the proof.

In [1], p. 142, Gabovič gives tables of all essentially different linearly ordered semigroups of order 4. A direct comparison shows that at least 11 of these semigroups satisfy the conditions of Theorem 2. These are the semigroups with the following numbers of rows and columns of their situation: 1.11; 1.12; 2.2; 2.4; 2.8; 2.9; 6.12; 8.2; 8.4; 8.6; 8.8.

At last the author would like to express his special thanks to E. Gabovič, whose critical notes contributed considerably to the final completion of the results in their present form.

References

[1] Gabovič, E. Ya., *Totally ordered semigroups and their applications* (Russian) Uspehi Mat. Nauk **31** (1976), no. 1(187), 137–201.

[2] Gabovič, E. Ya., *Ordered semigroups*, (Russian) Authorreferat, Leningrad 1968, 198–199.

[3] Zybina, L. D., *Semigroups in which every ordering is two-sided stable*, Leningrad. Gos. Ped. Inst. Ucen. Zap. **302** (1976), 70–76.

[4] Iordjev, Kr. Ya., K. J. Todorov, *On the totally orderability of finite semigroups*, Mathematics and Education in Mathematics, 1985, 258–262 (Bulgarian, English summary).

[5] Lallement, G., *Semigroups and Combinatorial Applications*, John Wiley and Sons, New York, 1979, 1–440.

[6] Gabovič, E., *The orderability of some classes of groupoids*, Semigroup Forum **9** (1974), 139–154.

[7] Merlier, T., *Sur les 0-bandes finies et les demi-groupes totalement ordonne 0-simples*, Semigroup Forum **4:2** (1972), 124–149.

[8] Schein, B. M., *Problem* 8, Semigroup Forum **1** (1970), 90–92.

[9] Bijev, G., K. Todorov, *Idempotent-generated subsemigroups of the symmetric semigroup of degree four; computer investigations*, Semigroup Forum **31** (1985), 119–122.

Bulgarien Academy of Sciences
Mathem. Institute
1090 Sofia, P. O. Box 373

Received December 3, 1990
and in final form July 31, 1991

Semigroup Forum Vol. 45 (1992) 77–91

RESEARCH ARTICLE

Some Remarks on the Two–dimensional Dirac Equation

Adam Bobrowski*

Communicated by J. A. Goldstein

Introduction

The paper deals with the two–dimensional Dirac equation. It falls into two parts. In the first part we give a simple proof of the probabilistic formula for solutions of the Dirac equation due to Ph. Blanchard, Ph. Combe, M. Sirugue and M. Sirugue–Colin. (See [1–4]. The other proof of this formula was given by T. Zastawniak [15].) Our proof is based on M. Kac and J. Kisyński's ideas ([11],[12]) which are as follows.

Consider the Cauchy problem

$$\begin{aligned} x''(t) + 2ax'(t) &= v^2 Ax(t) \ , \\ x(0) = x_0 \quad &, \quad x'(0) = x_1 \end{aligned} \tag{0°}$$

where A is the infinitesimal generator of a cosine operator family acting in a certain Banach space $\mathbf{E}$; x is $\mathbf{E}$–valued function and a, v are positive constants. The solution of this problem can be obtained by putting random time

$$\xi(t) = \int_0^t (-1)^{N_a(s)}\, ds \ ,$$

where $N_a(t)$ is a homogenous Poisson process with the mean value $EN_a(t) = at$. The solution is given by the formula

$$x(t) = E\widetilde{x}\Big(v\int_0^t (-1)^{N_a(s)}\, ds\Big) \ ,$$

where $\widetilde{x}$ is the solution of the following Cauchy problem:

$$\begin{aligned} x''(t) &= Ax(t) \ , \\ x(0) = x_0 \quad &, \quad x'(0) = x_1 \ . \end{aligned}$$

An interesting proof of this fact, due to J. Kisyński (see [9] p.13 and [12]), is based on the fact that

$$\Big(\int_0^t (-1)^{N_a(s)}\, ds \ , \ (-1)^{N_a(t)}\Big)_{t\geq 0}$$

is the time–homogeneous process with independent left increments in the group $\mathcal{G} = \mathbf{R}\times\{-1,1\}$ with the multiplication rule

$$(\xi,k)\circ(\eta,l) = (\xi l + \eta, kl) \ , \quad \xi,\eta\in\mathbf{R} \ , \ k,l=\pm 1 \ .$$

---

* The author wishes to thank Prof. J. Kisyński for drawing his attention to this problem and for helpful, stimulating discussions.

A similar approach, but with no group–theoretical aspect is presented in [5] pp. 468–471, where the references are given to [7], [8].

The two–dimensional Dirac equation may be viewed as the equation of the form

$$x'(t) = Ax(t) + Bx(t) ,$$

where A is the infinitesimal generator of a one–parameter group $G(t)$, $t \in \mathbf{R}$ acting in a Banach space $\mathbf{E}$ and B is a bounded linear operator such that $BG(t) = G(-t)B$ for all $t \in \mathbf{R}$. The usage of the J. Kisyński method gives the probabilistic solution of the Dirac equation (Th. 1).

The authors of [1–4] have also proved a formula for the two–dimensional Dirac equation with potentials, which is similar to the Feynman–Kac one. (For another proof see [15]). To prove that this similarity is not accidental is the purpose of the second part of the paper. It was done in Theorem 2, in which the Feynman–Kac formula for equations connected with Markov processes with values in a certain group is proved. A particular case of these equations is the two dimensional Dirac equation.

Note that some explicit calculations of the distribution of the random variable $\int_0^t (-1)^{N_a(s)}\, ds$ related to the problem (0°) may be found in [10].

1. The Poisson–Kac process as a process with independent increments in a group

Let $N(t)$, $t \geq 0$ be a homogenous Poisson stochastic process with the mean value $EN(t) = t$. The stochastic variables

$$v(t) = (-1)^{N(t)} ,$$
$$\xi(t) = \int_0^t (-1)^{N(s)}\, ds$$

describe the velocity and position of a point of $\mathbf{R}^1$, respectively, which moves with the velocity equal to $+1$ or -1. The velocity changes randomly in such a manner, that a number of changes through the time interval from 0 to t is equal to $N(t)$; the point at the moment $t = 0$ is at the origin of $\mathbf{R}^1$ and has the velocity $+1$. The natural phase space for such a motion is the non–commutative group

$$\mathcal{G} = \Big(\mathbf{R} \times \{-1, 1\}, \circ\Big),$$

where "$\circ$" denotes multiplication defined by the formula (see [12])

$$(\xi, k) \circ (\eta, l) = (\xi l + \eta, kl) \quad \xi, \eta \in \mathbf{R},\ k, l = \pm 1 .$$

Let $\mathcal{G} \times \mathbf{Z}$ be the direct product of the group $\mathcal{G}$ and the group of integers $\mathbf{Z}$. The set

$$\mathcal{A} = \big\{(\xi, k, z) : k = (-1)^z\big\}$$

is a subgroup of $\mathcal{G} \times \mathbf{Z}$ (if $k = (-1)^z$ and $l = (-1)^w$ then $kl = (-1)^{z-w}$). This subgroup is isomorphic with the group

$$\mathcal{H} = (\mathbf{R} \times \mathbf{Z}, \,\cdot\,) ,$$
$$(\xi, z) \cdot (\eta, w) = [\, \xi(-1)^w + \eta, z + w \,] .$$

The mapping

$$I : (\xi, (-1)^w, w) \longrightarrow (\xi, w)$$

is an isomorphism.

The stochastic process g_t, $t \geq 0$ with values in $\mathcal{G}$ defined by

$$g_t = \Big(\int_0^t (-1)^{N(s)}\, ds, (-1)^{N(t)} \Big)$$

has the following properties (see [12]) :

(*) the random variables $g_{t+h}\, g_t^{-1}$ and g_t are independent,

(**) the random variables $g_{t+h}\, g_t^{-1}$ and g_h have the same distribution.

Let

$$\widetilde{g}_t = (g_t, N(t)) \ , \quad t \geq 0$$

be a stochastic process with values in $\mathcal{G} \times \mathbf{Z}$. We have :

$$\widetilde{g}_{t+h} \widetilde{g}_t^{-1} = \Big(g_{t+h}\, g_t^{-1}, N(t+h) - N(t) \Big) \ ,$$

then the process $\tilde{g}_t$ has the properties (*), (**), too. Thus the Borel measures on $\mathcal{H}$

$$\mu_t(\mathcal{B}) = P(h_t \in \mathcal{B}) \ ,$$

$$h_t = I \circ \widetilde{g}_t = \Big(\int_0^t (-1)^{N(s)}\, ds, N(t) \Big)$$

form a convolution semigroup, i.e.

$$\mu_t * \mu_s = \mu_{t+s} \ .$$

2. The probabilistic solution of the two–dimensional homogeneous Dirac equation

Let $G(t)$, $t \in \mathbf{R}$ be a one–parameter strongly continuous group of bounded linear automorphisms of a Banach space $\mathbf{E}$. Let B be a linear, bounded automorphism of this space; A the infinitesimal generator of the group G, and U the mapping from $\mathcal{H}$ into $L(\mathbf{E}, \mathbf{E})$ given by

$$U(\xi, k) = B^k G(\xi) \ .$$

Theorem 1. *Suppose that, for all $t \in \mathbf{R}$, the operator B satisfies condition*

$$G(t)B = BG(-t) \ . \tag{I}$$

Then the mapping U is a strongly continuous representation of the group $\mathcal{H}$ by linear automorphisms of $\mathbf{E}$. Furthermore, for $t \geq 0$, the one–parameter group of operators generated by the operator $A + B$ has the form :

$$T(t)x = \int_{\mathcal{H}} e^t U(g) x \mu_t(dg) \ , \quad x \in \mathbf{E} \ . \tag{*}$$

Proof. We have

$$U(\xi, z)U(\eta, w) = B^z G(\xi) B^w G(\eta) = B^{z+w} G\big[(-1)^w \xi + \eta \big] = U\big((\xi, z) \circ (\eta, w)\big) \ ,$$

hence the mapping U is a representation. The formula $(*)$ defines a strongly continuous semigroup of operators because the measures μ_t form a convolution semigroup, and because μ_t converges in distribution to the probabilistic measure concentrated in the point $(0,0)$, when $t \to 0$. Theorem 1 will follow if we show that the operator $A+B$ is the generator of the semigroup T.

Set

$$\mu_t^n(\mathcal{B}) = \mu_t(\mathcal{B} \times \{n\}) \ .$$

(Measures μ_t^n are Borel measures on the real line.)

Then

$$T(t)x = \sum_{n=0}^{\infty} \int_{\mathbf{R}} e^t B^n G(s)x \, \mu_t^n(ds) \tag{1°}$$

and

$$\begin{aligned}
&\mu_t^0(\mathcal{B}) = e^{-t}\delta(\mathcal{B} - t) \ , \\
&\operatorname{supp} \mu_t^n \subset \langle -t, t\rangle \ , \quad t \in \mathbf{R},\ n \in \mathbf{N} \ , \\
&\mu_t^n(\mathbf{R}) = P\big(N(t) = n\big) = \frac{t^n}{n!}\, e^{-t} \ .
\end{aligned}$$

Thus

$$\int_{\mathbf{R}} e^t G(s)x \, \mu_t^0(ds) = G(t)x \ , \tag{2°}$$

$$\Big\| \frac{1}{t} \sum_{n=2}^{\infty} \int_{\mathbf{R}} e^t B^n G(s)x \, \mu_t^n(ds) \Big\| \underset{t \to 0}{\longrightarrow} 0 \ .$$

Given $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ such that the inequality

$$\|G(s)x - x\| < \frac{\varepsilon}{\|B\|}$$

holds for $|s| < \delta(\varepsilon)$. As a consequence :

$$\begin{aligned}
&\Big\| \int_{\mathbf{R}} \frac{e^t BG(s)x \, \mu_t^1(ds)}{t} - Bx \Big\| = \Big\| \int_{\mathbf{R}} \frac{BG(s)x - Bx}{t} \, e^t \, \mu_t^1(ds) \Big\| \le \\
&\int_{\mathbf{R}} \|B\| \frac{\|G(s)x - x\|}{t} \, e^t \mu_t^1(ds) \le \varepsilon \ \text{ for } t \text{ small enough} \ .
\end{aligned} \tag{3°}$$

Applying 1°, 2°, 3° we see that

$$\lim_{t \to 0} \Big(\frac{T(t)x - x}{t} - \frac{G(t)x - x}{t} \Big) = Bx \ , \quad x \in E \ ,$$

which proves that the domain of the generator of the semigroup T is the same as the domain $\mathcal{D}(A)$ of the operator A, and that the generator of T is the operator $A+B$. Theorem 1 is proved.

Remark. The above proof shows that it is worth noting that the Poisson–Kac process has independent identically distributed left increments in the group $\mathcal{H}$. It suggests the use of the Markov property to obtain some other results concerning the Dirac equation. For example, it is possible, after easy algebraical calculations, to prove the so called "path–space measures iteration formula" (see [15]). The proof is omitted.

Examples. 1°. The two–dimensional Dirac equation.
Let

$$C_0 = \{u \; ; \; u : (-\infty, \infty) \to C \; , \; \lim_{x \to \pm\infty} u(x) = 0\} \; , \quad \mathbf{E} = C_0 \times C_0 \; ,$$

z be a complex number and A, B operators defined as follows

$$A : \mathcal{D}(A) \to \mathbf{E} \quad , \quad B : \mathbf{E} \to \mathbf{E}$$
$$\mathcal{D}(A) = \{(u, v) \in E \; ; \; u, v \text{ are differentiable}, u', v' \in C_0\} \; ,$$

$$A \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{du}{dx} \\ -\frac{dv}{dx} \end{bmatrix} \quad , \quad B \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} zv \\ zu \end{bmatrix} \; .$$

Then the operator A generates the one–parameter group

$$G(t) \begin{bmatrix} u \\ v \end{bmatrix} (x) = \begin{bmatrix} u(x+t) \\ v(x-t) \end{bmatrix} \; , \; t \in \mathbf{R} \; .$$

Since

$$BG(t) \begin{bmatrix} u \\ v \end{bmatrix} (x) = \begin{bmatrix} zv(x-t) \\ zu(x+t) \end{bmatrix} = G(-t)B \begin{bmatrix} u \\ v \end{bmatrix} (x) \; ,$$

the condition (I) is satisfied and the solution of the two–dimensional Dirac equation

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} + zv \; ,$$
$$\frac{\partial v}{\partial t} = -\frac{\partial v}{\partial x} + zu \; ,$$
$$u(0, x) = u(x) \quad , \quad v(0, x) = v(x) \quad , \quad x \in \mathbf{R} \; , \; t \geq 0$$

is given by the formula

$$\begin{bmatrix} u(t, x) \\ v(t, x) \end{bmatrix} = e^t E z^{N(t)} B^{N(t)} \begin{bmatrix} u(x + \xi(t)) \\ v(x - \xi(t)) \end{bmatrix} \; ,$$

where

$$\xi(t) = \int_0^t (-1)^{N(s)} \, ds \; ,$$

(E denotes the expectation value.) The above result is due to Ph. Blanchard, Ph. Combe, M. Sirugue, M. Sirugue–Collin (see [4], where an abundant bibliography is given, compare also [15]).

Remark. When $z = 1$ then the operator B fulfills, additionally, the condition $B^2 = \mathrm{Id}_{\mathbf{E}}$ and one can use lemma 1 of [12].

2°. The probabilistic formula for the solution of the equation :

$$\frac{\partial u(t, x)}{\partial t} = a \frac{\partial u(t, x)}{\partial x} + bu(t, -x) \; , \; u(0, x) = u_0(x)$$
$$a, b \in \mathbf{R} \quad , \quad t \geq 0 \quad , \quad x \in \mathbf{R}$$

is the following :

$$u(t,x) = E\ e^{t}b^{N(t)}u_0\left[(-1)^{N(t)}(x - a\xi(t))\right] .$$

3. A Feynman–Kac type formula

It is known that for the two–dimensional Dirac equation with potentials there exists a formula similar to the Feynman – Kac one. As we will see, it is a property of a fairly large class of differential equations, solutions of which take on a probabilistic form. To prove this we need the proposition which we will show below.

Let an operator $A : \mathcal{D}(A) \to \mathbf{E}$ be the generator of a strongly continuous semigroup $T(t)$, $t \geq 0$ and suppose that, for some $\lambda \in \mathbf{R}$ and for every $t \geq 0$, the inequality

$$\|T(t)\| \leq e^{\lambda t}$$

holds. Furthermore, let $V(t) \in L(\mathbf{E},\mathbf{E})$, $t \in \langle 0,\Theta\rangle$ be a strongly continuous family of operators such that, for every $x \in \mathcal{D}(A)$, the function $t \to V(t)x$ is strongly continuously differentiable. It is well known (see for example [13] ch.2, Theorems 3.6, 3.7) that the family $A + V(t)$, $t \in \langle 0,\Theta\rangle$ generates the evolution operator $S(t,s)$, $\Theta \geq t \geq s \geq 0$ associated with the well–posed initial value problem. Let $R(t,s)$, $\Theta \geq t \geq s \geq 0$ denote the evolution operator generated by the family $V(t)$, $t \in \langle 0,\Theta\rangle$.

Proposition. (The Lie – Trotter formula). *Under the above assumptions we have*

$$S(t,s)x = \lim_{n\to\infty} S_n(t,s)x \ ,$$

$$S_n(t,s) = T(\Delta_n)R(t,t-\Delta_n)T(\Delta_n)R(t-\Delta_n,t-2\Delta_n)\ldots T(\Delta_n)R(s+\Delta_n,s) \ ,$$

where $x \in \mathbf{E}$, $\Delta_n = \dfrac{t-s}{n}$.

Proof. Since the proof follows the pattern of the proof of the Lie – Trotter formula for semigroups (see Goldstein [6] or Reed and Simon [14]), we give it in the appendix.

Let $\mathbf{E}$ be a commutative Banach algebra, and let U be a strongly continuous representation of a locally compact group $\mathcal{G}$ by linear continuous automorphisms of $\mathbf{E}$. Furthermore, let g_t, $t \geq 0$ be a stochastic process with values in the group $\mathcal{G}$. Suppose that this process has the properties (*), (**) of Paragraph 1 and, additionally, that sample paths of it have left limits and are right continuous (it is a natural condition -see [5] ch.4.2 Proposition 2.4). Finally, let A denote the generator of the semigroup $T(t)$, $t \geq 0$ defined by the formula :

$$T(t)x = E\ U(g_t)x = \int_{\mathcal{G}} U(g)x\ \mu_t(dg)$$

and let us assume that, for some $\lambda \in \mathbf{R}$ and for every $t \geq 0$, the inequality

$$\|T(t)\| \leq e^{\lambda t} \tag{II}$$

holds.

Theorem 2 (The Feynman – Kac formula). *Suppose that W is a strongly continuous representation of the group $\mathcal{G}$ by linear continuous automorphisms of* $\mathbf{E}$ *such that :*

$$\|W(g)\| \le M \quad , \quad g \in \mathcal{G} \, , \tag{$*$}$$

and

$$U(g)(x \circ y) = W(g)x \circ U(g)y \quad , \quad W(g)(x \circ y) = W(g)x \circ W(g)y \tag{$**$}$$
$$x, y \in \mathbf{E} \quad , \quad g \in \mathcal{G} \quad , \quad \text{“}\circ\text{” denotes multiplication in } \mathbf{E} \, .$$

(The “$\circ$” will be omitted below.)

Assume that a function $V : \langle 0, \Theta \rangle \to \mathbf{E}$ is strongly continuously differentiable. Then the evolution operator generated by the equation

$$\frac{dx}{dt} = Ax + V(t)x \quad , \quad t \in \langle 0, \Theta \rangle$$

is given by :

$$S(t,s)x = E \exp\Big[\int_s^t W(g_{t-r})V(r)\,dr\Big] U(g_{t-s})x \, . \tag{III}$$

Remark. Continuity of the derivative of V guarantees that the evolution operator S is associated with the well–posed initial value problem. (see [13] ch.2, theorems 3.6, 3.7)

Proof. From the assumption concerning continuity of the function V and inequality (II) it follows that we can use the Lie–Trotter formula. Thus

$$S(t,s)x = \lim_{n \to \infty} S_n(t,s)x \; ,$$
$$S_n(t,s) = T(\Delta_n)R(t, t-\Delta_n)T(\Delta_n)R(t-\Delta_n, t-2\Delta_n)\ldots T(\Delta_n)R(s+\Delta_n, s) \; ,$$

where $x \in \mathbf{E}$, $\Delta_n = \frac{t-s}{n}$ and $R(t,s)$, $\Theta \ge t \ge s \ge 0$ is the evolution operator generated by the equation

$$\frac{dx}{dt} = V(t)x \quad , \quad t \in \langle 0, \Theta \rangle \; .$$

The algebra $\mathbf{E}$ is commutative, hence

$$R(t,s)x = \Big(\exp\Big[\int_s^t V(r)\,dr\Big]\Big)x \; . \tag{1°}$$

Finally, the property $(**)$ together with the commutativity of $\mathbf{E}$ implies

$$W(g)(e^y) = W(g)\sum_{n=0}^{\infty} \frac{y^n}{n!} = \sum_{n=0}^{\infty} \frac{[W(g)y]^n}{n!} = e^{W(g)y} \; . \tag{2°}$$

To avoid mistakes let us admit the following convention. For operators A, B and vectors x, y, z the value of the expression $BzAyx$ is calculated in the following way : we multiplicate vectors x and y, then we take the value of A on xy, multiplicate it by z and finally take the value of B on $zAyx$. The value of $Bz(Ay)x$ is calculated by the multiplication of vectors x, Ay, z and taking the value of B on this product. With such a convention, according to the property $(**)$, for every $g_i \in \mathcal{G}$, $x, y_i \in \mathbf{E}$, $i = 1, 2, \ldots, n$, we have

$$U(g_n)y_nU(g_{n-1})y_{n-1}\ldots U(g_1)y_1x = \tag{3°}$$
$$= [W(g_n)y_n]\,[W(g_ng_{n-1})y_{n-1}]\ldots[W(g_ng_{n-1}\ldots g_1)y_1][U(g_ng_{n-1}\ldots g_1)x].$$

Indeed, when $n = 1$, the above formula is equivalent to the first equation in $(**)$. Moreover, we have, by induction:

$$\begin{aligned}
&U(g_{n+1})y_{n+1}[U(g_n)y_n\ldots U(g_1)y_1x] = \\
&U(g_{n+1})y_{n+1}[W(g_n)y_n][W(g_ng_{n-1})y_{n-1}]\ldots[W(g_ng_{n-1}\ldots g_1)y_1] \\
&\quad[U(g_ng_{n-1}\ldots g_1)x] = \\
&W(g_{n+1})\{y_{n+1}[W(g_n)y_n]\,[W(g_ng_{n-1})y_{n-1}]\ldots \\
&\quad[W(g_ng_{n-1}\ldots g_1)y_1\}[U(g_{n+1})U(g_ng_{n-1}\ldots g_1)x] = \\
&[W(g_{n+1})y_{n+1}]\,[W(g_{n+1}g_n)y_n]\,[W(g_{n+1}g_ng_{n-1})y_{n-1}]\ldots \\
&\quad[W(g_{n+1}g_n\ldots g_1)y_1]\,[U(g_{n+1}g_n\ldots g_1)x].
\end{aligned}$$

(We have used the first equation in $(**)$ and then $n+1$–times the second one, together with the fact that U, W are representations of the group $\mathcal{G}$.) So (3°) is proved.

For the time being let us fix $x \in \mathbf{E}$, $t, s \in \langle 0, \Theta\rangle$, $n \in \mathbf{N}$ and set $V_i = \int\limits_{s_{i-1}}^{s_i} V(u)\,du$, $s_i = s + i\Delta$, $i = 1, 2, \ldots, n$, $\Delta = \frac{t-s}{n}$. We have, by (1°) and the definition of the semigroup $T(t)$, $t \geq 0$,

$$\begin{aligned}
&S_n(t,s) = T(\Delta)e^{V_n}\ldots T(\Delta)e^{V_1}x = \\
&\int_\Omega\cdots\int_\Omega U(g_\Delta(\omega_n))e^{V_n}\ldots U(g_\Delta(\omega_1))e^{V_1}x\,dP(\omega_1)\ldots dP(\omega_n)\,.
\end{aligned}$$

By virtue of (3°), the integrand equals

$$\begin{aligned}
&[W(g_\Delta(\omega_n))e^{V_n}][W(g_\Delta(\omega_n)g_\Delta(\omega_{n-1}))e^{V_{n-1}}]\ldots \\
&[W(g_\Delta(\omega_n)g_\Delta(\omega_{n-1})\ldots g_\Delta(\omega_1))e^{V_1}] \\
&[U(g_\Delta(\omega_n)g_\Delta(\omega_{n-1})\ldots g_\Delta(\omega_1))x]\,.
\end{aligned}$$

Applying (2°) and the commutativity of the algebra $\mathbf{E}$ we see that

$$\begin{aligned}
S_n(t,s)x = &\int_\Omega\cdots\int_\Omega \exp[W(g_\Delta(\omega_n))V_n + W(g_\Delta(\omega_n)g_\Delta(\omega_{n-1}))V_{n-1}+ \\
&+\cdots+W(g_\Delta(\omega_n)g_\Delta(\omega_{n-1})\ldots g_\Delta(\omega_1))V_1] \\
&U(g_\Delta(\omega_n)g_\Delta(\omega_{n-1})\ldots g_\Delta(\omega_1))x\,dP(\omega_1)\ldots dP(\omega_n)\,.
\end{aligned}$$

The stochastic process g_t, $t \geq 0$ fulfills the conditions $(*)$, $(**)$ from Paragraph 1, so we deduce that the above expression is the expectation value of the random vector

$$F_n(t,s) = U(g_{t-s})x\,\exp\Big[\sum_{i=1}^n W(g_{t-s_{i-1}})\int_{s_{i-1}}^{s_i} V(r)\,dr\Big]\,.$$

When n tends to infinity, $F_n(t,s)$ tends almost surely to the random vector:

$$F(t,s) = U(g_{t-s})x \exp\left[\int_s^t W(g_{t-r})V(r)\,dr\right] .$$

Indeed, the representation W is bounded, thus

$$\left\|\sum_{i=1}^{n} W(g_{t-s_{i-1}}) \int_{s_{i-1}}^{s_i} V(r)\,dr - \Delta \sum_{i=1}^{n} W(g_{t-s_{i-1}})V(s_{i-1})\right\| \le$$
$$M \sum_{i=0}^{n-1} (s_{i+1} - s_i)\omega(\Delta) = M(t-s)\omega(\Delta) ,$$

where $\omega(\ )$ is modulus of continuity of the function $V(\ )$, and M is the constant, which was defined in $(*)$. But sample paths of our process are right continuous and have left limits, hence the function $r \to W(g_{t-r})V(r)$, $r \in \langle s,t\rangle$ is left continuous and has right limits, and therefore it is Riemann integrable. The expression

$$\Delta \sum_{i=1}^{n} W(g_{t-s_{i-1}})V(s_{i-1})$$

is its integral sum. Hence $F_n(t,s)$ tends to $F(t,s)$ a.s. By the Lebesque dominated convergence theorem and the inequality

$$\|F_n(t,s)x\| \le e^{ML(t-s)}\|U(h_{t-s})x\|$$
$$(\text{where } L \stackrel{\text{df}}{=} \sup_{s\le r\le t} \|V(r)\|\ ) ,$$

the theorem is proved.

4. Examples

1°. The two–dimensional Dirac equation.

In this example the group $\mathcal{G}$ and the process g_t, $t \ge 0$ discussed in Theorem 2 are replaced by the group $\mathcal{H}$ and the process h_t, $t \ge 0$ defined in Paragraph 1. We use the same notation as in Example 1 of the above mentioned paragraph. Let us define the multiplication in $\mathbf{E} = C_0 \times C_0$ by :

$$(u,v)\,(f,g) = (uf, vg) .$$

In such a manner $\mathbf{E}$ becomes a commutative Banach algebra. By virtue of the definition of the Poisson process, the semigroup

$$T(t)\begin{bmatrix} u \\ v \end{bmatrix}(x) = E\, e^t z^{N(t)} B^{N(t)} \begin{bmatrix} u(x+\xi(t)) \\ v(x-\xi(t)) \end{bmatrix}$$

satisfies the condition

$$\|T(t)\| \le e^{|z|t} .$$

Furthermore the representation U of the group $\mathcal{H}$ (see Paragraph 1) defined by

$$U(\xi,k)\begin{bmatrix} u \\ v \end{bmatrix}(x) = z^k B^k \begin{bmatrix} u(x+\xi) \\ v(x-\xi) \end{bmatrix}$$
$$B\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} v \\ u \end{bmatrix}$$

satisfies the condition (**) of Theorem 2 with respect to the multiplication in $\mathbf{E}$, where the representation W is given by the formula

$$W(\xi,k)\begin{bmatrix} u \\ v \end{bmatrix}(x) = B^k \begin{bmatrix} u(x+\xi) \\ v(x-\xi) \end{bmatrix} .$$

Indeed

$$U(\xi,k)\begin{bmatrix} uf \\ vg \end{bmatrix}(x) = z^k B^k \begin{bmatrix} u(x+\xi)f(x+\xi) \\ v(x-\xi)g(x-\xi) \end{bmatrix} =$$
$$W(\xi,k)\begin{bmatrix} u \\ v \end{bmatrix}(x)\, U(\xi,k)\begin{bmatrix} f \\ g \end{bmatrix}(x) .$$

The rest of assumptions of Theorem 2 is clearly satisfied: for every $h \in \mathcal{H}$ the inequality

$$\|W(h)\| \le 1$$

holds and, furthermore, the points of discontinuity of the process h_t, $t \ge 0$ are exactly the points of jumps of the Poisson process.

So, if the function $V : \langle 0,\Theta\rangle \to \mathbf{E}$, $V(s) = \begin{bmatrix} G(s) \\ H(s) \end{bmatrix}$ is strongly continuously differentiable, the solution of the equation

$$\begin{aligned}
&\frac{\partial u}{\partial t} = \frac{\partial u}{\partial xt} + zv + G(t)u \ , \\
&\frac{\partial v}{\partial t} = -\frac{\partial v}{\partial x} + zu + H(t)v \ , \\
&u(s,x) = u(x) \quad , \quad v(s,x) = v(x) \quad , \quad x \in \mathbf{R} \quad , \quad \Theta \ge t \ge s \ge 0
\end{aligned}$$

is given by :

$$\begin{bmatrix} u(t,x) \\ v(t,x) \end{bmatrix} = E\, e^{t-s} z^{N(t-s)} B^{N(t-s)} \begin{bmatrix} u[x+\xi(t-s)] \\ v[x-\xi(t-s)] \end{bmatrix}$$
$$\exp\Big[\int_s^t B^{N(t-r)} \begin{bmatrix} G[r,x+\xi(t-r)] \\ H[r,x-\xi(t-r)] \end{bmatrix} dr\Big] ,$$

(E denotes the expectation value) where

$$\xi(t) = \int_0^t (-1)^{N(s)}\, ds .$$

The above formula was originally obtained in [1], [2], see also [15].

2°. The heat equation.

The operator $\frac{1}{2}\frac{\partial^2}{\partial x^2}$ with the domain $\mathcal{D}\left(\frac{1}{2}\frac{\partial^2}{\partial x^2}\right) = \{f \in C_0 \ ; \ f'' \in C_0\}$ is the generator of the strongly continuous semigroup $T(t)$, $t \geq 0$; $T(t) : C_0 \to C_0$,

$$T(t)f = E\, U[w(t)]f \ ,$$

where $w(t)$ is a Wiener process, and the operator $U(t)$ is given by

$$U(t)f(x) = f(x+t) \ .$$

Since, for all $f, g \in C_0$,

$$U(t)fg = U(t)f\, U(t)g \ ,$$

the representation W is again U. The operators $U(t)$, $t \in \mathbf{R}$, are contractions, the sample paths of the Wiener process are continuous and, as a consequence, the assumptions of Theorem 2 are fulfilled. Formula (III) gives the well-known Feynman – Kac formula. For every $s \geq 0$ the function

$$u(t,x) = S(t,s)f(x) = E\ \exp\left[\int_s^t V[r, x + w(t-r)]\, dr\right] f[w(t-s)]$$

is the solution of

$$\begin{aligned} &\frac{\partial u(t,x)}{\partial t} = \frac{\partial^2 u(t,x)}{\partial x^2} + V(t,x)u(t,x) \quad , \quad \Theta \geq t \geq s \ , \ x \in \mathbf{R} \ , \\ &u(s,x) = f(x) \ . \end{aligned}$$

3°. The probabilistic formula for the equation :

$$\begin{aligned} &\frac{\partial u(t,x)}{\partial t} = a\frac{\partial u(t,x)}{\partial x} + bu(t,-x) + V(t,x)u(t,x) \ , \quad a \in \mathbf{R} \\ &u(s,x) = u(x) \ , \qquad x \in \mathbf{R} \ , \ \Theta \geq t \geq s \geq 0 \end{aligned}$$

is the following :

$$\begin{aligned} u(t,x) = E\ e^{t-s} b^{N(t-s)} u\left[(-1)^{N(t-s)}(x - a\xi(t-s))\right] \\ \exp\left[\int_s^t V[r, (-1)^{N(t-r)}(x - a\xi(t-r)]\, dr\right] . \end{aligned}$$

5. Appendix

Lemma 1. *Let us suppose that the family $A(t)$, $0 \leq t \leq \Theta$, of closed operators with the common domain $\mathcal{D}(A)$ generates an evolution operator $S(t,s)$, $\Theta \geq t \geq s \geq 0$. (The definitions we use are to be found in* [13] ch2.*)*
Then

$$\lim_{\Delta \to 0} \sup_{s \leq t \leq \Theta - \Delta} \left\| \frac{S(s+\Delta, s)x - x}{\Delta} - A(s)x \right\| = 0 \quad , \quad x \in \mathcal{D}(A)$$

(it is uniform convergence of difference quotients).

Proof. We have

$$S(s+\Delta,s)x = x + \int_s^{s+\Delta} S(s+\Delta,p)A(p)x\,dp \quad , \quad x \in \mathcal{D}(A) \ ,$$

hence

$$\sup_{0\le s\le \Theta-\Delta}\left\|\frac{S(s+\Delta,s)x - x}{\Delta} - A(s)x\right\| =$$
$$\sup_{0\le s\le \Theta-\Delta}\left\|\frac{1}{\Delta}\int_s^{s+\Delta}\left[S(s+\Delta,p)A(p)x - A(s)x\right]dp\right\| \le$$
$$\sup_{\substack{0\le s\le \Theta-\Delta\\ s\le p\le s+\Delta}} \|S(s+\Delta,s)A(p)x - A(s)\| \ .$$

But $S(s+\Delta,s)x \underset{\Delta\to 0}{\longrightarrow} x$ uniformly in s on $\langle 0,\Theta\rangle$. Additionally, operators $S(t,s)$ are equibounded in $\Theta \ge t \ge s \ge 0$ and for a fixed $x \in \mathbf{E}$ the set $\{y \ ; \ \exists\ s \in \langle 0,\Theta\rangle \ , \ y = A(s)x\}$ is compact. This completes the proof.

The proof of Proposition of Paragraph 3. For all fixed operators A_i, B_i, $i = 1,2,\ldots,n$

$$\prod_{i=1}^n A_i - \prod_{i=1}^n B_i = \sum_{i=1}^n A_nA_{n-1}\ldots A_{i+1}(A_i - B_i)B_{i-1}\ldots B_1 \ ,$$

where, by definition, $B_0 = A_0 = I$. Therefore, putting $s_k = s + k\Delta_n$, we have

$$\|S_n(t,s)x - S(t,s)x\| = \|S_n(t,s)x - S(t,s_{n-1})S(s_{n-1},s_{n-2})\ldots S(s_1,s)x\| =$$
$$\Big\|\sum_{i=1}^n T(\Delta_n)R(t,s_{n-1})T(\Delta_n)R(s_{n-1},s_{n-2})\ldots T(\Delta_n)R(s_{i-1},s_{i-2})$$
$$[T(\Delta_n)R(s_i,s_{i-1}) - S(s_i,s_{i-1})]S(s_{i-1},s)x\Big\| \le$$
$$ne^{M(t-s)} \sup_{s\le\tau\le t-\Delta_n} \|\ [T(\Delta_n)R(\tau+\Delta_n,\tau) - S(\tau+\Delta_n,\tau)]\ S(\tau,s)x\| =$$
$$(t-s)e^{M(t-s)} \sup_{s\le\tau\le t-\Delta_n}\left\|\frac{1}{\Delta_n}\,[T(\Delta_n)R(\tau+\Delta_n,\tau) - S(\tau+\Delta_n,\tau)]S(\tau,s)x\right\|,$$

where

$$M \overset{\text{df}}{=} \sup\{\tau : \lambda \vee \|V(\tau)\| \ , \ 0 \le \tau \le \Theta\} \ .$$

The proposition will follow if we show that

$$\lim_{\Delta\to 0}\ \sup_{s\le\tau\le t-\Delta}\left\|\frac{1}{\Delta}\,[T(\Delta)R(\tau+\Delta,\tau) - S(\tau+\Delta,\tau)]S(\tau,s)x\right\| = 0 \qquad (1°)$$

for $x \in \mathcal{D}(A)$.

Indeed, the set $\mathcal{D}(A)$ is dense in $\mathbf{E}$, and the operators $S_n(t,s)$ are (for fixed t,s) equibounded in $n \in \mathbf{N}$:

$$\|S_n(t,s)\| \le [e^{M\Delta_n}]^{2n} = e^{2M(t-s)} \ .$$

To prove (1°) note first that

$$\sup_{s\le\tau\le t-\Delta}\left\|\frac{T(\Delta)R(\tau+\Delta,\tau)x-x}{\Delta}-Ax-V(\tau)x\right\|\le$$
$$\sup_{s\le\tau\le t-\Delta}\left\|T(\Delta)\frac{R(\tau+\Delta,\tau)x-x}{\Delta}-T(\Delta)V(\tau)x\right\|+$$
$$\left\|\frac{T(\Delta)x-x}{\Delta}-Ax\right\|+\sup_{s\le\tau\le t-\Delta}\|T(\Delta)V(\tau)x-V(\tau)x\|\le$$
$$\left\|\frac{T(\Delta)x-x}{\Delta}-Ax\right\|+e^{\lambda\Delta}\sup_{s\le\tau\le t-\Delta}\left\|\frac{R(\tau+\Delta,\tau)x-x}{\Delta}-V(\tau)x\right\|+$$
$$\sup_{s\le\tau\le t-\Delta}\|T(\Delta)V(\tau)x-V(\tau)x\|\ .$$

Following this estimation we conclude that

$$\lim_{\Delta\to0}\sup_{s\le\tau\le t-\Delta}\left\|\frac{T(\Delta)R(\tau+\Delta,\tau)x-x}{\Delta}-Ax-V(\tau)x\right\|=0$$
for $x\in\mathcal{D}(A)$.

Hence, by Lemma 1,

$$\lim_{\Delta\to0}\sup_{s\le\tau\le t-\Delta}\left\|\frac{1}{\Delta}\left[T(\Delta)R(\tau+\Delta,\tau)-S(\tau+\Delta,\tau)\right]x\right\|=0 \qquad (2^\circ)$$
for $x\in\mathcal{D}(A)$.

Let Y denote a linear space $\mathcal{D}(A)$ with the norm

$$\|x\|_*=\|x\|+\sup_{s\le\tau\le t}\|Ax+V(\tau)x\|\ .$$

Y is a Banach space. (We have $\|x\|\le\|x\|_*$ and $\|Ax+V(\tau)x\|\le\|x\|_*$ for each τ, $s\le\tau\le t$. So, if x_n, $n\ge1$ is a fundamental sequence in Y then there exist $x_0=\lim_{n\to\infty}x_n$ and $y=\lim_{n\to\infty}Ax_n$. Because operator A is closed x_0 belongs to $\mathcal{D}(A)$ and $y=Ax_0$. Finally $\|x_n-x_0\|_*\le\|Ax_n-Ax_0\|+(M+1)\|x_n-x_0\|$, so x_n tends to x_0 in Y.)

Let $O(\tau,\Delta)$, $s\le\tau\le t$, $0\le\Delta\le1$, $\tau+\Delta\le t$ denote a family of bounded operators acting in Y with values in $\mathbf{E}$

$$O(\tau,\Delta)=\frac{1}{\Delta}\left[T(\Delta)R(\tau+\Delta,\tau)-S(\tau+\Delta,\tau)\right]\ .$$

Formula (2°) and the Banach – Steinhaus theorem proves that they are equibounded.

For a fixed $x\in\mathcal{D}(A)$ the function $\tau\to[A+V(\tau)]S(\tau,s)x$, $\tau\in\langle s,t\rangle$ is continuous, as a derivative of an evolution operator S. Thus the function $\tau\to AS(\tau,s)x$, $\tau\in\langle s,t\rangle$ is continuous, too. It implies that the function $\tau\to S(\tau,s)x$, $\tau\in\langle s,t\rangle$ is continuous in the topology of Y. Indeed :

$$\sup_{s\le r\le t}\|\,[A+V(r)]\,[S(\tau+\Delta,s)x-S(\tau,s)x]\,\|+\|S(\tau+\Delta,s)x-S(\tau,s)x\|\le$$
$$\|A[S(\tau+\Delta,s)x-S(\tau,s)x]\,\|+(M+1)\|S(\tau+\Delta,s)x-S(\tau,s)x\|\ .$$

Hence the set $K \stackrel{\mathrm{df}}{=} \{y\ ;\ \exists\, \tau \in \langle s,t\rangle\ ,\ y = S(\tau,s)x\}$ is compact in Y. Because of the equiboundendness of the operators $O(\tau,\Delta)$ and the condition (2°)

$$\lim_{\Delta\to 0}\ \sup_{s\le\tau\le t-\Delta}\ \sup_{y\in K}\|O(\tau,\Delta)y\| = 0\ ,$$

$$\lim_{\Delta\to 0}\ \sup_{s\le\tau\le t-\Delta}\ \sup_{s\le h\le t}\|O(\tau,\Delta)S(h,s)x\| = 0 \quad,\quad x\in\mathcal{D}(A)\ ,$$

$$\lim_{\Delta\to 0}\ \sup_{s\le\tau\le t-\Delta}\|O(\tau,\Delta)S(\tau,s)x\| = 0 \quad,\quad x\in\mathcal{D}(A)\ .$$

The formula (1°) as well as Proposition is proved.

References

[1] Blanchard, Ph., Ph. Combe , M. Sirugue, M. Sirugue–Collin, *Path integral representation of the Dirac equation in presence of an external electromagnetic field*, Path integrals from meV to MeV (Bielefeld, 1985), 396–413, Bielefeld Encount. Phys. Math., VII, World Sci. Publishing, Singapore, 1986.

[2] Blanchard, Ph., Ph. Combe , M. Sirugue, M. Sirugue–Collin, *Jump processes related to the two-dimensional Dirac equation*, Stochastic processes–mathematics and physics, II, (Bielefeld, 1985), 1–13, Lecture Notes in Math. **1250** Berlin–New York, 1987.

[3] Blanchard, Ph., Ph. Combe , M. Sirugue, M. Sirugue–Collin, *Stochastic jump processes associated with Dirac equation*, Stochastic processes in classical and quantum system, (Ascona, 1985), 87–104, Lecture Notes in Phys. **262** Springer, Berlin–New York, 1986.

[4] Blanchard, Ph., Combe, Ph., Sirugue, M., Sirugue–Collin, M., *Jump processes. An introduction and some applications to quantum physics*, in /Functional integration with emphasis on the Feynman integral/ Rend. Circ. Math. Palermo (2) Suppl. No 17 (1987), pp. 47–104 (1988).

[5] Ethier, S. N. and T. G. Kurtz, "Markov processes. Characterisation and convergence", John Wiley & Sons., Wiley Series In Probability and Mathematical Statistics, New York, 1986.

[6] Goldstein, J.A., "Semigroups of Linear Operators and Applications", Oxford Mathematical Monographs, 1985.

[7] Griego, R.J. and R. Hersh, *Random evolutions, Markov chains and systems of partial differential equations*, Proc. Nat. Acad. Sci. USA **62** (1969), 305–308.

[8] Griego, R.J. and R. Hersh, *Theory of random equations with applications to partial differential equations*, Trans. Amer. Math. Soc. **156** (1971), 405–418.

[9] Heyer, M., "Probability measures on locally compact groups", Springer–Verlag, Berlin Heidelberg New York 1977.

[10] Janssen, A. and E. Siebert, *Convolution semigroups and generalized telegraph equations*, Math. Z. **177** (1981), 519–532.

[11] Kac, M. "Some Stochastic Problems in Physics and Mathematics", Magnolia Petrolum Company (Polish translation).

[12] Kisyński, J., *On M. Kac's probabilistic formula for the solution of the telegraphist's equation*, Ann. Pol. Math. **29** (1974), 259–272.

[13] Krein, S.G., "Linear Differential Equations in Banach Space", Amer. Math. Soc. Trans. of Math. Monographs, vol 29, Providence, 1971.

[14] Simon, B. and M. Reed, "Methods of modern mathematical physics" parts 1-2, Academic Press, New York–London ,1972, 1975.

[15] Zastawniak, T. *Path integrals for the telegraphers and Dirac equations: the analytic family of measures and the underlying Poisson process*, to be published in Bull. Polish Acad. Sci..

Balladyny 12/57
20–601 LUBLIN
Poland

Technical University, Lublin

Received January 9, 1991
and in final form April 29, 1991

Semigroup Forum Vol. 45 (1992) 92–119

RESEARCH ARTICLE

BOUNDARY CONDITIONS FOR ONE-DIMENSIONAL POSITIVE SEMIGROUPS

Marlene G. Ulmet

Communicated by Jerome A. Goldstein

1. INTRODUCTION

Of concern are partial differential equations of the type

$$\begin{aligned} \frac{\partial u(t,x)}{\partial t} &= Au(t,x) \qquad x \in \Omega,\ t > 0 \\ u(0,x) &= f(x) \qquad x \in \bar{\Omega}, \end{aligned} \tag{IP}$$

where Ω denotes a bounded domain in $\mathbb{R}^n$, A a linear partial differential operator and f a continuous function on $\bar{\Omega}$. Common examples are electrolysis and catalytic reactions with A a first order differential operator, and diffusion processes with A a differential operator of second order.

The basic question is, whether we need additional conditions on the boundary $\partial\Omega$ of the domain Ω, such that the initial value problem (IP) is well posed, i.e., it has a unique solution u depending continuously on the initial data f. We call conditions leading to well posed problems *admissible boundary conditions.* In this paper we give a complete answer to this question for the case $\Omega = [0,1]$. More then that, we succeed in classifying and finding a representation for all admissible boundary conditions that lead to positive solutions for positive initial data. It will turn out that some of them are of nonlocal character. W. Feller called such conditions "lateral conditions" rather than boundary conditions, see [Fe1].

For this purpose we choose a functional analytical setting, i.e., the theory of one-parameter semigroups, see [Go], [Na]. In this new setting, our basic question turns into the problem of determing the domains $D(A) \subset C[0,1]$ of a given differential operator A, such that A with domain $D(A)$ generates a strongly continuous semigroup, especially a positive one.

We only consider first and second order differential operators because S. Miyajima and N. Okazawa proved in [MO] that if a differential operator generates a *positive* strongly continuous semigroup on the space $L^p(\mathbb{R}^n)$, then it must be (degenerate) elliptic and of order at most 2. Recently, W. Arendt, together with C.J.K. Batty and D.W. Robinson proved the above result for the Banach space of all continuous functions defined on a domain of $\mathbb{R}^n$, see [ABR].

The problem of determing all domains $D(A) \subset C[0,1]$ such that the operator A defined on $C[0,1]$ with domain $D(A)$ generates a positive semigroup was solved for the operator $Af = f'$ in [MS], and for the operator $Af = f''$ in [St]. In this paper we consider first and second order differential operators perturbed in a multiplicative and additive way as well.

If the differential operator A defined on $C[0,1]$ is of order one, then we also deal with singularities in the sense that we allow the coefficient of the derivative to vanish. In the second part of this paper we consider the second order differential operator $Af = mf'' + nf' + pf$ with $m > 0$ on $(0,1)$ and $m, n, p \in C[0,1]$. Many authors have considered boundary value problems for one-dimensional diffusion processes. Similar results to those of Section 3.3 were proved for the case of *contraction* semigroups; for details see the last section.

The following theorem which is due to Arendt, Chernoff and Kato, represents a new and efficient tool for characterizing positive semigroups. Some of our results based on this theorem are therefore easy to apply in the case of specific examples.

Theorem 1.1. *A densely defined linear operator A on $C(K)$, K compact, is the generator of a positive semigroup if and only if*

i) $(\lambda - A)$ is surjective for sufficiently large λ,

ii) A satisfies the "positive minimum principle" (P), that is

$$0 \le f \in D(A),\ x \in K,\ f(x) = 0 \quad \textit{implies} \quad (Af)(x) \ge 0. \qquad (P)$$

2. THE FIRST ORDER DIFFERENTIAL OPERATOR

Given $m \in C[0,1]$, we define a maximal operator A_m on $C[0,1]$ by

$$A_m f = mf', \quad m \in C[0,1], \quad \text{with domain}$$

$$D(A_m) = \{f \in C[0,1] : f \text{ is continuously differentiable in } x \text{ if } m(x) \ne 0 \text{ and } mf' \text{ has a continuous extension for } x \text{ if } m(x) = 0\}. \tag{2.1}$$

The purpose of this section to determine those restrictions of A_m which are generators of positive semigroups on $C[0,1]$. In the sequel we use the notation $C^1_m[0,1]$ for the maximal domain $D(A_m)$ defined above.Therefore shall perform the following steps.

a) We solve the problem considering the maximal operator $B_{\max} f = f'$ defined on $C^1[a,b]$, with $-\infty \le a < b \le \infty$.

b) Next we allow a nonvanishing multiplicative perturbation, i.e., we consider the maximal operator $A_m f = f'$ on $C[0,1]$ with $m \ne 0$ on $(0,1)$.

c) Finally we consider the general case, where A_m is defined by (2.1).

The operator $Bf = mf' + nf$ defined on X, with $m, n \in C[0,1]$ and domain $D(B) = D(A)$ represents only a bounded additive perturbation of A and generates a positive semigroup if and only if A does.

2.1 The first derivative on a compact interval. The starting point for this section yield characterization theorems of generators of positive translation semigroups on spaces of continuous functions on compact intervals. We consider the maximal operator $B_{\max}$ defined on $C[a,b]$, $-\infty \le a < b \le \infty$, with $D(B_{\max}) = C^1[a,b]$ and $B_{\max} f = f'$. If a or b are infinite, then $C[a,b] = \{f \in C(a,b) : \lim_{x\to a} f(x) \text{ and } \lim_{x \to b} f(x) \text{ exist}\}$, and $C^1[a,b] = \{f \in C[a,b] \cap C^1(a,b) : \lim_{x\to a} f'(x) \text{ and } \lim_{x\to b} f'(x) \text{ exist}\}$. For the Banach space $X = C[a,\infty]$, with $-\infty \le a$ the following proposition holds (see [Yo], p.242).

Proposition 2.1. *Consider the maximal operator $B_{\max} f = f'$ with domain $D(B_{\max}) = C^1[a,\infty]$, $-\infty \le a$. Then $B_{\max}$ generates the following positive translation semigroup on $X = C[a,\infty]$*

$$(T(t)f)(x) = f(x+t), \qquad \textit{for all} \quad x \in [a,\infty],\ t \ge 0. \qquad \blacksquare$$

Of concern is now the Banach space $X = C[a,b]$, with $b < \infty$ and $-\infty \le a$. We start with a finite. For this case we have the following characterization theorem, see [MS].

Definition 2.2. A *positive boundary condition at* $x = b$ is a linear form Φ_b on $C^1[a,b]$ given by

$$\Phi_b(f) = c'f'(b) + \int_a^b \frac{f(b)-f(s)}{b-s}d\mu(s) + cf(b), \tag{2.2}$$

with $c, c' \in \mathbb{R}$, $c' \geq 0$, μ a positive bounded Borel measure on $[a,b]$ such that $\mu(b) = 0$, and

$$c' = 0 \quad \text{implies} \quad \int_a^b \frac{1}{b-s}d\mu(s) = \infty. \tag{2.3}$$

Theorem 2.3. *A restriction B of $B_{\max}$ is the generator of a positive semigroup on $C[a,b]$, a,b finite, if and only if there is a positive boundary condition Φ_b at b such that $D(B) = \ker \Phi_b$.* ∎

Next we generalize the above theorem to the more involved case where a is infinite. Let $X = C[-\infty, b]$ and recall that the operator $B_{\max}$ defined on X is given by $B_{\max}f = f'$ with domain $D(B_{\max}) = C^1[-\infty, b]$, where $C^1[-\infty,b] = \{f \in C[-\infty,b] :\ f$ is continuously differentiable for $-\infty < x \leq b$ and $\lim_{x\to-\infty} f'(x)$ exists$\}$. The following definition gives a representation for all positive boundary conditions at $x = b$ for the case when the Banach space $X = C[-\infty, b]$.

Definition 2.4. Let $z \in (-\infty, b)$ be fixed. A *positive boundary condition at* $x = b$ is a linear form Φ_b on $C^1[-\infty,b]$ given by

$$\Phi_b(f) = c'f'(b) + \int_z^b \frac{f(b)-f(s)}{b-s}d\mu(s) - \int_{-\infty}^z f(s)d\nu(s) + cf(b) \tag{2.4}$$

with $c, c' \in \mathbb{R}$, $c' \geq 0$, μ and ν positive bounded Borel measures, μ defined on $[z,b]$ with $\mu(b) = 0$, ν defined on $[-\infty, z]$ and

$$c' = 0 \quad \text{implies} \quad \int_z^b \frac{1}{b-s}d\mu(s) = \infty. \tag{2.5}$$

Theorem 2.5. *A restriction B of $B_{\max}$ genereates a positive semigroup on $C[-\infty,b]$, b finite, if and only if there is a positive boundary condition Φ_b at $x = b$ (given by Definition 2.4) such that $D(B) = \ker\Phi_b$.*

Proof. First we prove that if B generates a positive semigroup then $D(B) \subset \ker\Phi_b$ and secondly we show that the operator B with domain $\ker\Phi_b$ generates a semigroup. This implies $D(B) = \ker\Phi_b$.

1. Assuming that B generates a positive semigroup $T(t)$, we obtain by the definition of a generator

$$(Bf)(b) = \lim_{t\downarrow 0} \frac{1}{t}[< f - f_b, T(t)'\delta_b > + < f_b, T(t)'\delta_b > - f(b)],$$

where f_b denotes the constant function $f_b(s) = f(b)$ for all $s \in [-\infty, b]$. Let $z \in (-\infty, b)$. Then there exist probability measures μ_t and ν_t with $\mu_t(\{b\}) = 0$ and positive functions $q(t), \bar{p}(t)$ for $t > 0$ such that

$$T(t)'\delta_b = q(t)\nu_t + \bar{p}_t\mu_t,$$

with $supp\ \nu_t \subset [-\infty, z]$ and $supp\ \mu_t \subset [z, b]$. With the above notations and $p(t) := (b-s)\bar{p}(t) \geq 0$ one obtains

$$(Bf)(b) = \lim_{t\downarrow 0} \frac{1}{t}\Big[\int_{-\infty}^{z} f(s)d\nu_t(s)q(t) + \int_z^b \frac{f(s)-f(b)}{b-s}d\mu_t(s)p(t) + \\ + [\int_z^b d\mu_t(s)\bar{p}(t) - 1]f(b)\Big].$$

We denote by $r : (0,\infty) \to \mathbb{R}$ the mapping $t \to \int_z^b d\mu_t(s)\bar{p}(t) - 1$; then r is a continuous function, and we have

$$Bf(b) = \lim_{t\downarrow 0} \frac{1}{t}\Big[\int_{-\infty}^{z} f(s)d\nu_t(s)q(t) - \int_z^b \frac{f(b)-f(s)}{b-s}d\mu_t(s)p(t) + r(t)f(b)\Big]. \quad (2.6)$$

Since X is separable, the bounded sets $\{\mu_t : t \in (0,1]\}$, $\{\nu_t : t \in (0,1]\}$ are $\sigma(X', X)$ relatively sequentially compact. For this reason one can choose a sequence $(t_n) \subset (0,\infty)$ such that $t_n \to 0$, $\nu_{t_n} \to \nu$, $\mu_{t_n} \to \mu$, $\frac{q(t_n)}{t_n} \to Q \geq 0$, $\frac{p(t_n)}{t_n} \to P \geq 0$, $\frac{r(t_n)}{t_n} \to R$, where ν, μ are positive measures and Q, P, R (possibly infinite) numbers.

Case i) If Q, P and R are finite, we obtain that $f \in D(B)$ satisfies $f \in \ker \Phi_b$, where Φ_b is given by

$$\Phi_b(f) = f'(b) - Q\int_{-\infty}^{z} f(s)d\nu(s) + P\int_z^b \frac{f(b)-f(s)}{b-s}d\mu(s) + Rf(b),$$

with $Q, P \geq 0$ and $R \in \mathbb{R}$.

Case ii) If one of these limits is not finite, we divide relation (2.6) with $g(t_n)/t_n$, where $g(t_n) := q(t_n) + p(t_n) + |r(t_n)|$ in order to obtain (2.4).

Assuming $c' = 0$ and $\int_z^b \frac{1}{b-s}d\mu(s) < \infty$, the linear form Φ_b would be continuous with respect to the supremum norm, which contradicts that $\ker \Phi_b$ is dense in $C[-\infty, b]$. This proves (2.5).

2. Conversely, we have to verify the conditions of the Arendt-Chernoff-Kato characterization theorem.

a) The operator B is densely defined. A simple regularization procedure shows that $D(B_{\max})$ is dense in X. Consider the normed space $Y = D(B_{\max})$ equipped with hte supremum norm. Then $\Phi_b : Y \to \mathbb{R}$ given by (2.4) is a discontinuous linear form on Y due to (2.5). Consequently $D(B) = \ker \Phi_b$ is dense in Y. Since Y is dense in X, $D(B)$ is dense in X as well.

b) B satisfies (P). This is clear for $x \in (-\infty, b)$. Let $x = -\infty$, $0 \leq f \in D(B)$ with $f(-\infty) = 0$. Then, by the mean value theorem there exists $\xi_x \in [x, x+1]$ such that $f(x+1) - f(x) = f'(\xi_x)$. We obtain $(Bf)(-\infty) = f'(-\infty) = \lim_{x\to-\infty} f'(\xi_x) = \lim_{x\to-\infty}(f(x+1) - f(x)) = 0$. Take now $x = b$ and $0 \leq f \in D(B)$ with $f(b) = 0$.

i) If $c' > 0$ in (2.4), then $f \in D(B)$ implies

$$0 = \Phi_b(f) = c'f'(b) + \int_z^b \frac{-f(s)}{b-s}d\mu(s) - \int_{-\infty}^{z} f(s)d\nu(s),$$

and since $f \geq 0$ we have

$$(Bf)(b) = f'(b) = \frac{1}{c'} \int_z^b \frac{f(s)}{b-s} d\mu(s) + \frac{1}{c'} \int_{-\infty}^z f(s) d\nu(s) \geq 0.$$

ii) If $c' = 0$ in (2.4), then we have for $f \in D(B)$

$$\Phi_b(f) = \int_z^b \frac{-f(s)}{b-s} d\mu(s) - \int_{-\infty}^z f(s) d\nu(s) = 0.$$

Both terms have the same sign so that each one must be zero. But $\int_z^b \frac{f(s)}{b-s} d\mu(s) = 0$ implies together with (2.5) that there is a sequence (s_n), $(s_n) \subset [a,b] \cap supp\ \mu$, with $s_n \to 0$. Since f is positive and $\Phi_b(f) = 0$, we conclude $f(s_n) = 0$, and therefore

$$(Bf)(b) = f'(b) = \lim_{n \to \infty} \frac{f(b) - f(s_n)}{b-s} = 0.$$

c) $(\lambda - B)$ is surjective for λ big enough. For every $g \in C[-\infty, b]$ we consider the equation $\lambda f - f' = g$. The solutions are given by

$$f(s) = \int_s^b e^{\lambda(s-t)} g(t) dt + k e^{\lambda s}, \qquad k \in \mathbb{R},$$

and simple computations show that $f \in D(B_{\max})$. Now $f \in \ker \Phi_b = D(B)$ if and only if

$$k \Phi_b \left(e^{\lambda s}\right) + \Phi_b \left(\int_s^b e^{\lambda(s-t)} g(t) dt \right) = 0.$$

If $\Phi_b(e^{\lambda s}) \neq 0$, then we determine the constant k as the solution of the above equation, and $(\lambda - B)$ is surjective. In fact, we have

$$e^{-\lambda b} \Phi_b(e^{\lambda s}) = c'\lambda + c + \underbrace{\int_z^b \frac{1 - e^{\lambda(s-b)}}{b-s} d\mu(s)}_{=:D>0} - \int_{-\infty}^z e^{\lambda(s-b)} d\nu(s).$$

The last integral is bounded for all $\lambda > 0$ since $s - b < 0$ and ν is a bounded Borel measure. Thus we can estimate $\Phi_b(e^{\lambda s})$ by

$$e^{-\lambda b} \Phi_b(e^{\lambda s}) > (c'\lambda + c - M) + D.$$

If $c' > 0$ then there is a λ_0 such that for $\lambda > \lambda_0$ we have $e^{-\lambda b} \Phi_b(e^{\lambda s}) > 0$; thus $\Phi_b(e^{\lambda s}) > 0$. If $c' = 0$ then condition (2.5) implies $D \to \infty$ for $\lambda \to \infty$. ■

Let us point out the following properties of the Banach spaces $C[-\infty, b]$ or $C[a, \infty]$ that we met in the previous proof.

Remark 2.6. i) $f \in C^1[-\infty, b]$ implies $f'(-\infty) = 0$ and $f \in C^1[a, \infty]$ implies $f'(\infty) = 0$ respectivly.

ii) If $c' = 0$ in (1.4), then $f \geq 0$, $f(x) = 0$ implies $f'(x) = 0$ for every $x \in [-\infty, b]$.

2.2 The operator Af=mf' for non-vanishing m. Since for the case $m(x) = 1$ for all $x \in [0,1]$ we already have the desired characterization, a first attempt is to apply perturbation theory, but J.R. Dorroh's result (see [Do]) on the multiplicative perturbation m of a generator A, in the case when A is dissipative, assumes m strictly positive. (See also [GL]). We start by looking at the case $m(x) \neq 0$ for $x \in (0,1)$, i.e., m may vanish for $x = 0$ and $x = 1$ respectively, and try to find a characterization for generators of positive semigroups. The underlying idea is to reduce our problem, via similarity transformation, to the situation $m(x) = 1$.

Let $m \in C[0,1]$, $m(x) \neq 0$ for $x \in (0,1)$. We fix $z \in (0,x)$, and define $\varphi(x) := \int_z^x \frac{dy}{m(y)}$. The mapping φ defines a homeomorphism from $[0,1]$ to $[a,b]$ and a diffeomorphism from $(0,1)$ to (a,b). We define the transformation $V : C[a,b] \to C[0.1]$, $Vf := f \circ \varphi$, which represents a lattice isomorphism from $C[a,b]$ to $C[0,1]$.

Definition 2.7. Let A be an operator on X, $V : X \to Y$ a lattice isomorphism. Then $V^{-1}AV$ is defined on Y by $D(V^{-1}AV) = \{f \in Y : Vf \in D(A)\}$ and $(V^{-1}AV)f = V^{-1}A(Vf)$.

Remark 2.8. A has the generator property if and only if $V^{-1}AV$ does.

Consider the maximal operator A_m defined on $X = C[0,1]$ by $A_m f = mf'$, $m \in C[0,1]$, $m \neq 0$ on $(0,1)$ with domain $D(A_m) = C^1_m[0,1]$. Proposition 2.9 shows that if we apply a similarity transformation with $Vf = f \circ \varphi$, $\varphi(x) := \displaystyle\int_z^x 1/m$, $z \in (0,x)$ to the maximal operator A_m, then we obtain an "unperturbed" maximal operator B.

Proposition 2.9. *Let the operator B defined on $C[a,b]$ with $B = V^{-1}A_m V$. Then $Bf = f'$ and $D(B) = C^1[a,b]$.*

Notation. *In the sequel $\lim\limits_{x \to a,b} f(x)$ exist means that $\lim\limits_{x \to a} f(x)$ and $\lim\limits_{x \to b} f(x)$ exist.*

Proof. The domain $D(B)$ is determined by

$$\begin{aligned} D(B) &= \{f \in C[a,b] : \ f \circ \varphi \in C^1_m[0,1]\} \\ &= \{f \in C[a,b] : \ f \circ \varphi \in C^1(0,1) \cap C[0,1], \ \lim_{x \to 0,1} [A(f \circ \varphi)](x) \text{ exists}\}. \end{aligned}$$

Since φ is a homeomorphism from $[0,1] \to [a,b]$ and a diffeomorphism from $(0,1) \to (a,b)$, we obtain

$$\begin{aligned} D(B) &= \{f \in C^1(a,b) \cap C[a,b] : \ \lim_{x \to 0,1} [m(f' \circ \varphi)\varphi'](x) \text{ exists }\} \\ &= \{f \in C^1(a,b) \cap C[a,b] : \ \lim_{x \to 0,1} (f' \circ \varphi)(x) \text{ exists }\} \\ &= \{f \in C^1(a,b) \cap C[a,b] : \ \lim_{x \to 0,1} f'(x) \text{ exists }\} = C^1[a,b]. \end{aligned}$$

We have

$$\begin{aligned} Bf &= (V^{-1}AV)f = V^{-1}A(f \circ \varphi) = V^{-1}m(f \circ \varphi)^{-1} \\ &= V^{-1}m(f' \circ \varphi)\varphi' = V^{-1}(f' \circ \varphi) = f'. \end{aligned}$$

∎

Thus the operator $A_m f = mf'$ with domain $D(A_m) = C^1_m[0,1]$, $m \in C[0,1]$,
$m(x) > 0$ for $x \in (0,1)$ generates a positive strongly continuous semigroup on $C[0,1]$ if and only if the operator $Bf = f'$ on $C[a,b]$ with domain $D(B) = C^1[a,b]$ does.

If for example, $m(1) = 0$ and $\varphi(1) = b = \infty$, then we reduce our problem via similarity transformation to the question whether the operator $Bf = f'$ defined on $C[a,\infty]$ with domain $C^1[a,\infty]$ has the generator property or not. Section 2.1 gives the necessary tools to formulate a first characterization for the case $m(x) \neq 0$ for $x \in (0,1)$. Consider $z \in (0,x)$ be fixed for $x \in (0,1)$. Using Remark 2.8. and Proposition 2.1. we obtain the following characterization.

Proposition 2.10.
a) If $m(x) > 0$ for $x \in (0,1)$, then A_m generates a positive semigroup if and only if $\int_z^1 \frac{1}{m(y)} dy = \infty$.
b) If $m(x) < 0$ for $x \in (0,1)$, then A_m generates a positive semigroup if and only if $\int_0^z \frac{1}{m(y)} dy = -\infty$.

Remark 2.11. If $m(x) > 0$ ($m(x) < 0$) for $x \in (0,1)$, then it does not matter whether $m(0) = 0$ ($m(1) = 0$) or not.

The following results describe the situation when we do need boundary conditions. Again, we denote by $\varphi(x) := \int_\eta^x \frac{dy}{m(y)}$, $\eta \in (0,x)$.

Definition 2.12. Consider $z \in (0,1)$ fixed. A *positive boundary condition at* $x = 1$ is a linear form Φ_1 on $C^1_m[0,1]$ given by

$$\Phi_1(f) = d'(Af)(1) + \int_z^1 \frac{f(1) - f(s)}{\varphi(1) - \varphi(s)} d\mu(s) - \int_0^z f(s) d\nu(s) + df(1), \tag{2.7}$$

with $d, d' \in \mathbb{R}$, $d' \geq 0$, μ and ν positive bounded Borel measures, μ defined on $[z,1]$ with $\mu(1) = 0$, ν defined on $[0,z]$ and

$$d' = 0 \quad \text{implies} \quad \int_z^1 \frac{1}{\varphi(1) - \varphi(s)} d\mu(s) = \infty. \tag{2.8}$$

Theorem 2.13. *Let $m(x) > 0$ if $x \in (0,1)$. Then a restriction A of A_m generates a positive semigroup on $C[0,1]$ if and only if there is a positive boundary condition Φ_1 at $x = 1$ such that $D(A) = \ker \Phi_1$.*

Proof. First we note that $s \to \frac{f(1) - f(s)}{\varphi(1) - \varphi(s)} \in C[0,1]$ for $f \in D(A_m)$ and $s \in [z,1]$.

1. Assume that A with domain $D(A)$ generates a positive semigroup $(T(t))_{t \geq 0}$ on $X = C[0,1]$. By definition, one has for $f \in D(A)$

$$(Af)(1) = \lim_{t \to 0} \frac{1}{t} \Big[< f - f_1, T(t)'\delta_1 > + < f_1, T(t)'\delta_1 > - f(1) \Big],$$

with f_1 the constant function $f_1(s) = f(1)$, for all $s \in [0,1]$. Let $z \in (0,1)$. Then there exist the probability measures μ_t and ν_t with $\mu_t(\{1\}) = 0$, *supp* $\mu_t \subset [z,1]$, *supp* $\nu_t \subset [0,z]$ and the positive functions $q(t), \bar{p}(t)$ such that

$$T(t)'\delta_1 = q(t)\nu_t + \bar{p}(t)\mu_t, \quad t > 0.$$

By hypothesis $\varphi(s) < \infty$ for $s \in [z,1]$, and $\varphi(1) - \varphi(s) \neq 0$ because $m > 0$ on $(0,1)$. We denote by $p(t) := (\varphi(1) - \varphi(s))\bar{p}(t)$ and together with the above notations we obtain

$$\begin{aligned}(Af)(1) &= \lim_{t\to 0} \frac{1}{t}\Big[\int_0^z [f(s) - f(1)]d\nu_t(s)q(t) + \int_z^1 \frac{f(s) - f(b)}{\varphi(b) - \varphi(s)}d\mu_t(s)p(t) + \\ &\quad + \int_0^z f(1)d\nu_t(s)q(t) + \int_z^1 f(1)d\mu_t(s)\bar{p}(t) - f(1)\Big] \\ &= \lim_{t\to 0} \frac{1}{t}\Big[\int_0^z f(s)d\nu_t(s)q(t) + \int_z^1 \frac{f(s) - f(b)}{\varphi(b) - \varphi(s)}d\mu_t(s)p(t) + \\ &\quad + [\int_z^1 d\mu_t(s)\bar{p}(t) - 1]f(1)\Big].\end{aligned}$$

The further considerations are simlar to those from the proof of Theorem 2.5. It follows that $D(A) \subset \ker\Phi_1$.

2. We show that the operator A with domain $\ker\Phi_1$ generates a positive semigroup. This implies $D(A) = \ker\Phi_1$. We consider the operator $B := V^{-1}AV$ (see Definition 2.7), with $Vf = f \circ \varphi$ from the beginning of this section. Then the operator $Bf = f'$ and is defined on $C[a,b]$ with $-\infty \leq a < b < \infty$ and we determine $D(B)$ as follows.

$$\begin{aligned}D(B) &= \{f \in C[a,b] : \; Vf \in D(A)\} \\ &= \{f \in C[a,b] : \; f \circ \varphi \in C^1_m[0,1], \; \Phi_1(f \circ \varphi) = 0\} \\ &= \{f \in C^1[a,b] : \; \Phi_1(f \circ \varphi) = 0\}.\end{aligned}$$

We compute $\Phi_1(f \circ \varphi)$. If we assume that there exist $d', d \in \mathbb{R}$, $d' \geq 0$, bounded positive Borel measures μ on $[z,1]$, ν on $[0,z]$, with $\mu(\{1\}) = 0$ and such that relation (2.7) holds, then $\Phi_1(f \circ \varphi) = 0$ is

$$\begin{aligned}d'm(1)f'(\varphi(1))\varphi'(1) + \int_z^1 \frac{f(\varphi(1)) - f(\varphi(s))}{\varphi(1) - \varphi(s)}d\mu(s) - \\ - \int_0^z f(\varphi(s))d\nu(s) + df(\varphi(1)) = 0.\end{aligned}$$

We substitute $\varphi(s) = t$, $\varphi(0) = a, \varphi(1) = b$ and obtain

$$d'f(b) + \int_y^b \frac{f(b) - f(t)}{b - t}d\tilde{\mu}(t) - \int_a^y f(b)d\tilde{\nu}(t) + df(b) = 0$$

with $d, d' \in \mathbb{R}$, $d' \geq 0$ and bounded positive Borel measures $\tilde{\mu}(t) = \mu(\psi(t))$ and $\tilde{\nu}(t) = \nu(\psi(t))$ defined on $[a,b]$. ($\tilde{\mu}, \tilde{\nu}$ are the image measures of μ, ν through φ with φ continuous.) Also $\tilde{\mu}(b) = \mu(\psi(b)) = \mu(\{1\}) = 0$ and relation (2.8) reads

$$d' = 0 \quad \text{implies} \quad \int_y^b \frac{1}{b - t}d\tilde{\mu}(t) = \infty.$$

We are in the framework of Theorem 2.5. (and Theorem 2.3 respectively) and obtain that B generates a positive semigroup. Then also A does. ■

Theorem 2.13 has an analogue for $m(x) < 0$ if $x \in (0,1)$. This can be proved with exactly the same methods. Note that using the similarity transformation described before we obtain an operator $Bf = f'$ defined on $C[b,a]$ with $a = \varphi(0) < \infty$ and $b = \varphi(1)$ possibly infinite.

Definition 2.14. Consider $z \in (0,1)$ fixed. A *positive boundary condition at* $x = 0$ is a linear form Φ_0 on $C^1_m[0,1]$ given by

$$\Phi_0(f) = d'(Af)(0) + \int_0^z \frac{f(0)-f(s)}{\varphi(0)-\varphi(s)} d\mu(s) - \int_z^1 f(s)d\nu(s) + df(0), \qquad (2.9)$$

with $d, d' \in \mathbb{R}$, $d' \geq 0$, μ and ν positive bounded Borel measures, μ defined on $[0,z]$ with $\mu(0) = 0$, ν defined on $[z,1]$ and

$$d' = 0 \quad \text{implies} \quad \int_0^z \frac{1}{\varphi(0)-\varphi(s)} d\mu(s) = \infty. \qquad (2.10)$$

Theorem 2.15. *Let $m(x) < 0$ for $x \in (0,1)$. Then a restriction A of A_m generates a positive semigroup on $C[0,1]$ if and only if there is a positive boundary condition Φ_0 at 0 such that $D(A) = \ker \Phi_0$.* ∎

2.3 The main result. In applications, the perturbation m is in general a polynomial function that may vanish on the underlying spatial interval. Therefore we try to find a characterization, similar to that from the previous section for the differential operator

$Af = mf'$, $m \in C[0,1]$ be defined on $C[0,1]$ with maximal domain $D(A_m) = \{f \in C[0,1] : f$ is differentiable in x if $m(x) \neq 0$ and Af has a continuous extension on $C[0,1]\} =: C^1_m[0,1]$.

A first attempt is to specialize our problem as follows. If m vanishes at a finite number of points $a_1, ..., a_k$, $a_1 \neq 0$, $a_k \neq 1$, the question when A generates a positive semigroup on $C[0,a_1]$, $C[a_i, a_{i+1}]$ for $i = 1, ..., k-1$ and on $C[a_k, 1]$ is answered in Section 2.1. All we have to do is to join the respective semigroups into one semigroup on $C[0,1]$ together. The gap of this procedure is, that we do not obtain *all* positive semigroups generated by A. For example on $[a_k, 1]$ we have to impose a boundary condition at $x = 1$ if $m(1) > 0$. The boundary condition would be given only by linear forms defined on $C^1_m[a_k, 1]$, but still there are other linear forms defined on $C^1_m[0,1]$ which lead to well posed problems.

For the characterization theorem in the "general case" we adapt the proof sketch of Theorem 2.5. But first let us specify what we mean by the "general case".

Definition 2.16. A continuous function $m : [0,1] \to \mathbb{R}$ is called *admissible* on $(0,1)$ if for every $a_i \in (0,1)$ such that $m(a_i) = 0$ we have

$$\int_{a_i-\varepsilon}^{a_i} \frac{1}{|m(y)|} dy = \int_{a_i}^{a_i+\varepsilon} \frac{1}{|m(y)|} dy = \infty,$$

with $\varepsilon > 0$ such that m has no zero on $(a_i - \varepsilon, a_i)$ and $(a_i, a_i + \varepsilon)$.

Throughout this section the assumptions on the multiplicative perturbation m are the following. Let $m : [0,1] \to \mathbb{R}$ be continuous, admissible in the sense of Definition 2.16, and has at most a finite number of zeroes $a_i, i = 1, ..., k$ on $(0,1)$ with $0 < a_1 < ... < a_k < 1$. The assumptions on m are quite natural since every Lipschitz continuous function $m : [a,b] \to \mathbb{R}$, $-\infty < a < b < \infty$ is admissible. As example for the multiplicative perturbation m we can take any polynomial function.

We consider the maximal operator A_m on $C[0,1]$ defined by $D(A_m) = C^1_m[0,1]$ and $A_m f = mf'$. For $0 < z_0 < a_1$ and $a_k < z_1 < 1$ we define φ_0 on $(0, z_0)$ and φ_1 on $(z_1, 1)$ by

$$\varphi_0(x) = \int_{z_0}^x \frac{1}{m(y)} dy, \qquad \varphi_1(x) = \int_{z_1}^x \frac{1}{m(y)} dy.$$

Remark 2.17. If $m : [0,1] \to \mathbb{R}$ is continuous and admissible and $a_i \in (0,1)$ denotes a zero of m, then $(Af)(a_i) = 0$.

Proof. Assume $m > 0$ on (a_{i-1}, a_i) and use the similarity transformation described at the beginning of Section 1.2. Then m admissible implies that the similar operator B is defined on $[\varphi(a_{i-1}), \infty]$. By Remark 1.4 we have $(Bf)(\infty) = 0$ thus, transformimg backwards, $(Af)(a_i) = m(a_i)f'(a_i) = 0$. ∎

Definition 2.18. A *positive boundary condition at 0* is a linear form Φ_0 on $C^1_m[0,1]$ given by

$$\Phi_0(f) = d_0'(Af)(0) + \int_0^{z_0} \frac{f(s) - f(0)}{\varphi_0(s) - \varphi_0(0)} d\mu_0(s) - \int_{z_0}^1 f(s) d\nu_0(s) + d_0 f(0), \quad (2.11)$$

with $d_0, d_0' \in \mathbb{R}$, $d_0' \geq 0$, μ_0, ν_0 positive bounded Borel measures on $[0, z_0]$ and $[z_0, 1]$ respectively, such that $\mu_0(\{0\}) = 0$ and

$$d_0' = 0 \quad \text{implies} \quad \int_0^{z_0} \frac{1}{\varphi_0(s) - \varphi_0(0)} d\mu_0(s) = \infty. \quad (2.12)$$

A *positive boundary condition at 1* is a linear form Φ_1 on $C^1_m[0,1]$ given by

$$\Phi_1(f) = d_1'(Af)(1) + \int_{z_1}^{1} \frac{f(1) - f(s)}{\varphi_1(1) - \varphi_1(s)} d\mu_1(s) - \int_0^{z_1} f(s) d\nu_1(s) + d_1 f(1), \quad (2.13)$$

with $d_1, d_1' \in \mathbb{R}$, $d_1' \geq 0$, μ_1, ν_1 positive bounded Borel measures on $[z_1, 1]$ and $[0, z_1]$ respectively, such that $\mu_1(\{1\}) = 0$ and

$$d_1' = 0 \quad \text{implies} \quad \int_{z_1}^{1} \frac{1}{\varphi_1(1) - \varphi_1(s)} d\mu_1(s) = \infty. \quad (2.14)$$

We show that the necessity of boundary conditions at the endpoints of the interval $[0,1]$ depends on the sign of m at these points and if m vanishes at $x = 0$ or $x = 1$, it depends whether the integrals $\int_0^{z_0} \frac{1}{|m|}$ or $\int_{z_1}^1 \frac{1}{|m|}$ are finite or not. We start with the case when m does not vanish at the endpoints.

Theorem 2.19. *Let $m(0) \neq 0$ and $m(1) \neq 0$ and consider the restriction A of A_m defined on $X = C[0,1]$.*

Case 1. If $m(0) > 0, m(1) < 0$ then A_m generates a positive semigroup on X.

Case 2. If $m(0) > 0, m(1) > 0$ then A generates a positive semigroup on X if and only if there is a positive boundary condition Φ_1 at 1 such that $D(A) = \ker \Phi_1$.

Case 3. If $m(0) < 0, m(1) < 0$ then A generates a positive semigroup on X if and only if there is a positive boundary condition Φ_0 at 0 such that $D(A) = \ker \Phi_0$.

Case 4. If $m(0) < 0, m(1) > 0$ then A generates a positive semigroup on X if and only if there are the positive boundary conditions Φ_0 at 0 and Φ_1 at 1 such that $D(A) = \ker \Phi_0 \cap \ker \Phi_1$.

Proof. Case 1. We verify the conditions of the Arendt-Chernoff-Kato characterization theorem.

a) $(\lambda - A)$ is surjective for λ big enough, i.e., for every $g \in C[0,1]$ there exists $f \in D(A_m)$ such that $\lambda f - mf' = g$. We divide our problem as follows. We search a solution of the equation on each subinterval $[0, a_1], [a_i, a_{i+1}], i = 1, ..., k-1$, and $[a_k, 1]$. Then we show that f defined on subintervals is continuous on $[0,1]$ and belongs indeed to $D(A_m)$. The solution f of the equation $\lambda f - mf' = g$ has either the representation

$$f(x) = k_1 e^{-\lambda \int_x^{a_{i+1}} \frac{dy}{m(y)}} + \int_x^{a_{i+1}} \frac{g(t)}{m(t)} e^{-\lambda \int_x^t \frac{dy}{m(y)}} dt, \quad x \in [a_i, a_{i+1}] \tag{2.15}$$

or

$$f(x) = k_2 e^{\lambda \int_{a_i}^x \frac{dy}{m(y)}} - \int_{a_i}^x \frac{g(t)}{m(t)} e^{-\lambda \int_x^t \frac{dy}{m(y)}} dt, \quad x \in [a_i, a_{i+1}]. \tag{2.16}$$

If $m > 0$ on (a_i, a_{i+1}) then we choose (2.15) and if $m < 0$ on (a_i, a_{i+1}) then (2.16). Since m is admissible on $(0,1)$ the first term of the solution vanishes. Without loss of generality we assume that $m > 0$ on $(a_i, a_{i+1}), i = 1, ..., k-1$. Then the solution f is

$$f(x) = \begin{cases} \int_x^{a_i} \frac{g(t)}{m(t)} e^{-\lambda \int_x^t \frac{dy}{m(y)}} dt & \text{if } x \in [0, a_1] \cup [a_{i-1}, a_i], i = 2, ..., k, \\ -\int_{a_k}^x \frac{g(t)}{m(t)} e^{-\lambda \int_x^t \frac{dy}{m(y)}} dt & \text{if } x \in [a_k, 1]. \end{cases}$$

Computing $(Af)(x) = m(x)f'(x)$ we obtain that $(Af)(a_i) = 0$ for $i = 1, ..., k$. (We can use also Remark 2.18.) Then the equation $\lambda f - mf' = g$ implies $f(a_i) = g(a_i)/\lambda$ for $i = 1, ..., k$. Thus $f \in C[0,1]$ and using the explicit form of the solution, we obtain $f \in D(A_m)$.

b) A verifies the positive minimum principle (P) for every $x \in [0,1]$, i.e., $0 \le f \in D(A_m), f(x) = 0$ implies $(Af)(x) \ge 0$. This is clear for $x \in (0,1)$ such that $m(x) \ne 0$.
Let $x = 0$, $0 \le f \in D(A_m)$ and $f(0) = 0$. Then $f(0) = 0$ implies $f'(0) \ge 0$, otherwise it would be a contradiction to the positivity of f. Since by assumption $m(0) > 0$, we obtain $(Af)(0) = m(0)f'(0) \ge 0$. Similarly for $x = 1$, $f(1) = 0$ implies $f'(1) \le 0$ but by assumption $m(1) < 0$ thus $(Af)(1) \ge 0$.
If $x = a_i$, $i = 1, ..., k$ is a zero of m, then we use Remark 1.13 and obtain $(Af)(a_i) = 0$.

c) A is densely defined. A simple regularization procedure shows that $D(A_m)$ is dense in $X = C[0,1]$.

Case 2. We keep the proof scheme of Theorem 2.5, i.e., first we show that if A generates a positive semigroup, then $D(A) \subset \ker \Phi_1$ and secondly that A with domain $\ker \Phi_1$ generates a positive semigroup. This implies $D(A) = \ker \Phi_1$.

1. We assume that A with domain $D(A)$ generates a positive semigroup $(T(t))_{t \ge 0}$ on $X = C[0,1]$. By definition, one has, for $f \in D(A)$,

$$(Af)(1) = \lim_{t \to 0} \frac{1}{t} [< f - f_1, T(t)'\delta_1 > + < f_1, T(t)'\delta_1 > - f(1)],$$

with f_1 the constant function $f_1(s) = f(1)$ for all $s \in [0,1]$. Let $z_1 \in (a_k, 1)$. Then there exist probability measures μ_{1_t} with *supp* $\mu_{1_t} \subset [z_1, 1]$, $\mu_{1_t}(\{1\}) = 0$ and ν_{1_t} with *supp* $\nu_{1_t} \subset [0, z_1]$, and the positive functions $q(t), \bar{p}(t)$ for $t > 0$ such that $T(t)'\delta_1 = q(t)\nu_{1_t} + \bar{p}(t)\mu_{1_t}$. By hypothesis $\varphi_1(s) < \infty$ for $s \in [z_1, 1]$

and $\varphi_1(1)-\varphi_1(s)>0$ for $s\in[z_1,1]$ since $m>0$ on $(a_k,1)$. Thus the function $p(t):=(\varphi_1(1)-\varphi_1(s))\bar{p}(t)$ is positive as well and we obtain

$$(Af)(1)=\lim_{t\to 0}\frac{1}{t}\Big[\int_{z_1}^{1}\frac{f(s)-f(1)}{\varphi_1(1)-\varphi_1(s)}p(t)d\mu_{1_t}(s)+\int_0^{z_1}f(s)q(t)d\nu_{1_t}(s)+ \\ +[\int_{z_1}^{1}d\mu_{1_t}(s)\bar{p}(t)-1]f(1)\Big].$$

We conclude similarly to Theorem 2.5 that $D(A)\subset\ker\Phi_1$.

2. We prove that A with domain $\ker\Phi_1$ verifies the conditions of Theorem 1.1, i.e., it generates a positive semigroup.

a) $(\lambda-A)$ is surjective for λ big enough. For $x\in[0,a_k]$ we take over the considerations of case 1. Let $x\in[a_k,1]$. Then the solution of the equation $\lambda f-mf'=g$ for every $g\in C[a_k,1]$ is given by

$$f(x)=ke^{-\lambda\int_x^1\frac{dy}{m(y)}}+\int_x^1\frac{g(t)}{m(t)}e^{-\lambda\int_x^t\frac{dy}{m(y)}}dt,\quad x\in[a_k,1].$$

Note, that by hypothesis, $\int_{z_1}^1 1/m<\infty$ and $\int_{a_k}^1 1/m=\infty$, $z_1\in(a_k,1)$ because m is admissible. Then the solution f is in $C_m^1[a_k,1]$, moreover, it is in $\ker\Phi_1$ if and only if

$$k\Phi_1\left(e^{-\lambda\int_x^1 1/m}\right)+\Phi_1\left(\int_x^1\frac{g(t)}{m(t)}e^{-\lambda\int_x^t 1/m}dt\right)=0.$$

Since for $z_1\in(a_k,1)$,

$$\Phi_1\left(e^{-\lambda\int_x^1 1/m}\right)=d_1'\lambda+\int_{z_1}^1\frac{1-e^{-\lambda\int_s^1 1/m}}{\varphi_1(1)-\varphi_1(s)}d\mu_1(s)- \\ -\int_{a_k}^{z_1}e^{-\lambda\int_s^1 1/m}d\nu_1(s)+d_1\to\infty\quad\text{as}\quad\lambda\to\infty,$$

the constant k is uniquely determined. Moreover, since $\int_{a_k}^z 1/m=\infty$, again $(Af)(a_k)=0$ and $f(a_k)=\frac{g(a_k)}{\lambda}$. If we join together the solutions on the subintervals of $[0,1]$, we obtain a continuous function $f\in C[0,1]$ that belongs to $D(A)$. The computations for b) and c) are similar to those from the proof of Theorem 2.5. The proof of case 3) and case 4) works with the same techniques.■

It remains to specify the situations when $m(0)$ and $m(1)$ respectively vanish. Using the same proof techniques as in the previous theorem we can show the following characterizations. We start with the case when we do not need boundary conditions.

Theorem 2.20. *Let $m(0)=0$ and $m(1)=0$, $z_0\in(0,a_1)$, $z_1\in(a_k,1)$ and consider the maximal operator A_m defined on $X=C[0,1]$.*

Case 1. If $m>0$ on $(0,a_1)$ and $m<0$ on $(a_k,1)$ then A_m generates a positive semigroup on X.

Case 2. If $m<0$ on $(0,a_1)$ and $m>0$ on $(a_k,1)$ then A_m generates a positive semigroup on X if and only if

$$\int_0^{z_0}\frac{1}{|m|}=\int_{z_1}^1\frac{1}{m}=\infty.\qquad ■$$

Note that if $m>0$ on $(0,a_1)$ and $m<0$ on $(a_k,1)$ then it is not relevant whether m vanishes in the endpoints of the interval $[0,1]$ or not. Next we describe the situation with positive boundary conditions.

Theorem 2.21. *Let $m(0) = 0$ and $m(1) = 0$. Consider a restriction A of the maximal operator A_m defined on $X = C[0,1]$.*

Case 1. If $m < 0$ on $(0, a_1)$, $\int_0^{z_0} \frac{1}{|m|} < \infty$ and $m > 0$ on $(a_k, 1)$, $\int_{z_1}^1 \frac{1}{m} < \infty$ then A generates a positive semigroup on X if and only if $D(A) = \ker \Phi_0 \cap \ker \Phi_1$.

Case 2. If $m < 0$ on $(0, a_1)$, $\int_0^{z_0} \frac{1}{|m|} < \infty$ and $m < 0$ on $(a_k, 1)$ then A generates a positive semigoup on X if and only if $D(A) = \ker \Phi_0$.

Case 3. If $m < 0$ on $(0, a_1)$, $\int_0^{z_0} \frac{1}{|m|} < \infty$ and $m > 0$ on $(a_k, 1)$ with $\int_{z_1}^1 \frac{1}{m} = \infty$ then A generates a positive semigoup on X if and only if $D(A) = \ker \Phi_0$.

Case 4. If $m > 0$ on $(0, a_1)$ and $m > 0$ on $(a_k, 1)$, $\int_{z_1}^1 \frac{1}{m} < \infty$ then A generates a positive semigoup on X if and only if $D(A) = \ker \Phi_1$.

Case 5. If $m < 0$ on $(0, a_1)$, $\int_0^{z_0} \frac{1}{|m|} = \infty$ and $m > 0$ on $(a_k, 1)$, $\int_{z_1}^1 \frac{1}{m} < \infty$ then A generates a positive semigoup on X if and only if $D(A) = \ker \Phi_1$. ∎

To get an overview of the results, let us summarize them briefly. We need a boundary condition at $x = 0$ if

i) $m(0) < 0$, or

ii) $m(0) = 0$, $m < 0$ on $(0, a_1)$ and $\int_0^{z_0} 1/|m| < \infty$,

and we need a boundary condition at $x = 1$ if

i) $m(1) > 0$, or

ii) $m(1) = 0$, $m > 0$ on $(a_k, 1)$ and $\int_{z_1}^1 1/m < \infty$.

The above theorems give an explicit form of *all* possible boundary conditions. In all other situations there is no need for boundary conditions at the endpoints $x = 0$ and $x = 1$.

C.J.K. Batty in [Ba] and R. DeLaubenfels in [De] have studied the question when the maximal operator A_m generates a strongly continuous semigroup and a strongly continuous group respectively. They considered only those cases when we do not need boundary conditions, but allowed for this situation a more general multiplicative perturbation m. For illustration we give one result of Batty; it includes Theorem 4 of [De], that gives necessary and sufficient conditions for a "well behaved $*$-derivation" to generate a positive contractions semigroup.

C.J.K. Batty proved in [Ba] that the closure of the derivation $Af : C_c^1(\mathbb{R}) \to C_0(\mathbb{R})$ defined by

$$Af(x) = \begin{cases} m(x)f'(x) & \text{if} \quad m(x) \neq 0 \\ 0 & \text{if} \quad m(x) = 0, \end{cases}$$

where $m : \mathbb{R} \to \mathbb{R}$ is continuous, generates a strongly continuous group on $C_0(\mathbb{R})$ (corresponding to a flow on $\mathbb{R}$) if and only if $1/m$ is not locally integrable on either side of any zero of m or at $\pm\infty$. (In our terminology, this means m is admissible.) For the semigroup case, we need the following notations.

If a, b are two consecutive zeroes of m, then we mean by $1/m$ is not locally integrable at b^- that $\int_z^b \frac{dx}{|m(x)|} = \infty$ for $z \in (a, b)$. Similarly, $1/m$ is not locally integrable at a^+ if $\int_a^z \frac{dx}{|m(x)|} = \infty$ for $z \in (a, b)$.

The following proposition holds, see[Ba].

Proposition 2.22. *Let $m : \mathbb{R} \to \mathbb{R}$ be continuous. The following are equivalent.*

i) $\overline{A|_{C_c^\infty(\mathbb{R})}}$ *generates a semigroup on* $C_0(\mathbb{R})$.

ii) Let a and b be two consecutive zeroes of m. Then if $m > 0$ on (a, b), $1/m$ is not locally integrable at b^-; if $m < 0$ on (a, b), $1/m$ is not locally integrable at a^+. ∎

The above mentioned results of Batty have analogues for $x \in [0, 1]$ in [Ba]. The one corresponding to Proposition 2.22 is

Proposition 2.23. *Let $m \in C[0, 1]$. The follwoing statements are equivalent.*

i) $\bar{A}$ generates a strongly continuous semigroup on $C[0, 1]$.

ii) $m(0) \geq 0$, $m(1) \leq 0$ and let a and b be two consequently zeroes of m. Then if $m > 0$ on (a, b), $1/m$ is not locally integrable at b^-; if $m < 0$ on (a, b), $1/m$ is not locally integrable at a^+. ∎

Let us compare our results with those mentioned before. In Proposition 2.23 it is assumed that $1/m$ is not locally integrable only in the neighbourhood of those zeroes of m, where m changes the sign from + to -. For the other situation, $1/m$ might be also locally integrable in a neighbourhood of a zero of m. These assumptions are weaker than the ones we use in this section (see Definition 2.16).

We also point out that in [Ba] the case when we need boundary conditions is not considered. Secondly, the operator is defined in a different manner. For those x for which $m(x) = 0$, $Af(x)$ is zero by definition, while we define $Af(x) = m(x)f'(x)$ for $f \in D(A_m)$. Therefore it is not obvious how to weaken the assumption concerning m in the theorems we proved. The reason is that for those zeroes of m, where m changes the sign from - to +, we would not longer get that $\lambda - A$ is surjective for λ big enough. One difficulty occurs from the fact that we did not define the differential operator $(Af)(x) = 0$ for those x for which $m(x) = 0$. If m is admissible in every a_i, $i = 1, ..., k$, then due to Remark 2.17 we obtain that $(Af)(a_i) = 0$, but if $1/m$ is locally integrable in a neighbourhood of one a_j, $j = 1, ..., k$, then it is not clear which value $(Af)(a_j)$ takes.

3. THE SECOND ORDER DIFFERENTIAL OPERATOR

Throughout this section we work on the same Banach space $X = C[0, 1]$ and consider the maximal operator $A_m f = mf'' + nf' + pf$, with $m, n, p \in C[0, 1]$, $m(x) > 0$ for all $x \in (0, 1)$ with domain $D(A_m) = \{f \in C[0, 1] \cap C^2(0, 1) : \lim_{x \to 0} Af(x)$ and $\lim_{x \to 1} Af(x)$ exist$\} =: C_m^2[0, 1]$. Of concern is the question whether A_m generates a *positive* semigroup or not, and if not, which are the restrictions A of A_m that have the required property.

Difficulties arise if $m(0) = 0$ or $m(1) = 0$. The problem of characterizing all domains $D(A) \subset C_m^2[0, 1]$ such that a restriction A of A_m generates a *contractive* semigroup was already solved (see §3.4). Our point here is not to rederive these results. We give a different approach, and the results we obtain are easier to handle. Anyway we shall have to impose stronger assumptions on the differentiability of m and n. Mainly we follow the scheme of Section 2, and use the proof techniques developed there.

3.1 Preliminaries. We start with a characterization of all domains $D(B) \subset D(B_{\max}) = C^2[a,b]$ such that the operator $Bf = f''$ generates a *positive* semigroup on $X = C[a,b]$ with $-\infty < a < b < \infty$, see [St].

Definition 3.1. A *positive boundary condition at* $x = a$ *(*$x = b$ *respectively)* is a linear form Φ_a (Φ_b) on $C^2[a,b]$ given by

$$\begin{aligned} \Phi_a f &= \alpha f''(a) + \int_a^b \frac{f(a)-f(s)}{a-s} d\mu(s) + \beta f(a), \\ (\Phi_b f &= \gamma f''(b) + \int_a^b \frac{f(b)-f(s)}{b-s} d\nu(s) + \delta f(b)), \end{aligned} \tag{3.1}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, $\alpha, \gamma \geq 0$, μ and ν positive bounded Borel measures on $[a,b]$ with $\mu(a) = 0$, $\nu(b) = 0$ and

$$\begin{aligned} \alpha = 0 \quad &\text{implies} \quad \int_a^b \frac{1}{s-a} d\mu(s) = \infty \\ (\gamma = 0 \quad &\text{implies} \quad \int_a^b \frac{1}{b-s} d\nu(s) = \infty). \end{aligned} \tag{3.2}$$

Theorem 3.2. *A restriction B of $B_{\max}$ generates a positive semigroup on $C[a,b]$ if and only if there are positive boundary conditions Φ_a at $x = a$ and Φ_b at $x = b$ such that $D(B) = \ker \Phi_a \bigcap \ker \Phi_b$.* ■

Our aim is to obtain similar results for the operator A defined at the beginning of this section. We try to reduce this case via similarity transformations to situations where our central question is solved. For this purpose we need to extend the above theorem to the Banach spaces $C[-\infty,\infty]$, $C[-\infty,b]$ and $C[a,\infty]$, $a, b \in \mathbb{R}$ and finite. We omit the proofs since they are very similar to those of Section 2.

First we consider the Banach space $X = C[-\infty,\infty] = \{f \in C(-\infty,\infty) : \lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ exist$\}$. Then the domain of the maximal operator $B_{\max} f = f''$ is $D(B_{\max}) = \{f \in C[-\infty,\infty] \cap C^2(-\infty,\infty) : \lim_{x\to\infty} f''(x)$ and $\lim_{x\to-\infty} f''(x)$ exist$\}$. For this situation the following proposition holds (see [Yo], p.235).

Proposition 3.3. *The maximal operator $B_{\max}$ generates a positive semigroup on X.* ■

Consider the Banach space $X = C[-\infty, b]$, $b < \infty$, $b \in \mathbb{R}$ and the maximal operator $B_{\max} f = f''$ be defined on X with maximal domain $D(B_{\max}) = \{f \in C[-\infty,b] \cap C^2(-\infty,b] : \lim_{x\to-\infty} f(x)$ exists$\}$. Let $z \in (-\infty, b)$ be fixed.

Definition 3.4. Let $z \in (-\infty, b)$ be fixed. A *positive boundary condition at* $x = b$ is a linear form Φ_b on $D(B_{\max})$ given by

$$\Phi_b f = \gamma f''(b) + \int_z^b \frac{f(b)-f(s)}{b-s} d\nu(s) - \int_{-\infty}^z f(s) d\mu(s) + \delta f(b), \tag{3.3}$$

with $\gamma, \delta \in \mathbb{R}$, $\gamma \geq 0$, μ and ν positive bounded Borel measures on $[-\infty, z]$ and $[z, b]$ respectively, $\nu(b) = 0$ and

$$\gamma = 0 \quad \text{implies} \quad \int_z^b \frac{1}{b-s} d\nu(s) = \infty. \tag{3.4}$$

Theorem 3.5. *A restriction b of $B_{\max}$ generates a positive semigroup on $C[-\infty, b]$ if and only if there is a positive boundary condition Φ_b at $x = b$ such that $D(B) = \ker \Phi_b$.* ■

It remains to give a characterization for the Banach space $X = C[a, \infty]$ with $-\infty < a \in \mathbb{R}$. Let the maximal operator $B_{\max} f = f''$ be defined on X with domain $D(B_{\max}) = \{f \in C[a, \infty] \cap C^2[a, \infty) : \lim_{x \to \infty} f(x) \text{ exists}\}$ and consider $z \in (a, \infty)$ be fixed.

Definition 3.6. Let $z \in (a, \infty)$ be fixed. A *positive boundary condition at* $x = a$ is a linear form Φ_a defined on the maximal domain $D(B_{\max})$ given by

$$\Phi_a f = \alpha f''(a) + \int_a^z \frac{f(a) - f(s)}{a - s} d\nu(s) - \int_z^\infty f(s) d\mu(s) + \beta f(a), \tag{3.5}$$

with $\alpha, \beta \in \mathbb{R}$, $\alpha \geq 0$, μ and ν positive bounded Borel measures defined on $[z, \infty]$ and $[a, z]$ respectively, with $\nu(a) = 0$ and

$$\alpha = 0 \quad \text{implies} \quad \int_a^z \frac{d\nu(s)}{s - a} = \infty. \tag{3.6}$$

Theorem 3.7. *A restriction B of $B_{\max}$ generates a positive semigroup on $C[a, \infty]$ if and only if there exists a positive bouundary condition at $x = a$ such that $D(B) = \ker \Phi_b$.* ■

3.2 The similarity transformation. Our objective is now to transform the operator A_m via similarity transformations into an operator $B_{\max}$, where

> $A_m f = mf'' + nf' + pf$, with $m, n, p \in C[0,1]$ and $m(x) > 0$ for all $x \in (0,1)$ defined on $C[0,1]$, with domain $D(A_m) = \{f \in C[0,1] \cap C^2(0,1) : \lim_{x \to 0} A_m f(x) \text{ and } \lim_{x \to 1} A_m f(x) \text{ exist}\}$,

and

> $B_{\max} f = f'' + cf$, defined on $C[a, b]$, $-\infty \leq a < b \leq \infty$, $c \in C[a, b]$ with domain $D(B_{\max}) = \{f \in C[a, b] \cap C^2(a, b) : \lim_{x \to a} f''(x) \text{ and } \lim_{x \to b} f''(x) \text{ exist}\}$.

If $\tilde{B} f = f''$ defined on $C[a, b]$ with domain $D(\tilde{B})$ generates a semigroup, then due to perturbation results (see [Na], p.44) the restriction B of $B_{\max}$ defined on $C[a, b]$ with domain $D(B) = D(\tilde{B})$ also does. Moreover, B generates a *positive* semigroup if and only if $\tilde{B}$ does. This follows immediately from Theorem 1.1.

For our purpose we need two transformation V_1 and V_2. We define the operator $C := V_2^{-1} V_1^{-1} A V_1 V_2$ with domain $D(C) := \{f \in C[a, b] : V_1 V_2 f \in D(A)\}$ and determine V_1, V_2 such that C coincides with $B_{\max}$. The first transformation V_1 reduces our problem to the case $m(x) = 1$ while the second transformation V_2 anihilates the coefficient of f'.

We consider $m \in C[0,1]$, $m(x) > 0$ for $x \in (0,1)$. Moreover, we fix $z \in (0, x)$ and define $\varphi(x) := \int_z^x \frac{dy}{\sqrt{m(y)}}$ with $a := \varphi(0)$, $b := \varphi(1)$ $\psi := \varphi^{-1}$ and $V_1 f := f \circ \varphi$. The mapping φ defines a homeomorphism from $[0,1]$ to $[a, b]$ and a diffeomorphism from $(0,1)$ to (a, b). The transformation V_1 is even a lattice isomorphism from $C[a, b]$ to $C[0,1]$. The role of the second transformation $V_2 f := Mf$ with $M \gg 0$, $M \in C[a, b]$ is, as already mentioned, to annihilate the coefficient of f'. In this sense we shall determine M later. Let us start computing.

$$\begin{aligned} AV_1 f &= A(f \circ \varphi) \\ &= m((f'' \circ \varphi)\varphi'^2 + (f' \circ \varphi)\varphi'') + n(f' \circ \varphi)\varphi' + p(f \circ \varphi) \\ &= f'' \circ \varphi + m\varphi''(f' \circ \varphi) + n\varphi'(f' \circ \varphi) + p(f \circ \varphi) \\ V_1^{-1} AV_1 f &= f'' + [(m\varphi'' + n\varphi') \circ \psi] f' + (p \circ \psi) f \end{aligned}$$

$$\begin{aligned} V_1^{-1} AV_1 V_2 f &= M'' f + 2M' f' + M f'' \\ &\quad + [(m\varphi'' + n\varphi') \circ \psi](M' f + m f') + (p \circ \psi) M f \\ &= M f'' + \{2M' + M[(m\varphi'' + n\varphi') \circ \psi]\} f' \\ &\quad + \{M'' + M'[(m\varphi'' + n\varphi') \circ \psi] + M(p \circ \psi)\} f \\ V_2^{-1} V_1^{-1} AV_1 V_2 f &= f'' + \left[\frac{2M'}{M} + (m\varphi'' + n\varphi') \circ \psi\right] f' \\ &\quad + \left\{\frac{M''}{M} + \frac{m'}{M}[(m\varphi'' + n\varphi') \circ \psi] + p \circ \psi\right\} f. \end{aligned}$$

The condition $\frac{2M'}{M} + (m\varphi'' + n\varphi') \circ \psi = 0$ leads to the explicit form of M:

$$M(x) = e^{-\frac{1}{2}\int_z^x [(m\varphi'' + n\varphi') \circ \psi](y)dy} \tag{3.7}$$

for $y \in (0, x)$, $x \in [a, b]$, and $z \in (0, 1)$. With M defined in (2.7) the operator C becomes

$$V_2^{-1} V_1^{-1} AV_1 V_2 f = f'' + \left[\frac{M''}{M} - 2\Big(\frac{M'}{M}\Big)^2 + p \circ \psi\right] f. \tag{3.8}$$

The transformations serve our purpose only if the additive perturbation of the second derivative is a bounded operator, i.e.,

$$\frac{M''}{M} - 2\Big(\frac{M'}{M}\Big)^2 + p \circ \psi \in C[a, b]. \tag{3.9}$$

This is the condition that requires the stronger assumptions $m \in C^2[0, 1]$, $n \in C^1[0, 1]$, $p \in C[0, 1]$. Moreover, we have to make sure that M is indeed strictly positive and belongs to $C[a, b]$; since $\varphi' = \frac{1}{\sqrt{m}}$ implies $\varphi'' = \frac{-m'}{2m\sqrt{m}}$ the representation (3.7) becomes

$$\begin{aligned} M(x) &= e^{-\frac{1}{2}\int_z^x (\frac{m'}{2\sqrt{m}} \circ \psi)(y)dy} \cdot e^{-\frac{1}{2}\int_z^x (\frac{n}{\sqrt{m}} \circ \psi)(y)dy} \\ &= K \underbrace{e^{\frac{1}{2}\sqrt{(m \circ \psi)(x)}}}_{I_1} \cdot \underbrace{e^{-\frac{1}{2}\int_z^x (\frac{n}{\sqrt{m}} \circ \psi)(y)dy}}_{I_2}, \end{aligned}$$

for all $x \in [a, b]$, $y \in (z, x)$ and $z \in (0, 1)$. The term I_1 is bounded and differs from zero even if $m(z) = 0$ for $z = 0, 1$. The term I_2 has also to be bounded and nonzero, thus we impose the condition

$$\left|\int_z^x \Big(\frac{n}{\sqrt{m}} \circ \psi\Big)(y)dy\right| < \infty \tag{3.10}$$

for $z \in (a,b)$, $x \in [a,b]$ and $y \in (z,x)$. Note that $\psi'(y) = \frac{1}{\varphi'(\psi(y))} = \sqrt{m(\psi(y))}$, and substituting $\psi(y) = s$, $s \in [0,1]$, one obtains

$$\int \frac{n(\psi(y))}{\sqrt{m(\psi(y))}}dy = \int \frac{n(\psi(y))}{\sqrt{m(\psi(y))}}\frac{1}{\sqrt{m(\psi(y))}}\psi'(y)dy = \int \frac{n(s)}{m(s)}ds.$$

With the above computations, condition (3.10) reads

$$\left|\int_0^z \frac{n}{m}\right| < \infty, \qquad \left|\int_z^1 \frac{n}{m}\right| < \infty \quad \text{for all} \quad z \in (0,1).$$

Because $m(x) > 0$ for all $x \in (0,1)$, the above conditions reduce to

$$\begin{aligned} n(1) \geq 0 \quad &\text{implies} \quad \int_z^1 \frac{n}{m} < \infty \\ n(1) \leq 0 \quad &\text{implies} \quad \int_z^1 \frac{n}{m} > -\infty \\ n(0) \geq 0 \quad &\text{implies} \quad \int_0^z \frac{n}{m} < \infty \\ n(0) \leq 0 \quad &\text{implies} \quad \int_0^z \frac{n}{m} > -\infty. \end{aligned}$$

Note that if $n(x) = 0$, for all $x \in [0,1]$, then the above conditions are always satisfied. (For example $Af = mf'' + pf$.)

Transformation of domains. We compute $D(C)$, and prove that it coincides with $D(B_{\max})$.

$$\begin{aligned} D(C) &= \{f \in C[a,b] : \; V_1V_2f \in D(A_m)\} \\ &= \{f \in C[a,b] : \; (Mf)\circ\varphi \in D(A_m)\} \\ &= \{f \in C[a,b] : \; Mf\circ\varphi \in C[0,1] \cap C^2(0,1), \; A(Mf\circ\varphi) \in C[0,1]\}. \end{aligned}$$

Since we assumed $m \in C^2[0,1]$, φ is a C^2-diffeomorphism from $(0,1)$ to (a,b), and because φ is a homeomorphism from $[0,1]$ into $[a,b]$, we obtain

$$\begin{aligned} D(C) = \{f \in C[a,b] : \; &Mf \in C[a,b] \cap C^2(a,b), \\ &Mf'' + (M'' - 2\frac{M'^2}{M} + (p\circ\psi)M)f \in C[a,b]\}. \end{aligned}$$

Now we use (3.9) and the fact that $M \in C[a,b] \cap C^2(a,b)$; thus

$$\begin{aligned} D(C) &= \{f \in C[a,b] \cap C^2(a,b) : \; Mf \in C[a,b] \cap C^2(a,b), \; Mf'' \in C[a,b]\} \\ &= \{f \in C[a,b] \cap C^2(a,b) : \; f'' \in C[a,b]\} \\ &= \{f \in C[a,b] \cap C^2(a,b) : \; \lim_{x\to a} f''(x) \text{ and } \lim_{x\to b} f''(x) \in C[a,b]\} \\ &= D(B_{\max}). \end{aligned}$$

Let A_m be the operator defined at the beginning of this section. We have proved the following proposition.

Proposition 3.8. *Let $m \in C^2[0,1]$, $n \in C^1[0,1]$, $p \in C[0,1]$ and $m(x) > 0$ for $x \in (0,1)$ such that $|\int_z^1 n/m| < \infty$ and $|\int_0^z n/m| < \infty$ for $z \in (0,1)$. Consider the operator B on $C[a,b]$ given by $B = V_2^{-1}V_1^{-1}A_mV_1V_2$.*

Then $Bf = f'' + cf$ with $c \in C[a,b]$, $-\infty \le a < b \le \infty$, and the domain of B coincides with its maximal domain.

Remark 3.9. This transformation is positivity preserving. ■

3.3 The main result. Let the Banach space X be $C[0,1]$ and consider the differential operator

$$\begin{aligned} Af(x) &= m(x)f''(x) + n(x)f'(x) + p(x)f(x), \\ D(A_m) &= \{f \in C[0,1] \cap C^2(0,1) : \lim_{x\to 0} Af(x) \text{ and } \lim_{x\to 1} Af(x) \text{ exist}\} \end{aligned} \tag{3.11}$$

where m, n, p always satisfy the assumptions

$$\begin{aligned} &m \in C^2[0,1],\ n \in C^1[0,1],\ p \in C[0,1], \\ &m(x) > 0 \text{ for all } x \in (0,1), \\ &\left|\int_z^1 \frac{n(s)}{m(s)} ds\right| < \infty, \quad \left|\int_0^z \frac{n(s)}{m(s)} ds\right| < \infty \quad \text{for } z \in (0,1). \end{aligned} \tag{3.12}$$

The next results give necessary and sufficient conditions such that a restriction A of A_m (or even A_m) generates a positive strongly continuous semigroup. All proofs are based on the same idea, namely we reduce our problem via similarity transformations to cases where our problem is solved (see Theorem 3.5, 3.7 and Proposition 3.3). In Section 3.2 we proved that the operator (3.11) satisfying the conditions (3.12) can be transformed into the similar maximal operator

$$\begin{aligned} B_{\max}f(x) &= f''(x) + c(x)f(x), \quad c \in C[a,b] \\ D(B_{\max}) &= \{f \in C[a,b] \cap C^2(a,b) : \lim_{x\to a} B_{\max}f(x) \text{ and} \\ &\qquad \lim_{x\to b} B_{\max}f(x) \text{ exist}\}. \end{aligned} \tag{3.13}$$

In the following we keep the notations from Section 3.2. The necessity of a boundary condition at an endpoint will depend upon whether $\varphi(0)$ and $\varphi(1)$ respectively are finite or not. We have four situations, and treat each of them separately.

Theorem 3.10. *The maximal operator A_m generates a positive semigroup if and only if $|\varphi(0)| = \varphi(1) = \infty$, i.e.,*

$$\int_0^z \frac{1}{\sqrt{m(x)}} dx = \int_z^1 \frac{1}{\sqrt{m(x)}} dx = \infty \quad \textit{for } z \in (0,1).$$

Proof. Assume that $a = \varphi(0) = -\infty$ and $b = \varphi(1) = \infty$. We use the similarity transformation described in Section 3.2 and obtain an operator $B_{\max}f = f'' + cf$, $c \in C[a,b]$, defined on $C[-\infty,\infty]$. Proposition 3.3 assures the statement. Conversely, let us assume the contrary, i.e., $b \neq \infty$, Then $e^{\sqrt{\lambda}x}$ satisfies $\lambda f - f'' = 0$ thus $\sigma(B_{\max})$ would contain $\mathbb{R}_+$ which contradicts the generator property of $B_{\max}$, and thus of A_m. ■

Definition 3.11. Let $\varphi(0) = -\infty$ and $\varphi(1) < \infty$ and $z \in (0,1)$ be fixed. A *positive boundary condition at* $x = 1$ is a linear form Φ_1 defined on $D(A_m)$ given by

$$\Phi_1 f = c_1 Af(1) + \int_z^1 \frac{f(1) - f(s)}{\varphi(1) - \varphi(s)} d\nu(s) - \int_0^z f(s) d\mu(s) + c_2 f(1), \qquad (3.14)$$

with $c_1, c_2 \in \mathbb{R}$, $c_1 \geq 0$, μ and ν bounded positive Borel measures defined on $[0,z]$ and $[z,1]$ respectively, with $\nu(\{1\}) = 0$ and

$$c_1 = 0 \quad \text{implies} \quad \int_z^1 \frac{d\nu(s)}{\varphi(1) - \varphi(s)} = \infty. \qquad (3.15)$$

Theorem 3.12. *Let* $\varphi(0) = -\infty$ *and* $\varphi(1) < \infty$. *A restriction* A *of* A_m *generates a positive semigroup on* $C[0,1]$ *if and only if there is a linear form* Φ_1 *at* $x = 1$ *(given by Definition 3.11) such that* $D(A) = \ker \Phi_1$.

Proof. The main idea of the proof is to transform the claim of the theorem via the similarity transformations described in Section 3.2 into the statement of Theorem 3.5. The operator $A \subset A_m$ with domain $D(A)$ generates a positive semigroup if and only if the operator $B \subset B_{\max}$ with domain $D(B) = \{f \in C[a,b] : V_1 V_2 f \in D(A)\}$ does. In the following we compute the domain of B.

1. Assume now that there exist constants $c_1, c_2 \in \mathbb{R}$, $c_1 \geq 0$, and positive bounded Borel measures μ and ν such that (3.14) and (3.15) hold. Then

$$\begin{aligned} D(B) &= \{f \in C[a,b] : V_1 V_2 f \in D(A)\} \\ &= \{f \in D(B_{\max}) : \Phi_1[(Mf) \circ \varphi] = 0\}, \end{aligned}$$

with $\varphi(0) = a = -\infty$ and $\varphi(1) = b < \infty$. We compute $\Phi_1[(Mf) \circ \varphi]$ separately.

$$\begin{aligned} \Phi_1(Mf \circ \varphi) &= c_1 [A(Mf \circ \varphi)](1) + \int_z^1 \frac{Mf(\varphi(1)) - Mf(\varphi(s))}{\varphi(1) - \varphi(s)} d\nu(s) \\ &\quad - \int_0^z Mf(\varphi(s)) d\mu(s) + c_2 (Mf)(\varphi(1)) \\ &= c_1 M(b) f''(b) + \int_z^1 \frac{M(b) f(b) - M(\varphi(s)) f(\varphi(s))}{\varphi(1) - \varphi(s)} d\nu(s) \\ &\quad - \int_0^z M(\varphi(s)) f(\varphi(s)) d\mu(s) + c_2 M(b) f(b) + c(b) M(b) f(b). \end{aligned}$$

Here we used the fact that

$$\begin{aligned} [A(Mf \circ \varphi)](1) &= (A V_1 V_2 f)(1) \\ &= (V_1 V_2 B f)(1) \\ &= V_1 V_2 (f'' + cf)(1) \\ &= [(Mf'') \circ \varphi + (c \cdot Mf) \circ \varphi](1), \end{aligned}$$

where c coincides with expression (3.9). We divide the boundary condition equality with $M(b) > 0$ and obtain

$$\begin{aligned} & c_1 f''(b) + \int_z^1 \frac{f(b) - f(\varphi(s))}{\varphi(1) - \varphi(s)} \frac{M(\varphi(s))}{M(b)} d\nu(s) - \int_0^z f(\varphi(s)) \left[\frac{M(\varphi(s))}{M(b)} d\mu(s) \right] \\ & + \left[\frac{1}{M(b)} \int_z^1 \frac{M(b) - M(\varphi(s))}{\varphi(1) - \varphi(s)} d\nu(s) + c(b) + c_2 \right] f(b) = 0. \end{aligned}$$

Now we substitute $\varphi(s) = t,\ s \in [0,1], t \in [a,b]$ and redefine the constants. Then

$$c_1 f''(b) + \int_z^1 \frac{f(b)-f(t)}{b-t}\left[\frac{M(t)}{M(b)}d\tilde{\nu}(t)\right] - \int_0^z f(t)\left[\frac{M(t)}{M(b)}d\tilde{\mu}(t)\right] + \gamma f(b) = 0$$

with $c_1, \gamma \in \mathbb{R},\ c_1 \geq 0$, where $\tilde{\nu}(t) = \nu(\psi(t))$ is the image measure of ν through the continuous mapping φ. Since $M \in C[a,b]$ is strictly positive, $\frac{M(t)}{M(b)}d\tilde{\nu}(t)$ is again a bounded positive Borel measure defined on $[y,b]$ for a fix $y \in (a,b)$ with no mass in b, and $\frac{M(t)}{M(b)}d\tilde{\mu}(t)$ a positive bounded Borel measure on $[-\infty, y]$. Relation (3.15) turns over into (3.4) exactly, that is

$$c_1 = 0 \quad \text{implies} \quad \int_y^b \frac{d\tilde{\nu}(t)}{b-t} = \infty.$$

We are in the framework of Theorem 3.5, which proves the claim.

2. For the other implication we can either use the proof idea from Theorem 2.5, or a similarity transformations argument. ∎

For the following two cases we give no proofs since the idea and the computations as well are analogous to the last ones. In the first situation we transform the characterization to one for an operator $Bf = f''$ defined on $C[a,\infty]$, $-\infty < a$. Thus we make use of Theorem 3.7, and in the second case we utilize the characterization for the operator $Bf = f''$ defined on $C[a,b]$, $-\infty < a < b < \infty$, given in Theorem 3.2.

Definition 3.13. Let $|\varphi(0)| < \infty$ and $\varphi(1) = \infty$ and $z \in (0,1)$ be fixed. A *positive boundary condition* at $x = 0$ is a linear form Φ_0 defined on $D(A_m)$ given by

$$\Phi_0 f = d_1 Af(0) + \int_0^z \frac{f(0)-f(s)}{\varphi(0)-\varphi(s)}d\tau(s) - \int_z^1 f(s)d\sigma(s) + d_2 f(0), \tag{3.16}$$

with $d_1, d_2 \in \mathbb{R},\ d_1 \geq 0$, τ and σ bounded positive Borel measures on $[0,z]$ and $[z,1]$ respectively with $\tau(\{0\}) = 0$ and

$$d_1 = 0 \quad \text{implies} \quad \int_0^z \frac{d\tau(s)}{\varphi(0)-\varphi(s)} = \infty. \tag{3.17}$$

Theorem 3.14. *Let $|\varphi(0)| < \infty$ and $\varphi(1) = \infty$. A restriction A of A_m generates a positive semigroup on $C[0,1]$ if and only if there is a linear form Φ_0 at $x = 0$ (see Definition 3.13) such that $D(A) = \ker \Phi_0$.*

Theorem 3.15. *If $\int_0^z \frac{1}{\sqrt{m(x)}}dx < \infty$, $\int_z^1 \frac{1}{\sqrt{m(x)}}dx < \infty$ for $z \in (0,1)$, then a restriction A of A_m generates a positive semigroup if and only if $D(A) = \{f \in D(A_m) : \Phi_0 f = \Phi_1 f = 0\}$ with Φ_0, Φ_1 satisfying (3.14)-(3.17).* ∎

Summarizing we can state that the necessity of a boundary condition at an endpoint of the underlying spatial interval is determined by the integrability of $1/\sqrt{m}$. But we have to keep in mind the assumptions (3.12).

3.4 Related results. With his paper [Fel] published in 1952, W. Feller started a program for a semigroup approach to determine all one dimensional

Markov processes of diffusion type ([Fe2],[Fe3],[Fe4]). His ideas were taken over and developed by many authors, for example P. Mandl [Ma], Ph. Clément and C.A. Timmermans [CT], K. Taira [Ta1], [Ta2], K. Ito [It] and many others.

Based on an idea of E. Hille (see [Hi]), W. Feller classified the boundary points r_1, r_2 of the underlying spatial interval $[r_1, r_2]$ into the regular, the exit, the entrance and the natural boundary as follows (see [Yo] p.403). We set formally

$$D_pD_q := m(x)\frac{d^2}{dx^2} + n(x)\frac{d}{dx},$$

where $B(x) := \int^x \frac{n(t)}{m(t)}dt$, $dp(x) := \frac{1}{m(x)}e^{B(x)}dx$, $dq(x) := e^{-B(x)}dx$, and finally $D_p := \frac{d}{dp(x)}$ and $D_q := \frac{d}{dq(x)}$. A function $f \in C([r_1, r_2], \mathbb{R})$ belongs to $\mathcal{D}(D_pD_q)$ if $D_pD_qf(x)$ exists for $x \in (r_1, r_2)$ and is continuous there, and if additionaly the limits of $D_pD_qf(r_i)$ exist for $i = 1, 2$. We introduce the four quantities

$$\sigma_1 = \iint_{r_1<y<x<z} dp(x)dq(y), \qquad \sigma_2 = \iint_{r_2>y>x>z} dp(x)dq(y)$$

$$\mu_1 = \iint_{r_1<y<x<z} dq(x)dp(y), \qquad \mu_2 = \iint_{r_2>y>x>z} dq(x)dp(y)$$

The boundary point r_i, $i = 1, 2$ is called

regular	in case	$\sigma_i < \infty,\ \mu_i < \infty$
exit	in case	$\sigma_i < \infty,\ \mu_i = \infty$
entrance	in case	$\sigma_i = \infty,\ \mu_i < \infty$
natural	in case	$\sigma_i = \infty,\ \mu_i = \infty.$

The regular and the exit boundary point are called *accessible* boundaries, and the other two *inaccessible* boundaries. W. Feller's probabilistic interpretation of the above classification is the following. The probability that a particle, located at first in the interior of the open interval (r_1, r_2), will, after a finite lapse of time, reach a regular boundary or an exit boundary point, is positive; while the particle can, after a finite lapse of time, neither reach an entrance boundary nor a natural boundary point.

It is natural to ask what kind of boundary points we have considered in the previous section. We shall prove that the question of integrability of $1/\sqrt{m}$ at the endpoints of the interval corresponds with the classification into accessible or inaccessible boundary points, i.e., if σ_i, $i = 1, 2$ is finite or not. Moreover, assumption (3.12) implies $|B(x)| < \infty$. Thus, in accordance with W. Feller's probabilistic interpretation, if at an endpoint $1/\sqrt{m}$ is integrable we have an accessible boundary, so the particle reaches the boundary, and we have to impose a boundary condition. If $1/\sqrt{m}$ is not integrable, we have an inaccessible boundary, the particle will not reach the boundary and we need no boundary conditions.

P. Mandl also gives a nomenclature of regular boundaries in dependence on the type of boundary condition (see [Ma], p.67).

1) If $\Phi_i(f) = f(r_i)$, the boundary point r_i is called *absorbing*. The trajectory vanishes after having reached the boundary.

2) If $\Phi_i(f) = D_pD_q(r_i)$, the boundary point r_i is called an *adhesive* boundary. The trajectory remains in the point r_i after having reached it.

3) If $\Phi_i(f) = D_p f(r_i)$, the boundary point r_i is called a *reflecting* boundary.

4) If $\Phi_i(f) = c_i f(r_i) - (-1)^i d_i D_p f(r_i)$, where $c_i, d_i > 0$, the boundary point r_i is called an *elastic* boundary.

5) If $\Phi_i(f) = c_i f(r_i) - \int f(x) d\mu_i(x) + d_i D_p D_q f(r_i)$ where $c_i, d_i > 0$, μ_i a positive bounded Borel measure on $[r_1, r_2]$, the boundary point r_i is called an *elementary return* boundary. When the trajectory reaches the boundary r_i it remains at r_i for a random interval of time, and then it vanishes.

a) The result of P. Mandl (see [Ma])

In his book "Analytical Treatment of One-Dimensional Markov Processes", P. Mandl characterized *all* domains of the operator (3.11), such that it generates a strongly continuous *contraction* semigroup by using W. Feller's classification of boundary points. If a boundary point r_i is an accessible (regular or exit) boundary point, then we need a boundary condition at $x = r_i$, and if the boundary point r_i is an inaccessible (natural or entrance) boundary point, then we do not have to impose a condition at $x = r_i$. Furthermore, he proved that all boundary conditions have a similar representation to (3.14)-(3.17). For the contraction property he assumed additionally that the coefficient of $f(r_i)$ in the boundary condition is nonnegative. The advantages of this results are the weak assumptions on the functions m and n, i.e., m and n are piecewise continuous and the limits $m(x \pm 0)$, $n(x \pm 0)$ are finite, $m(x \pm 0) > 0$ for $x \in (r_1, r_2)$. But actually many examples can be treated with the methods developed in §3.3, since the functions m, n are in general polynomial functions.

Let us now make a detailed analysis of the situation when $n(x) = 0$ for all $x \in [0,1]$. We consider $Af(x) = m(x) f''(x)$, $m \in C^2[0,1]$, $m(x) > 0$ for $x \in (0,1)$ defined on $C[0,1]$ with maximal domain $D(A_m) = \{f \in C[0,1] \cap C^2(0,1) : \lim_{x\to 0}(Af)(x)$ and $\lim_{x\to 1}(Af)(x)$ exists$\}$. P. Mandl affirms that A_m generates a positive semigroup if and only if 0 and 1 are inaccessible boundaries, i.e., for $z \in (0,1)$

$$\iint_{0<y<x<z} dp(x)dq(y) = \iint_{1>y>x>z} dp(x)dq(y) = \infty,$$

where $dp(x) = \frac{dx}{m(x)}$ and $dq(x) = dx$. On the other hand, Theorem 3.10 claims that A_m generates a positive semigroup if and only if

$$\int_0^z \frac{dx}{\sqrt{m(x)}} = \int_z^1 \frac{dx}{\sqrt{m(x)}} = \infty, \qquad \text{for} \quad z \in (0,1).$$

We prove that the two results coincide. We argue locally, e.g. for $x = 0$. Then, using Fubini's theorem we obtain

$$\iint_{0<y<x<z} dp(x)dq(y) = \iint_{0<y<x<z} \frac{1}{m(x)} dx dy = \int_0^z \left(\int_0^x dy \right) \frac{dx}{m(x)} = \int_0^z \frac{x}{m(x)} dx.$$

It remains to prove the following lemma.

Lemma 3.16. *Let $m \in C^2[0,1]$, $m(x) > 0$, $x \in (0,1)$. Then*

$$\int_0^z \frac{1}{\sqrt{m(x)}} dx = \infty \quad \textit{if and only if} \quad \int_0^z \frac{x}{m(x)} dx = \infty$$

for all $z \in (0,1)$.

Proof. Assume first $\int_0^z \frac{x}{m(x)}dx = \infty$. Then necessarily $m(0) = 0$ and $\frac{m(x)}{x} \to 0$ for $x \to 0$. We develop $m(x)$ using Taylor's formula $m(x) = m(0) + xm'(0) + \frac{x^2}{2}m''(\xi_x) = xm'(0) + \frac{x^2}{2}m''(\xi_x)$ for $\xi_x \in (0,1)$. Now $m \in C^2[0,1]$ implies $c_1 \le m''(\xi_x) \le c_2$ for all ξ_x. Moreover, $\frac{m(x)}{x} \to 0$ for $x \to 0$ implies $m'(0) = 0$. Then

$$\int_0^z \frac{1}{\sqrt{m(x)}}dx = \int_0^z \frac{1}{\sqrt{\frac{x^2}{2}m''(\xi_x)}}dx \ge \sqrt{2}\int_0^z \frac{dx}{\sqrt{c_2}\cdot x} = k(\ln x - \ln 0) = \infty,$$

with $k = \frac{\sqrt{2}}{\sqrt{c_2}}$. Thus $\int_0^z \frac{1}{\sqrt{m(x)}}dx = \infty$ for all $z \in (0,1)$.

Take now $\int_0^z \frac{dx}{\sqrt{m(x)}} = \infty$. This implies $m(0) = 0$ and we show that also $m'(0) = 0$. Assume the contrary, e.g. $m'(0) = 1$. Then $m(x) = m(0) + xm'(0) + \frac{x^2}{2}m''(\xi_x) = x + \frac{x^2}{2}m''(\xi_x)$, with $m''(\xi_x) \to m''(0)$. Since $m(x) \ge 0$ and $x \in [0,1]$ the above relation implies that $m''(\xi_x) \ge 0$. Let $m''(\xi_x) > 0$ in a neighbourhood of 0. Then $|x + \frac{x^2}{2}m''(\xi_x)| = x + \frac{x^2}{2}m''(\xi_x) \ge x$ and

$$\int_0^z \frac{dx}{\sqrt{m(x)}} = \int_0^z \frac{dx}{\sqrt{x + \frac{x^2}{2}m''(\xi_x)}} \le \int_0^z \frac{1}{\sqrt{x}}dx < \infty,$$

which contradicts the assumption. Thus $m'(0) = 0$ and $m(x) = \frac{x^2}{2}m''(\xi_x)$. Then $\frac{m(x)}{x} = \frac{x}{2}m''(\xi_x) \to 0$ for $x \to 0$ since $m \in C^2[0,1]$ and therefore $\int_0^z \frac{x}{m(x)}dx = \infty$. ∎

These computations also show that for the case $n(x) = 0$ for all $x \in [0,1]$ the integrability of σ_i is equivalent to the integrability of $1/\sqrt{m}$ at the respective endpoint. If $n(x) \ne 0$ for $x \in [0,1]$ then due to assumption (3.12) we have $|B(x)| < \infty$, therefore $e^{B(x)}$ and $e^{-B(x)}$ are bounded. We estimate σ_i for this case.

$$\begin{aligned}\sigma_0 &= \iint_{0<y<x<z} dp(x)dq(y) = \int_0^z \left[\int_y^z \frac{1}{m(x)}e^{-B(x)}dx\right] e^{B(y)}dy \\ &\le K\int_0^z \left[\int_y^z \frac{dx}{m(x)}\right] dy = K\int_0^z \left(\int_0^x dy\right)\frac{dx}{m(x)} \\ &= K\int_0^z \frac{x}{m(x)}dx, \qquad \text{with} \quad K \in \mathbb{R},\ K > 0.\end{aligned}$$

Here we used Fubini's theorem. A similar computation yields the other inequality $\sigma_0 \ge L\int_0^z \frac{x}{m(x)}dx$, where $L \in \mathbb{R}$, $L > 0$. These estimates, together with Lemma 3.11, prove that also in the case $n(x) \ne 0$ for $x \in [0,1]$, the integrability of σ_i is equivalent to the integrability of $1/\sqrt{m}$ at the respective endpoint.

Note that in the simple case when $Af(x) = f''(x)$ is defined on $C[0,1]$ with $D(A_m) = C^2[0,1]$, the points $x = 0$ and $x = 1$ are regular boundary points.

b) The result of Ph. Clément and C.A. Timmermans (see [CT])

The problem these two authors solved has a little different nature since their interest concerns Ventcel boundary conditions, i.e., $\lim_{x\to 0} Af(x) = 0$ and $\lim_{x\to 1} Af(x) = 0$, and they give necessary and sufficient conditions such that the below defined operator A generates a semigroup on $C[0,1]$. Let $I = [0,1]$ and α, β be real valued continuous functions on $(0,1)$ with $\alpha > 0$ on $(0,1)$ and consider

$$Af = \alpha f'' + \beta f' \qquad \text{with domain}$$
$$D(A) = \{f \in C[0,1] \cap C^2(0,1) : \lim_{x\to 0} Af(x) = 0 \text{ and } \lim_{x\to 1} Af(x) = 0\}.$$

Then the following theorem holds.

Theorem 3.17. *A is generator of a semigroup on $C[0,1]$ if and only if 0 and 1 are not entrance boundary points.* ∎

At a first reading, one may assume a contradiction to the results of P. Mandl or even to our results (see Theorem 3.10). But in fact, if the boundary point is an exit or a regular boundary, then the Ventcel condition acts as boundary condition, and if $x = 0$ or $x = 1$ is a natural boundary, then the Ventcel condition assures that $\lim_{x\to 0} Af(x)$ and $\lim_{x\to 1} Af(x)$ exists.

In order to prove that the results coincide, let us assume the simpler case of the operator $Af = mf''$, $m \in C[0,1]$, $m > 0$ on $(0,1)$, defined on $X = C[0,1]$ with domain $D(A_m) = \{f \in C[0,1] \cap C^2(0,1) : \lim_{x\to 0} Af(x)$ and $\lim_{x\to 1} Af(x)$ exist$\}$. We consider the situation, when we do not need additional boundary conditions. For this situation Theorem 3.17 states the following.

The operator $Af = mf''$ with domain $D(A) = \{f \in C[0,1] \cap C^2(0,1) : \lim_{x\to 0} Af(x) = 0,\ \lim_{x\to 1} Af(x) = 0\}$ generates a strongly continuous positive semigroup.

For the same situation, Theorem 3.10 reads as follows.

Let $m \in C^2[0,1]$. Then the maximal operator A_m generates a positive semigroup if and only if

$$\int_0^z \frac{1}{\sqrt{m}} = \int_z^1 \frac{1}{\sqrt{m}} = \infty \qquad \text{for all} \quad z \in (0,1).$$

Now we prove that there is no contradiction between these two statements; moreover we obtain a criterion when the maximal domain coincides with a domain including Ventcel boundary conditions. The lemma below is a local statement for $x = 0$; for $x = 1$ it works analogously.

Lemma 3.18. *If $m \in C^2[0,1]$, $m(x) > 0$ for $x \in (0,1)$, $\int_0^z \frac{dx}{\sqrt{m(x)}} = \infty$ and $f \in C[0,1] \cap C^2(0,1)$, then the existence of $\lim_{x\to 0} m(x)f''(x)$ implies that this limit is zero.*

Proof. First note that $m(0) = 0$ and $m'(0) = 0$ because $\int_0^z \frac{1}{\sqrt{m}} = \infty$. For details see the proof of Lemma 3.16. Again we use Taylor's theorem and obtain $m(x) = m(0) + xm'(0) + \frac{x^2}{2}m''(\xi_x) = \frac{x^2}{2}m''(\xi_x)$, hence

$$\lim_{x\to 0} m(x)f''(x) = \lim_{x\to 0} \frac{x^2}{2} m''(\xi_x)f''(x). \tag{3.18}$$

Assume the contrary, i.e., $\lim_{x\to 0} m(x)f''(x) = 1$. Since $m \in C^2[0,1]$ and we suppose $\lim_{x\to 0} \frac{x^2}{2} m''(\xi_x)f''(x) = 1$ we have $x^2 f''(x)m''(\xi_x) \geq x^2 f''(x)c_1 \geq K$, hence

$$f''(x) \geq \frac{K}{x^2 c_1}, \qquad \text{where} \quad c_1 = \inf m''(\xi_x). \tag{3.19}$$

We choose a small neighbourhood U of $x = 0$ such that $\inf_{x\in U} m''(\xi_x) \neq 0$. This works since $m(x) = \frac{x^2}{2} m''(\xi_x)$ and $m(x) > 0$ for $x \neq 0$. (If $m''(\xi_x) = 0$, we substitute in (3.18).) We compute now $f(x)$ and obtain a contradiction to $f \in C[0,1]$.

$$\begin{aligned} f(x) &= f(1/2) - \int_x^{1/2} f'(z)dz \\ &= f(1/2) - f'(1/2)(1/2 - x) + \int_x^{1/2} \int_z^{1/2} f'' dt dz. \end{aligned}$$

We use relation (3.19) and obtain

$$\begin{aligned} f(x) &\geq f(1/2) - f'(1/2)(1/2 - x) + \frac{K}{c_1} \int_x^{1/2} \int_z^{1/2} \frac{1}{t^2} dt dz \\ &= f(1/2) - f'(1/2)(1/2 - x) + \frac{K}{c_1} \left[-1 + \ln 1/2 + 2x - \ln x\right], \end{aligned}$$

which implies that $\lim_{x\to 0} f(x) \geq \infty$ because of the logarithm. This is a contradiction to $\lim_{x\to 0} f(x)$ exists ($f \in C[0,1]$). ■

c) The result of J. Goldstein and C.Y. Lin (see [GL])

Inspired by the work of Ph. Clément and C.A. Timmermans, these two authors proved even for the nonlinear case, that under a Ventcel type boundary condition, the coefficient of f'' can be allowed to approach zero *arbitrarily* rapidly.

Let $\Phi : [0,1] \times \mathbb{R} \to \mathbb{R}$ be continuous. Suppose $\Phi(x,\xi) \geq \Phi_0(x)$ on $[0,1] \times \mathbb{R}$, where $\Phi_0 > 0$ on $(0,1)$ and $\Phi_0 \in C[0,1]$. Define an operator A on the real space $X = C[0,1]$ by

$$\begin{aligned} &Au(x) = \Phi(x, u'(x))u''(x) \quad \text{for} \quad u \in D(A) \\ &D(A) = \{u \in X \cap C^2(0,1) : \lim_{x\to 0} Au(x) = 0 \text{ and } \lim_{x\to 1} Au(x) = 0\}. \end{aligned}$$

Then, the theorem reads as follows, see [GL1].

Theorem 3.19. *The operator A is densely defined and $m-$dissipative on X.* ■

Finally let us mention that, using the same arguments as in part b), one can see that there is no contradiction between Theorem 3.19. and the results of Section 3.

References

[ABR] Arendt, W., C.J.K. Batty and D.W. Robinson, *Positive semigroups generated by elliptic operators on Lie groups,* J. Operator Theory, to appear.

[ACK] Arendt, W., P. Chernoff and T. Kato, *A generalization of dissipativity and positive semigroups*, J. Operator Theory **8** (1982), 167-180.

[Ba] Batty, C.J.K., *Derivations on the line and flows along orbits*, Pacific J. Math. **126** (1987), 209-225.

[CT] Clément, Ph., and C.A. Timmermans, *On C_0-semigroups generated by differential operators satisfying Ventcel's boundary conditions*, Indag. Math. **89** (1986), 379-387.

[De] DeLaubenfels, R., *Well-behaved derivations on* $C[0,1]$, Pacific J. Math. **115** (1984), 73-80.

[Do] Dorroh, J.R., *Contraction semigroups in a function space*, Pacific J. Math. **19** (1966), 35-38.

[Fe1] Feller, W., *The parabolic differential equations and the associated semigroups of operators*, Ann. Math. **55** (1952), 468-519.

[Fe2] Feller, W., *The general diffusion operator and positivity preserving semigroups in one dimension*, Ann. Math. **10** (1954), 417-463.

[Fe3] Feller, W., *On second order differential operators*, Ann. Math. **61** (1955), 90-105.

[Fe4] Feller, W., *Generalized second order differential operators and their lateral conditions*, Illinois J. Math. **1** (1957), 459-504.

[Go] Goldstein, J.A., "Semigroups of Linear Operators and Applications," Oxford University Press, 1985.

[GL1] Goldstein, J.A., and C.Y. Lin, *Highly degenerate parabolic bounary value problems*, Differential and Integral Equations **2** (1989), 216-227.

[GL2] Goldstein, J.A., and C.Y. Lin, *Singular nonlinear parabolic boundary value problems in one space dimension*, J. Differential Equations **68** (1987), 429-443.

[GL] Gustafson, K., and G. Lumer, *Multiplicative perturbation of semigroup generators*, Pacific J. Math. **41** (1972), 731-742.

[Hi] Hille, E., *The abstract Cauchy problem and Cauchy's problem for parabolic differential equations*, J. d'Analyse Math. **3** (1954), 81- 196.

[IK] Itô K., and H.P. McKean, "Diffusion Processes and Their Sample Path," Springer-Verlag, Berlin-Heidelberg-New York, 1965.

[Ma] Mandl, P., "Analytical Treatment of One-Dimensional Markov Processes," Springer-Verlag, Berlin-Heidelberg-New York, 1968.

[MO] Miyajima, S., and N. Okazawa, *Generators of positive C_0- semigroups*, Pacific J. Math. **125** (1986), 161-175.

[MS] Munteanu, M., and M. Schwarz, *A characterization of generators of positive translation semigroups*, Semigroup Forum **38** (1989), 223- 231.

[Na]] Nagel, R., (ed.), "One-Parameter Semigroups of Positive Operators," Lect. Notes Math. **1184**, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo 1986.

[St] Steinmüller, E., *Differentialoperatoren zweiter Ordnung als Generatoren von Operatorhalbgruppen auf* $C[0,1]$, Diplomarbeit, Tübingen, 1982.

[Ta1] Taira, K., *Semigroups and boundary value problems*, Duke Math. J. **49** (1982), 287-320.

[Ta2] Taira, K., *Semigroups and boundary value problems II*, Proc. Japan Acad. **58A** (1982), 277-280.

[Yo] Yosida, K., "Functional Analysis," Springer-Verlag, Berlin- Heidelberg-New York, 1968.

Mathematisches Institut der
Universität Tübingen
Auf der Morgenstelle 10
7400 Tübingen, Germany

Received January 23, 1991
and in final form April 18, 1991

Semigroup Forum Vol. 45 (1992) 120–128

RESEARCH ARTICLE

Extensions of semigroup valued, finitely additive measures

K.P.S. Bhaskara Rao and R.M. Shortt

Communicated by K. H. Hofmann

Abstract

Let $\mathcal{C}$ be a field of subsets of a non-empty set X and let $\mu : \mathcal{C} \to E$ be a finitely additive measure (a "charge") taking values in a commutative semigroup E. We consider the problem of extending μ to a charge $\bar{\mu} : \mathcal{P}(X) \to E$ defined on the power set $\mathcal{P}(X)$, and we say that E has the charge extension property (CEP) if such extensions always exist. Łos and Marczewski proved [4] that the semigroup of non-negative reals has CEP, and Carlson and Prikry [2] have shown that every *group* has CEP. We prove that every compact semigroup has CEP and show that CEP follows from certain completeness and distributivity conditions. Specializing to the case of lattices considered as semigroups under the operation of supremum, we characterize the class of lattices with CEP. An application to closure operators in general topology is also discussed.

0. Semigroups: preliminary ideas

By a *semigroup* we mean a set E on which is defined an associative binary operation $E \times E \to E$ and which contains an identity element for this operation. We shall deal exclusively with commutative semigroups and shall employ additive notation, writing $+$ for the operation and 0 for the identity element. Most of our new results will concern *partially ordered semigroups*, i.e. semigroups $(E, +)$ together with a partial order $\leq$ such that $a \leq b$ implies $a + c \leq b + c$. A partially ordered semigroup $(E, +, \leq)$ such that $a \geq 0$ for each $a \in E$ will be called a *positively ordered semigroup* (POS). An interesting study of such structures has been written by F. Wehrung [5], who studied their behaviour *vis-a-vis* injectivity. (Actually, he considered the case where $\leq$ is not necessarily a partial order and calls these structures partially ordered monoids, or POM's.) His work is not unrelated to ours, and we shall employ one of his techniques in §3.

Say that a subset A of a POS E is *directed upwards* [resp. *downwards*] if for any $a, b \in A$, there is some $c \in A$ with $c \geq a$ and $c \geq b$ [resp. $c \leq a$ and $c \leq b$]. Say that a POS E is *directedly complete* [resp. *conditionally directedly complete*] if every upwards directed [resp. upwards directed and bounded above] subset of E has a supremum and if every downwards directed subset has an infinum. We call a POS E *directedly distributive* if

0.1) $a + \bigwedge A = \bigwedge \{a + u : u \in A\}$ for all $a \in E$ and all downwardly directed $A \subseteq E$ such that $\bigwedge A$ exists, and

0.2) $a + \bigvee A = \bigvee \{a + u : u \in A\}$ for all $a \in E$ and all upwardly directed $A \subseteq E$ such that $\bigvee A$ exists.

One class of POS will be given special attention, *viz.* lattices $(L, \leq)$ with a least element 0 considered as semigroups under the operation of supremum

$a + b = a \vee b$. Note that such a semigroup $(L, \vee)$ is (conditionally) directedly complete if and only if L is a (conditionally) complete lattice. Notice also that condition 0.2 is always satisfied and, more importantly, that condition 0.1 is not necessarily related to distributivity for L; e.g. every finite lattice satisfies 0.1. If a lattice has a greatest element, we indicate it with the symbol 1.

The power set of a set X is denoted by $\mathcal{P}(X)$. Along with the lattices with 0, it and other Boolean algebras are considered as POS's under $\vee$.

Let $(E, +)$ be a semigroup and let $\mathcal{F}$ be a (proper) filter of subsets of an index set I. We introduce an equivalence relation $\sim$ on E^I by saying that

$$[a_\lambda : \lambda \in I] \sim [b_\lambda : \lambda \in I]$$

just in case $\{\lambda : a_\lambda = b_\lambda\}$ belongs to $\mathcal{F}$, i.e. these vectors are equal "$\mathcal{F}$-almost everywhere". Denote by $E^I/\mathcal{F}$ the corresponding set of equivalence classes. Defining

$$[a_\lambda : \lambda \in I] + [b_\lambda : \lambda \in I] = [a_\lambda + b_\lambda : \lambda \in I].$$

we see that $E^I/\mathcal{F}$ becomes a semigroup in which E is embedded. In case E is ordered, or is a POS or Boolean algebra, so too is $E^I/\mathcal{F}$. The structure $E^I/\mathcal{F}$ is called a *reduced power* of E. See [3; p.168].

1. Extensions of charges; complete distributive semigroups

Let $\mathcal{C}$ be a field (algebra) of subsets of a non-empty set X and suppose that $(E, +)$ is a (commutative) semigroup with identity 0. A function $\mu : \mathcal{C} \to E$ is an (E-valued) *charge* if $\mu(\emptyset) = 0$ and $\mu(A \cup B) = \mu(A) + \mu(B)$ for all A, B in $\mathcal{C}$ with $A \cap B = \emptyset$. We say that E has the *charge extension property* if whenever $\mu : \mathcal{C} \to E$ is a charge, then there is a charge $\bar{\mu} : \mathcal{P}(X) \to E$ such that $\bar{\mu}(C) = \mu(C)$ for all $C \in \mathcal{C}$.

Lemma 1.1. *Every Abelian group G has the charge extension property.*

Indication. See [2].

The situation for semigroups is not so simple, as the following example shows:

Example 1.2. Take E be the semigroup of non-negative rationals under ordinary addition. Let $X = [0, 1]$ and let $\mathcal{C}$ be the field generated by intervals with rational endpoints contained in X. Define $\mu : \mathcal{C} \to E$ by letting $\mu(C)$ be the Lebesgue measure of C. Then μ has no extension to $\mathcal{P}(X)$. (There is no way to define $\bar{\mu}(0, 2^{-1/2})$ as a rational number.)

As the example suggests, it may be that some sort of order-completeness is needed in order for a semigroup to allow charge extensions. In keeping with this, we develop an extension procedure based on inner and outer measure. (Compare [4], [1; p.70].)

We begin by noting that charges $\mu : \mathcal{C} \to E$ taking values in a POS $(E, +)$ have a rather special property: they are monotone, i.e. given $A \subseteq B$, sets in $\mathcal{C}$, we have $\mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$. The importance of this property can be seen by comparing $\mathcal{Q}$ (rationals), which, being a group, has the CEP, and $\mathcal{Q}^+$ (non-negative rationals), which is a POS without CEP.

In results 1.3–1.6 we let X be a non-empty set and $\mathcal{C}$ a field of subsets of X and take $\mu : \mathcal{C} \to E$ to be a c harge with values in a conditionally directedly complete and directedly distributive POS. Define inner and outer charges μ_* and μ^* on subsets $A \subseteq X$ by

$$\mu_*(A) = \bigvee\{\mu(B) : B \subseteq A,\ B \in \mathcal{C}\}$$
$$\mu^*(A) = \bigwedge\{\mu(C) : A \subseteq C,\ C \in \mathcal{C}\}$$

We note that the sets to which sup and inf are applied are directed sets. We also note that $\mu_*(A) \le \mu^*(A) \le \mu(X)$.

Lemma 1.3. *If A and B are disjoint subsets of X, then*
1) $\mu_*(A \cup B) \le \mu_*(A) + \mu^*(B) \le \mu^*(A \cup B)$. *If also $A \cup B \in \mathcal{C}$, then*
2) $\mu(A \cup B) = \mu_*(A) + \mu^*(B) = \mu^*(A) + \mu_*(B)$.

Proof. 1) Let C and D be sets in $\mathcal{C}$ with $B \subseteq C$ and $D \subseteq A \cup B$. Then $D - C \subseteq A$, so that $\mu(D - C) \le \mu_*(A)$. Then $\mu(D) \le \mu_*(A) + \mu(D \cap C) \le \mu_*(A) + \mu(C)$. Taking an infimum over sets C yields $\mu(D) \le \mu_*(A) + \mu^*(B)$. (See 0.1.) Taking the supremum over sets D yields $\mu_*(A \cup B) \le \mu_*(A) + \mu^*(B)$. (See 0.2.) Note that the infimum and supremum in question are taken over directed sets.

Now let C and D be sets in $\mathcal{C}$ with $C \subseteq A$ and $A \cup B \subseteq D$. Then $B \subseteq D - C$, so that $\mu^*(B) \le \mu(D - C)$. Then $\mu(C) + \mu^*(B) \le \mu(D)$. Taking a supremum over sets C and then an infimum over sets D yields $\mu_*(A) + \mu^*(B) \le \mu^*(A \cup B)$. (Again, we have used 0.1.)

2) This part follows immediately. ■

Lemma 1.4. *If A and B are disjoint subsets of X, then*
1) $\mu_*(A) + \mu_*(B) \le \mu_*(A \cup B)$
2) $\mu^*(A \cup B) \le \mu^*(A) + \mu^*(B)$.

Proof. 1) Let C and D be sets in $\mathcal{C}$ with $C \subseteq A$ and $D \subseteq B$. Then $\mu(C) + \mu(D) = \mu(C \cup D) \le \mu_*(A \cup B)$. Taking a supremum over all such sets C and then over set s D yields (0.2) $\mu_*(A) + \mu_*(B) \le \mu_*(A \cup B)$.

2) Let C and D be sets in $\mathcal{C}$ with $A \subseteq C$ and $B \subseteq D$. Then $\mu^*(A \cup B) \le \mu(C \cup D) \le \mu(C) + \mu(D)$. Taking infima over C and D yields $\mu^*(A \cup B) \le \mu^*(A) + \mu^*(B)$. ■

Lemma 1.5. *If A, B, C, D are subsets of X with $A \subseteq C$, $B \subseteq D$, $C \cap D = \emptyset$, and C, $D \in \mathcal{C}$, then*
1) $\mu_*(A \cup B) = \mu_*(A) + \mu_*(B)$.
2) $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

Proof. 1) From lemma 1.3, we have the inequality $\ge$. Now suppose $U \subseteq A \cup B$ for $U \in \mathcal{C}$. Then $U \cap C \subseteq A$ and $U \cap D \subseteq B$, so that $\mu(U) = \mu(U \cap C) + \mu(U \cap D) \le \mu_*(A) + \mu_*(B)$. Taking a supremum over U yields $\mu_*(A \cup B) \le \mu_*(A) + \mu_*(B)$, as required.

2) From lemma 1.3, we have the inequality $\le$. Now suppose $A \cup B \subseteq U$ for some $U \in \mathcal{C}$. Then $A \subseteq U \cap C$ and $B \subseteq U \cap D$, so that $\mu(U) \ge \mu(U \cap C) + \mu(U \cap D) \ge \mu^*(A) + \mu^*(B)$. Taking an infimum over U yields $\mu^*(A \cup B) \ge \mu^*(A) + \mu^*(B)$. ■

Proposition 1.6. *Let A be an arbitrary subset of X and let $\mathcal{F} = \mathcal{F}(\mathcal{C}, A)$ be the field generated by $\mathcal{C} \cup \{A\}$. If $\mu : \mathcal{C} \to E$ is a charge taking values in a conditionally directedly complete, directedly distributive* POS, *then the set functions μ_1 and μ_2 defined by*

$$\mu_1(C) = \mu_*(C \cap A) + \mu^*(C \cap A^c)$$
$$\mu_2(C) = \mu^*(C \cap A) + \mu_*(C \cap A^c)$$

are E-valued charges on $\mathcal{F}$ extending μ.

Demonstration. Given $C \in \mathcal{C}$, we note that $(C \cap A) \cup (C \cap A^c) = C \in \mathcal{C}$. It follows from lemma 1.5.2 that $\mu_1(C) = \mu_2(C) = \mu(C)$. So μ_1 and μ_2 are extensions of μ. By symmetry, it is enough to prove that μ_1 is a charge. Suppose that C and D are disjoint members of $\mathcal{F}$. We may write

$$C = (C_1 \cap A) \cup (C_2 \cap A^c)D = (D_1 \cap A) \cup (D_2 \cap A^c)$$

where C_1, C_2, D_1, $D_2 \in \mathcal{C}$ and $C_1 \cap D_1 = C_2 \cap D_2 = \emptyset$. Then

$$\begin{aligned}
\mu_1(C \cup D) &= \mu_*((C \cup D) \cap A) + \mu^*((C \cup D) \cap A^c) \\
&= \mu_*((C \cap A) \cup (D \cap A)) + \mu^*((C \cap A^c) \cup (D \cap A^c)) \\
&= \mu_*((C_1 \cap A) \cup (D_1 \cap A)) + \mu^*((C_2 \cap A^c) \cup (D_2 \cap A^c)) \\
\text{(lemma 1.3)} &= \mu_*(C_1 \cap A) + \mu_*(D_1 \cap A) + \mu^*(C_2 \cap A^c) + \mu^*(D_2 \cap A^c) \\
&= (\mu_*(C_1 \cap A) + \mu^*(C_2 \cap A^c)) + (\mu_*(D_1 \cap A) + \mu^*(D_2 \cap A^c)) \\
&= (\mu_*(C \cap A) + \mu^*(C \cap A^c)) + (\mu_*(D \cap A) + \mu^*(D \cap A^c)) \\
&= \mu_1(C) + \mu_1(D)
\end{aligned}$$

■

Corollary 1.7. *Every conditionally directedly complete, directedly distributive* POS *has the charge extension property.*

Indication. This follows from proposition 1.6 through a straightforward application of Zorn's Lemma.

Corollary 1.8. *Every complete Boolean algebra, considered as a semigroup under the join operation, has the charge extension property.*

Example 1.9. Let E be the semigroup of non-negative real numbers under addition. Corollary 1.7 implies that E has the charge extension property. This result dates back to [4]. It is instructive to contrast examples 1.2 and 1.9.

2. Compact semigroups

A (commutative) semigroup $(E, +)$ is a *topological* semigroup if E is a topological space with the property that addition $+$ is a continuous mapping $E \times E \to E$. Say that E is a *compact semigroup* if the E is a topological semigroup whose topology is compact. We prove that every compact semigroup has the charge extension property. In particular, this applies to all finite semigroups.

Lemma 2.1. *Let $\mathcal{F}$ be a field of subsets of a non-empty set X and let $\mathcal{C}$ be a subfield of $\mathcal{F}$. If $\mathcal{C}$ is finite, then every charge $\mu : \mathcal{C} \to E$ taking values in a semigroup E has an extension to a charge $\bar{\mu} : \mathcal{F} \to E$.*

Proof. Without loss of generality, we take $\mathcal{F} = \mathcal{P}(X)$, the power set of X. Since $\mathcal{C}$ is finite, it is generated by a partition $X = C_1 \cup \ldots \cup C_n$. For each i, choose a point $x_i \in C_i$. Define $\bar{\mu}$ on $\mathcal{F}$ by $\bar{\mu}(F) = \sum\{\mu(C_i) : x_i \in F\}$. Clearly, $\bar{\mu}$ extends μ to all of $\mathcal{F}$. ∎

Proposition 2.2. *Every compact semigroup E has the charge extension property.*

Demonstration. Let $\mu : \mathcal{C} \to E$ be an E-valued charge on a field $\mathcal{C}$ of subsets of X and suppose that $\mathcal{F} \supseteq \mathcal{C}$ is another such field. For each finite sub-field $\mathcal{F}_0$ of $\mathcal{F}$, define $D(\mathcal{F}_0)$ to be the set of all functions $\bar{\mu} : \mathcal{F} \to E$ such that 1) the restriction of $\bar{\mu}$ to $\mathcal{F}_0$ is a charge, and 2) $\bar{\mu}(A) = \mu(A)$ for sets $A \in \mathcal{F}_0 \cap \mathcal{C}$. Regarded as a subset of $E^{\mathcal{F}}$, each set $D(\mathcal{F}_0)$ is closed. Given finite sub-fields $\mathcal{F}_1, ..., \mathcal{F}_n$ of $\mathcal{F}$, we have $D(\mathcal{F}_1) \cap \ldots \cap D(\mathcal{F}_n) \supseteq D(\mathcal{F}_0)$, where $\mathcal{F}_0$ is the field generated by $\mathcal{F}_1 \cup \ldots \cup \mathcal{F}_n$. It follows from lemma 2.1 that each set $D(\mathcal{F}_0)$ is non-empty. Since the collection of all such sets has the finite intersection property and since each is closed as a subset of the compact space $E^{\mathcal{F}}$, we may find a function $\bar{\mu} : \mathcal{F} \to E$ in all of these sets. Such a $\bar{\mu}$ must be a charge extending μ. ∎

Corollary 2.3. *Every finite semigroup has the charge extension property.*

Question 2.4. *What is the connection between proposition 2.2 and corollary 1.7? Is there an order-theoretic characterization of compact semigroups?*

3. Lattices with the charge extension property

We now restrict our attention to a particular class of semigroups that are idem-multiple ($x + x = x$), *viz.* the class of lattices with 0 under the supremum operation. One result follows immediately from results in §1.

Lemma 3.1. *Let L be a lattice with 0. If L is conditionally complete and satisfies the identity*

$$a \vee \bigwedge A = \bigwedge\{a \vee u : u \in A\}$$

for $a \in A$ and $A \subseteq L$, a (downward) directed set, then the semigroup $(L, \vee)$ has the charge extension property.

Indication. This follows straightaway from corollary 1.7. In fact, one can obtain a simpler form for the extension charge. Under these hypotheses, one sees without great difficulty that the outer measure

$$\mu^*(A) = \bigwedge\{\mu(C) : A \subseteq C, C \in \mathcal{C}\}$$

is defined and is an extension of μ from $\mathcal{C}$ to the power set $\mathcal{P}(X)$.

We now set about proving the converse of lemma 3.1.

Lemma 3.2. *Let L be a lattice with 0, considered as a semigroup under $\vee$. If L has the charge extension property, then L is conditionally complete.*

Proof. Let A be a subset of L bounded above by some element u. Put $X = \{x \in L : x \leq u.\}$. Then X is a sub-lattice of L. For each $x \in X$, we define $D(x) = \{y \in X : y \leq x\}$. Let $\mathcal{C}$ be the field of subsets of X generated by the sets $D(x)$ for $x \in X$.

We assert that every set in $\mathcal{C}$ has a supremum in L. To prove this, we note that each set in $\mathcal{C}$ is a finite union of sets of the form $C = D(a_1)\cap...\cap D(a_n)\cap D(b_1)^c \cap ... \cap D(b_m)^c$, where a_i, b_i are elements of X and complementation is taken relative to X. The supremum of such a set C is $a_1 \wedge ... \wedge a_n$ or 0 according as C is non-empty or empty.

So we define $\mu : \mathcal{C} \to L$ by the rule $\mu(C) = \bigvee C$. Clearly, μ is an L-valued charge on $\mathcal{C}$. If L has the charge extension property, then μ extends to a charge $\bar{\mu} : \mathcal{P}(X) \to L$. Define $A_0 = \cup\{D(a) : a \in A\}$. We assert that $\bar{\mu}(A_0)$ is the supremum in L of the set A. Clearly, $\bar{\mu}(A_0) \geq \bar{\mu}(D(a)) = \mu(D(a)) = a$ for each $a \in A$, so that $\bar{\mu}(A_0)$ is an upper bound of A. If b is any other upper bound for A, then $A_0 \subseteq D(b)$, so that $\bar{\mu}(A_0) \leq \bar{\mu}(D(b)) = b$. Thus $\bar{\mu}(A_0) = \bigvee A$, and L is conditionally complete. ■

Now we treat distributivity. First, however, we need a technical result concerning join-embeddings of lattices in Boolean algebras.

Lemma 3.3. *Let L be a complete lattice (with 0 and 1). Then there is a complete Boolean algebra B and a one-one map $j : L \to B$ such that $j(0) = 0$ and $j(a \vee b) = j(a) \vee j(b)$ for all a, b in L. Also, j and B can be chosen so that there is a mapping $\varphi : B \to L$ such that 1) $\varphi(a \vee b) = \varphi(a) \vee \varphi(b)$ for a, $b \in B$, and 2) $\varphi(j(a)) = a$ for $a \in L$.*

Proof. Put $L_1 = L - \{1\}$. We take $B = \mathcal{P}(L_1)$. For each $a \in L$, define $C(a) = \{x \in L_1 : x \geq a\}$ and put $j(a) = C(a)^c$. (Complementation with respect to L_1.) Clearly, j is one-one, and we have $j(a \vee b) = C(a \vee b)^c = [C(a) \cap C(b)]^c = j(a) \cup j(b)$, as well as $j(0) = \emptyset$. Next, define $\varphi : B \to L$ by the rule $\varphi(M) = \inf\{a \in L : M \subseteq j(a)\}$. If $M \subseteq N$, then $\varphi(M) \leq \varphi(N)$, so that $\varphi(M \cup N) \geq \varphi(M) \vee \varphi(N)$. Noting that for $\emptyset \neq M \subseteq L$, we have $\cap\{C(a)^c : a \in M\} = C(\inf\ M)^c$, it follows that $\cap\{C(a)^c : M \subseteq j(a)\} = C(\varphi(M))^c$.

Thus $M \subseteq j(\varphi(M))$, so that the infimum defining φ is actually attained. For M, $N \subseteq L$, we have $M \cup N \subseteq j(\varphi(M)) \cup j(\varphi(N)) = j(\varphi(M) \vee \varphi(N))$, so that $\varphi(M \cup N) \leq \varphi(M) \vee \varphi(N)$. We have proved that $\varphi(M \cup N) = \varphi(M) \vee \varphi(N)$. The equation $\varphi(j(a)) = a$, $a \in L$ is trivial. ■

The following result makes use of reduced powers. In this, we follow the example of Wehrung [5; p.44].

Proposition 3.4. *Let L be a lattice with 0 and 1. If $(L, \vee)$ has the charge extension property., then L is complete, and we have the identity*

$$(*) \qquad a \vee \bigwedge A = \bigwedge \{a \vee u : u \in A\}$$

for each $a \in L$ and all (downward) directed subsets A of L.

Demonstration. Completeness for L follows from lemma 3.2. Given a directed set $A \subseteq L$ and an element a in L, put $b = \bigwedge A$ and $c = \bigwedge\{a \vee u : u \in A\}$. Obviously, $a \vee b \leq c$. Now let I be the collection of all finite subsets of A that contain a least element. Let $\mathcal{F}$ be the filter of subsets of I generated by

$\{F_p : p \in I\}$, where $F_p = \{q \in I : p \subseteq q\}$. Let $L^I/\mathcal{F}$ be the reduced power of L.

By lemma 3.3, there is a Boolean algebra B and functions $j : L \to B$ and $\varphi : B \to L$ that preserve $\vee$ and such that $\varphi(j(a)) = a$ for $a \in L$. Let $B^I/\mathcal{F}$ be the corresponding reduced power of B and let $j_* : L^I/\mathcal{F} \to B^I/\mathcal{F}$ be the mapping induced by j. We regard L and B as subsets of $L^I/\mathcal{F}$ and $B^I/\mathcal{F}$, respectively. In particular, B can be thought of as a sub-algebra (or sub-field) of $B^I/\mathcal{F}$. Since L has the charge extension property, the charge φ may be extended to a charge $\bar{\varphi} : B^I/\mathcal{F} \to L$. Define $\rho : L^I/\mathcal{F} \to L$ as $\rho = \bar{\varphi} \circ j_*$. We have the following commutative diagram.

$$\begin{tikzcd}
L^I/\mathcal{F} \arrow[rr, "j_*"] \arrow[dr, "\rho"'] & & B^I/\mathcal{F} \arrow[dl, "\bar{\varphi}"] \\
& L & \\
L \arrow[uu] \arrow[ur, "="'] \arrow[rr, "j"'] & & B \arrow[uu] \arrow[ul, "\varphi"]
\end{tikzcd}$$

Now put $d = \rho([\bigwedge p : p \in I])$. For each $p \in I$, we have $c \le \bigwedge\{a \vee u : u \in p\} = a \vee \bigwedge p$. (This last equality follows from the fact that sets in I contain a least element.) Let $\bar{c}$ and $\bar{a}$ be the images of c and a under inclusion in $L^I/\mathcal{F}$. Then $\bar{c} \le \bar{a} \vee [\bigwedge p : p \in I]$. Application of ρ to both sides yields $c \le a \vee d$. Given $u \in A$, we have $\bigwedge p \le u$ for $\mathcal{F}$-almost all p, *viz.* for p in $F_{\{u\}}$. It follows that $[\bigwedge p : p \in I] \le \bar{u}$ for each $u \in A$. Applying ρ, we see that $d \le u$ for each such u. So $c \le a \vee d \le a \vee b$, as required.

Corollary 3.5. *Let L be a lattice with 0. Then $(L, \vee)$ has the charge extension property if and only if L is conditionally complete and satisfies the identity*

$$a \vee \bigwedge A = \bigwedge\{a \vee u : u \in A\}$$

for each $a \in L$ and all downward directed subsets A of L.

We can now complete corollary 1.8.

Corollary 3.6. *Let B be a Boolean algebra. The $(B, \vee)$ has the charge extension property if and only if B is complete.*

Note 3.7. It is well-known that the injective objects in the category of Boolean algebras are characterized by completeness. It would seem at first glance that corollary 3.6 merely restates this fact. However, the charges we consider respect only the supremum operation, and so are not Boolean homomorphisms. Thus, corollary 3.6 does not imply, nor is it implied by the Boolean injectivity theorem.

We conclude this section with a few remarks concerning generalizations to semilattices. Let $(L, \le)$ be a join-semilattice, i.e. a partially ordered set such that each pair of elements x, y have a supremum $x \vee y$. If L has a least element 0, then $(L, \vee)$ is a POM. In fact, we have

Lemma 3.8. *A* POM $(E, +, \le)$ *is a join-semilattice (with $+ = \vee$) if and only if*

i) $a \le b$ if and only if there is some $c \in E$ such that $b = a + c$; (in the language of [5], the order $\le$ is minimal) and *ii)* $a + a = a$ *for all* $a \in E$. *(E is idem-multiple.)*

Proof. The necessity of the two conditions is easily seen. Assume that they hold. For x, $y \in E$, certainly $x + y \geq x$ and $x + y \geq y$. Now suppose that z is another upper bound for x and y. By minimality, there are u, $v \in E$ with $z = x + u$ and $z = y + v$. Adding these, we have $z = z + z = x + y + u + v$, so that $z \geq x + y$. We have shown $x + y = x \vee y$. ■

We do not have a characterization of exactly which join-semilattices have CEP. There is, however, this positive result:

Proposition 3.9. *Let* $(E, +)$ *be a join-semilattice with* $+ = \vee$. *Suppose that*
i) $\bigwedge A$ *exists whenever* A *is a downward directed subset of* E;
ii) $a \vee \bigwedge A = \bigwedge(a + A)$ *for each downward directed subset of* E.
Then $(E, +)$ *has the* CEP.

Indication. As in the notes for lemma 3.1, one may observe that outer measure μ^* exists and defines an extension of the E-valued charge μ to all of $\mathcal{P}(X)$.

Corollary 3.10. *Let* $(E, +)$ *be a join-semilattice that is well-founded, i.e. without infinite descending sequences. Then* $(E, +)$ *has the* CEP.

Example 3.11. Let E be the set $\{0, 1, 2, ..., x, y, z\}$ with the usual ordering of integers extended so that $x \geq n$ and $y \geq n$ for $n = 0, 1, 2, \ldots$ and $z = x \vee y$. The elements x, y are incomparable. The $(E, \vee)$ is well-founded and has CEP.

4. Lattices of open sets

In this section, we explore a particular kind of lattice, *viz.* the lattice of open subsets of a zero-dimensional topological space, and investigate the charge extension property for such lattices.

Let X be a zero-dimensional (Hausdorff) topological space and let $\tau(X)$ denote the topology of X. We consider $\tau(X)$ as a semigroup under the operation of union $\cup$. Then $c(X)$, the collection of "clopen" (closed and open) subsets of X, is a subsemigroup of $\tau(X)$, as well as being a sub-field of $\mathcal{P}(X)$, the power set of X. The inclusion map $\iota : c(X) \to \tau(X)$ is a $\tau(X)$-valued (or $c(X)$-valued) charge. We consider the problem of extending ι to a $\tau(X)$-valued charge $\Phi : \mathcal{P}(X) \to \tau(X)$.

4.1. If ι has such an extension Φ, then ι has an extension to a charge $\chi : \mathcal{P}(X) \to c(X)$.

To see this, suppose that Φ is a $\tau(X)$-valued extension of ι. If U is an open set, $U = \cup_\alpha C_\alpha$ for clopen sets C_α. Then $\Phi(U) \supseteq \Phi(C_\alpha) = C_\alpha$, so that $U \subseteq \Phi(U)$. Thus $\Phi^n(A) \subseteq \Phi^{n+1}(A)$ for $n \geq 1$ and $A \subseteq X$. (Here, Φ^n represents the n-fold composition power of Φ.) In the same way, we find that $\Phi(F) \subseteq F$ for closed sets F. Therefore, $\Phi(K) = K$ if and only if K is clopen. We define composition powers Φ^γ for each ordinal γ: put $\Phi^{\gamma+1}(A) = \Phi(\Phi^\gamma(A))$ and set $\Phi^\gamma(A) = \cup\{\Phi^\beta(A) : \beta < \gamma\}$ for γ a limit ordinal. Note that each of these powers Φ^γ is a $\tau(X)$-valued charge on $\mathcal{P}(X)$. Since the transfinite series $\Phi^\gamma(A)$ is increasing in γ, there must be some ordinal γ such that $\Phi^{\gamma+1} = \Phi^\gamma$. It follows that $\Phi^\gamma(A)$ is clopen for every A, so that $\chi = \Phi^\gamma$ is an extension of ι to a $c(X)$-valued charge.

4.2. The charge ι extends to $\tau(X)$-valued (or $c(X)$-valued) charge χ on $\mathcal{P}(X)$ if and only if X is extremally disconnected.

Suppose that $\chi : \mathcal{P}(X) \to c(X)$ is an extension of ι. (From 4.1, this is no harder than extension to a $\tau(X)$-valued charge.) Given a family of clopen sets C_γ whose union is $U = \bigcup_\gamma C_\gamma$, we assert that $\chi(U)$ is the supremum in the Boolean algebra $c(X)$ of the sets C_γ. Certainly, $\chi(U) \supseteq \chi(C_\gamma) = C_\gamma$ for each γ. If K is a clopen set containing the sets C_γ, then $U \subseteq K$ and $\chi(U) \subseteq \chi(K) = K$, as required. Since $c(X)$ is complete, it follows that X is extremally disconnected. In fact, $\chi(A) = int(\overline{A}) = \overline{int(A)}$ for every $A \subseteq X$: this is in fact the unique extension of ι to a join-homomorphism on $\mathcal{P}(X)$. Indeed, this χ is a Boolean homomorphism. The kernel of χ is the ideal of nowhere dense subsets of X. There follows

4.3. If X is extremally disconnected, then $c(X)$ is isomorphic as a Boolean algebra to the quotient of $\mathcal{P}(X)$ by the ideal of nowhere dense sets.

4.4. A Hausdorff topology $\tau(X)$ has the charge extension property if and only if X is discrete.

This follows from corollary 3.5 and an elementary argument.

4.5. A Boolean algebra B has the charge extension property if and only if B is complete.

Of course, this is corollary 3.6. But here is another proof of "only if": Suppose B has the charge extension property. We use Stone's Theorem to represent B as $c(X)$, where X is the Stone space of B. Extend the identity map $\iota : c(X) \to c(X)$ to a charge $\chi : \mathcal{P}(X) \to c(X)$. Then 4.2 implies that X is extremally disconnected, so that B is complete.

References

[1] Bhaskara Rao, K.P.S. and Bhaskara Rao, M., "Theory of Charges", Academic Press, London–New York, 1983.

[2] Carlson, T. and Prikry, K., *Ranges of Signed Measures*, Periodica Math. Hungarica **13** (1982), 151–155.

[3] Chang, C.C., and Keisler, H.J., "Model Theory", North Holland/Elsevier, Amsterdam–New York, 1973.

[4] Łos, J. and Marczewski, E., *Extension of measures*, Fund. Math **36** (1949), 267–276.

[5] Wehrung, F., *Embedding into injective ordered monoids*, Rapport de recherche, Université de Caen **(4)** 1990.

Indian Statistical Institute
Bangalore Centre
Statistics and Mathematics Unit
8th Mile, Mysore Road
R.V. College Post
Bangalore 560 059
India

Department of Mathematics
Wesleyan University
Middletown, CT, 60459

Received May 10, 1991
and in final form January 19, 1992

Semigroup Forum Vol. 45 (1992) 129–134

BOOK REVIEW

The Analytical and Topological Theory of Semigroups

Book Review

Dennison R. Brown

Communicated by Boris M. Schein

1. Introduction

In January, 1989, a large portion of that part of the semigroup community not exclusively involved with the algebraic theory of the subject met in Oberwolfach to update one another on the current state of the art. As it was to be the first international tete-a-tete in quite some time, the organizers (who are also the editors)—Karl Hofmann, Jimmie Lawson, and John Pym—opted to eschew what had been the more or less usual format of such get togethers, that is the arrangement in which everyone talks and everyone publishes in the proceedings. In place of that, they featured sixteen invited addresses, each of one hour's duration. These talks were to cover various emerging topics, with updates stretching back at least to the previous Oberwolfach conference in 1981. The written versions of these remarks were then amassed and published as Volume One of De Gruyter's handsome new series, *Expositions in Mathematics* [4]. This work is thus not a standard conference proceedings, as the editors point out in their preface, since at least as many talks of shorter length were given. Moreover, several authors were unable to attend the conference. What it is, however, is a compilation of excellent articles entailing what has become of our subject during the 1980's. Reviewer cannot recall another anthology in the area in which every article has been so carefully prepared, in which the subjects treated have been so well laid out to lead the interested reader back to where things were ten years (or more) ago, then to bring one up to date. As a mirror of the health of the subject at present, then, this volume does an excellent job. What it indicates for the future of our endeavors is somewhat less clear—more on that subject in reviewer's closing remarks.

The book begins with a short preface, and then proceeds, divided more or less artificially, into four sections of four articles each. Indeed, anyone who could truly structure a volume of papers in our area could probably balance the US budget—we are a race of rugged individualists. After the *de rigeur* nod to the influence of A. D. Wallace on our subject, the introduction alerts the reader not to expect any papers on the general theory of topological semigroups. Humorously, the only other appearence of ADW within these bounds is a misattribution of one of Numakura's results to him. The preface goes on to regret that the subject has diversified to the degree that we can no longer meet profitably with the "algebraic semigroupers." This is probably true; however, one of these papers (Magill's) is, basically, about the algebraic theory, while another (Renner's) contains results both in that direction and in the direction of algebraic geometry. It is true that the Zariski topology is present in Renner's work, and that Magill uses topology frequently in his study of the semigroups of continuous selfmaps of topological spaces; presumably this squeezes them under the umbrella of the title. Nevertheless, in the reviewers opinion, conclusions are what should determine the description of such work. Indeed, the only valid (and mild) criticism of this book is that the title is perhaps somewhat misleading. In addition to Renner's paper, one notes three papers in geometry (the first three, all related at least somewhat to either Lie Theory, Lie Groups, or Lie Representations), four essentially in semitopological semigroups, and one in semilattices from the viewpoint of continuous lattices. The remaining six papers consist of four in functional analysis, one in geometric control theory, and one in probability. Given that much of semitopological semigroup theory is based on the theory of weakly almost periodic functions, and that Lie Representations require a solid background in classical analysis, and one sees that ten of the sixteen presentations will make you wish you had paid closer attention in your analysis classes. Topology is certainly present (how can one do analysis without it ?), but perhaps it should be spelled with a lower case **t** in the title. Reviewer feels that *geometric* semigroups might be a proper term for the study of the Lie Theory of semigroups, but up to now appears to be a cult of one on this subject.

While almost every article herein is a mini-goldmine in its own area, reviewer wishes to point out two results which, alone, might make the reader wish to add this to his library. One is the proof by Jimmie Lawson that a cancellative semigroup on a manifold is globally embeddable in a Lie group if (and, of course, only if) the semigroup is algebraically embeddable in a group. This proof, together with an accompaning example of such an object that is not group embeddable, reviewer finds quite remarkable. The second is the celebrated proof by Neil Hindman of the Van der Waerden Theorem on the existence of arithmetic progessions of arbitrary length in some member of any finite partition of the natural numbers. This argument requires a modicum of the theory of one-sided continuous semigroups, and the relationship of this theory to the algebraic structue of $\beta\mathbb{N}$, the Stone-Čech compactification of the natural numbers. Once these are understood, the proof is swift and elegant.

We pass now to the individual articles. As most of these have been reviewed, the remarks will be brief; in cases in which the reviewer is unlearned, even shorter.

2. The Review

The opening article, *Lie groups and semigroups* by Hofmann, is both an excellent primer on the Lie theory of semigroups and an updating of the subject since the completion of a giant monograph on it by Hilgert, Lawson, and himself [3]. There is, as always in an article by this author, an interesting history lesson, as well as a gracious acknowledgement of the work of the late Charles Loewner. From the outset of his efforts in this direction, the author has involved the Lie algebra in his studies of subsemigroups of Lie groups. Because these semigroups have the annoying habit of refusing to behave particularly well with regard to either the Lie multiplication or the Campbell-Hausdorff multiplication in the Lie algebra, it is necessary to nurse three theories along at the same time: the infinitesimal, local, and global. At the global level, there is not yet a version of Lie's fundamental theorem to relate when a given Lie wedge is the tangent wedge of some topological semigroup. The author discloses places in which the theory remains incomplete, and lists five problems and several heavyweight conjectures for the reader to mull. Whether one is versed in Lie theory or not, this article is an invaluable aid in enabling the reader to make sense of what is going on in the field.

In the second article, *Applications of Lie semigroups in analysis* , Joachim Hilgert illustrates the occurrence of the Lie theory of semigroups in classical analysis. While nothing is proved within this paper, there are several hard calculations indicated (author is a calculator non pareil), as well as an actual isomorphism between Howe's oscillator semigroup and the metaplectic semigroup of Brunet. But this is only the beginning. The author goes on to explain (well) G.I. Ol'shanskii's analytic continuation procedure and its use in creating a semigroup on a manifold which has a pre-chosen Lie group G as its Shilov boundary, in such a way that a pre-fixed infinite dimensional unitary representation of G will extend to a representation of this semigroup into the contractive operators. He continues with a quick look at present work in the direction of generalizing the Laplace transform to a class of semigroups broader than the positive real numbers, one involving Jacobi functions of the second kind, and another involving Volterra kernels. The article closes with an observation indicating that a certain infinite dimensional invariant cone plays a role in the proof of a theorem by Paneitz about the stability of certain causal differential equations.

The third article, *Embedding semigroups into Lie groups* by Lawson, has already been mentioned in the introduction. The author accomplishes the forementioned result by use of the free topological group on a topological semigroup, together with the local quotient technique of Raymond Houston and a lot of careful, clever work. There is a great deal more than that here; the analyticity of multiplication of a cancellative semigroup on a finite dimensional manifold is extended to those modeled on a Banach space, via a double sheaf covering of the semigroup and a Lie group locally isomorphic to a local quotient of the semigroup. This technique, due to Hofmann and Wolfgang Weiss jointly, introduces covering spaces into the investigation, and the theory therein is used extensively for applications, both global and local, involving the Lie theory of semigroups. This paper, at least the front of it, is possibly the most accessible in the book to someone with primary interest in what one might now call classical topological semigroups.

The final article in the geometric section of the book is *Algebraic varieties and semigroups* by Lex Renner. *Caveat lector* at this point. An algebraic semigroup is no longer one in which one is unconcerned with topological considerations, but rather an affine algebraic variety over an algebraically closed field k with an associative multiplication. Such an object is an algebraic monoid if it possesses an identity, and irreducible if it is closed in the Zariski topology and does not decompose into smaller Z-closed sets. Every algebraic monoid contains an irreducible one, namely the irreducible closed set containing the identity. In this case, the group of units is Z-dense. One of the basic results in this area, due to M.S. Putcha, states that every algebraic semigroup is (isomorphic to) a subsemigroup of $M_n(k)$ for some n, obviously under matrix multiplication. Moreover, such a subsemigroup may be taken, itself, to be closed in the Zariski topology. If k is the field of complex numbers, then clearly there should be some interesting things to study from the viewpoint of topological semigroups. However, the thrust of this work is primarily to educate the reader in the work of the author and of Putcha concerning regular irreducible algebraic monoids. That's semigroup regular, as in $x \in xSx$ for all $x \in S$. If S has a zero, then a well known result of both the author and of Putcha is that S is regular if and only if its group of units G is reductive. This allows the use of the Bruhat double coset decomposition of G. The author is able to extend this decomposition to S by replacing the Weyl group by R, the normalizer in S of a maximal torus in G. It turns out that R is a finite inverse semigroup with identity. The author can count the elements therein, as well as some important subsemigroups of R. In so doing, the number e^{-1} appears! Problem: find another book on topological semigroups where e^{-1} appears. The paper closes with a section on related developments, including the statement of an interesting result on the cohomology of compactifications of semisimple groups, due to DeConcini and Procesi.

Michael Mislove's paper on *Compact semilattices, partial orders and topology* opens the second section. This is another proofless offering; however, the author discusses the missing proofs at some length, and the paper is the better for it. The paper begins with the recent history of compact semilattices, going back to Lawson's extension of Nachbin's theorem. It proceeds to his proof of the intrinsic quality of the topology on a compact semilattice. This sets the stage for the definition of the Scott topology, and the reader is soon galloping off in the direction of the Compendium [2]. There is method in this categorical madness, however. The author discusses no fewer than nine different duality theorems between order theoretic objects and sets with various structures (sometimes the structure on the latter is, itself, order theoretic). These include the classical dualities of Lindenbaum-Tarski and of Stone. Other relationships among these categories are also indicated, in an effort to show the reader how everything fits together. The reader will get to see what Heyting algebras and sober spaces are without wading through lots and lots of pages in the official treatise on the subject. Reviewer considers this to be a very worthwhile effort. The paper closes with some nice applications, focused around the problem of when an entity contains a meaningful copy of the Cantor (semi)lattice 2^{ω}. These include harmonic analysis, theoretical computer science, and partially ordered sets in various forms.

At this juncture, the book leaves the domain of topological semigroups (actually, it was deserted somewhere in the middle of Mislove's article), not to return until the applications section. Reviewer, torn between his natural temptation to show off and the possibility that, out of his realm, he may completely miss the point of some article and wind up looking silly, has chosen the safer option. However, let it be written that, in every single one of the articles—and they were all read rather carefully—a great amount of attention has been given to making these surveys worthwhile for the non-expert. There are copious references provided, and numerous unsolved problems sprinkled throughout.

Wolfgang Ruppert, *Compact semitopological semigroups*, actually convinced the reviewer that the weakly almost periodic functions are understandable, a level that he has never been fooled up to before. J.-P. Troallic, *Semigroupes affines semitopologiques compacts*, gives a convincing demonstration that separately continuous affine semigroups are just about as nice as jointly continuous ones. John Pym, *Compact semigroups with one-sided continuity*, makes a valiant effort to straighten out the left-right confusion that abounds when continuity requirements are weakened far enough. There are excellent examples given to illustrate how badly things can go when one wanders into the Stone-Čech compactification of a discrete semigroup. In addition, he has added the non-word *oid* to reviewer's vocabulary.

In the section on functional analysis, John Baker, *Measure algebras on semigroups*, has given an extremely complete survey of the area, perhaps the most scholarly work in the book. He has included a section on invariant measures on semigroups; one hopes this will forestall for a while the next redundancy about invariant measures on compact topological semigroups. Christian Berg, *Positive definite and related functions on semigroups*, surveys the area of the title and moment functions as well, over commutative semigroups S with involution. He closes with remarks about negative definite functions and an intriguing result relating 2-divisibility in S to the size of the space of negative definite functions on S. In *Convolution semigroups and potential functions on a commutative hypergroup*, by Herbert Heyer, a hypergroup is a locally compact space K with a presumably non-trivial involution such that its space of bounded measures admits a convolution turning it into an involutive Banach algebra satisfying several conditions on its involution which won't be gone into here. In the commutative case (in the B-algebra), the theory of Gelfand pairs avails, and one can do harmonic analysis therein. In particular, Haar measure on K in the sense of an induced translation operator on K comes into play. The author provides many examples of commutative hypergroups, most of which have interesting names. In both of the prior two papers, semicharacters play a large role, as one would expect. The final paper in this section is by Anthony Lau, *Amenability of semigroups.* In addition to its mathematical worthiness, it is noteworthy on other two accounts. First, it includes a theorem with nineteen different equivalences to left amenability. This may not be a record, but it must be close to one. Second, there are no fewer than thirty open problems listed. There is also a reference list of 106 entries, befitting a subject that began in 1904.

The final section includes excellent application articles by Ivan Kupka, *Geometric control theory*, Neil Hindman, $\beta\mathbb{N}$ *and its uses*, R.W. Darling and A. Mukherjea, *Probability measures on non-negative matrix semigroups*, and Kenneth Magill, *Congruences on the semigroup of continuous selfmaps of a topological space* The first of these seems to the reviewer to offer the best opportunity for semigroups to impact (truly) applied mathematics. Reviewer has already discussed Hindman's work. The Darling-Mukherjea paper shows that there is more to be learned in the semigroups of stochastic and sub-stochastic matrices. Finally, Ken Magill's work within this huge semigroup continues to amaze the reviewer. This paper is a very tidy sample of what he has been doing for lo, these many years.

3. Closing Remarks

It remains only to amplify my earlier remarks, if I may be permitted to slip into the first person at this stage. While this has nothing to do with the book under scrutiny, I don't know of a better venue. Never have I seen a subject change so rapidly nor as drastically as topological semigroup theory since 1980. Even though a decent pair of introductory books in the area [1] came into being during the last decade, it would now appear that they are designed to prepare a student in an inactive area of the subject. While I can certainly educate my own students to do what I wish them to do (i.e., this is not a whine), it does seem to me that we need a new introductory treatment that is more oriented in the current directions. Lacking that, it is predictable that the next generation of students may have some problems conversing with one another. Fractionalization is healthy up to a point; however, if we consider ourselves bound together only by the associative law, the subject could gradually disappear into the morass of a dozen other fields represented within this very volume. Well, enough editorializing. I enjoyed the book—and the conference.

-D.R. BROWN-

References

[1] Carruth, J. H., J. A. Hildebrant, and R. J. Koch. "The Theory of Topological Semigroups," Vols. I and II, Marcel Dekker, New York, 1983 and 1986.

[2] Gierz, G., K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, and D. S. Scott. "A Compendium of Continuous Lattices," Springer-Verlag, Berlin, etc., 1980.

[3] Hilgert, J., K. H. Hofmann, and J. D. Lawson. "Lie Groups, Convex Cones, and Semigroups," Oxford University Press, 1989, xxxvii+645pp.

[4] Hofmann, K. H., J. D. Lawson, and J. S. Pym, Eds.. "The Analytic and Topological Theory of Semigroups. Trends and Developments," Walter de Gruyter, Berlin, New York, 1990 [Expositions in Mathematics, Volume I], xi+398pp, ISBN 3-11-012489-0.

Department of Mathematics
University of Houston
Houston, Texas 77204-3476

Received March 3, 1992

Semigroup Forum Vol. 45 (1992) 135–147

RESEARCH ARTICLE

On the Cauchy Problem for a Second Order Semilinear Parabolic Equation with Factored Linear Part

Angelo Favini and Alberto Venni

Communicated by J. A. Goldstein

1. Introduction

In this paper we study the Cauchy problem for a second-order semilinear equation on a bounded interval $[0,T]$. We assume that the linear part of the equation is given by the product of the linear differential operators $\left(\frac{d}{dt} - A_1\right)$ and $\left(\frac{d}{dt} - A_0\right)$ where A_0 and A_1 are (not necessarily commuting) infinitesimal generators of analytic semigroups in some Banach space.

The solutions u of the problem that we look for are strict (this means that the differential equation is satisfied up to 0); moreover we require that $\frac{d}{dt}(u' - A_0u)$ and $A_1(u' - A_0u)$ are Hölder continuous. Therefore we have to assume that the initial data belong to suitable interpolation spaces and that the nonlinear term of the equation has a convenient regularity.

In §2 we formulate precisely the problem under consideration, introducing some notations, and we state the main results. §3 contains a preliminary lemma on trace spaces that we need in order to reduce the Cauchy data of the problem to 0. §4 is devoted to the proof of the theorem.

In §5 we give some examples of general second order equations that can be reduced to equations of factored type. In particular we show that many well-known partial differential equations appearing in concrete applications can be handled by our method in the context of L^p or continuous function spaces.

We refer to [2] and [8] as the most recent papers on the subject, where further literature can be found too. We note that in both papers the operators corresponding to our A_0 and A_1 are infinitesimal generators of C_0-semigroups, and correspondingly mild solutions of the problem are investigated; moreover, the nonlinear term does not depend on the derivative of the solution. We also quote the paper [6], where the well-posedness is studied of the Cauchy problem for the complete second-order equation, in a situation to which our Example 6 is, formally, very similar (here again, however, we deal with analytic semigroups). Finally recall the results of the book [10] (see pp. 190–192), where a second-order semilinear equation is studied in a Hilbert space by means of a factorization of the linear term.

We also quote the paper [5] where mild solutions of the problem

$$\begin{cases} u''(t) + Au'(t) = f(t, u(t), u'(t)) \\ u(0) = u_0 \\ u'(0) = u_1 \end{cases}$$

are studied, where $-A$ is the generator of an analytic semigroup and f acts from $[0,T] \times \mathcal{D}(A) \times \mathcal{D}((-A)^\alpha)$ to X, with $0 < \alpha < 1$. In our Example 9 we show how to deal with a related equation in which the linear term is complete and f acts from $[0,T] \times \mathcal{D}(B) \times X$ to X, where B is a suitable closed operator such that $\mathcal{D}(B)$ is embedded in $\mathcal{D}(A)$.

2. Formulation of the problem

Let $\mathfrak{X}$ be a Banach space and A_j $(j = 0,1)$ infinitesimal generators of analytic semigroups in $\mathfrak{X}$, $t \mapsto \exp(tA_j)$. We do not suppose that the operators A_j are densely defined, so the semigroups are not necessarily strongly continuous at 0. We fix a positive real number T and consider a function $f : [0,t] \times \mathcal{D}(A_0) \times \mathfrak{X} \to \mathfrak{X}$. We shall study the following Cauchy problem on some interval $[0,T_1] \subseteq [0,T]$:

$$(1)\qquad \begin{cases} \left(\dfrac{d}{dt} - A_1\right)\left(\dfrac{d}{dt} - A_0\right) u(t) = f(t,u(t),u'(t)) \\ u(0) = u_0 \\ u'(0) = u_1 \ . \end{cases}$$

We shall make use of several Banach spaces, besides $\mathfrak{X}$, which we list below.

- When $\mathfrak{D}$, $\mathfrak{B}$ are Banach spaces, $\mathcal{L}(\mathfrak{D},\mathfrak{B})$ is the Banach space of the bounded linear operators from $\mathfrak{D}$ to $\mathfrak{B}$, and $\mathcal{L}(\mathfrak{D}) = \mathcal{L}(\mathfrak{D},\mathfrak{D})$.
- When $\mathfrak{D}$, $\mathfrak{B}$ are Banach spaces and $A : \mathcal{D}(A) \to \mathfrak{B}$ is a linear operator whose domain is a linear subspace of $\mathfrak{D}$ and whose graph is closed in $\mathfrak{D}\times\mathfrak{B}$, then $\mathcal{D}(A)$ is endowed with the graph norm, i.e. $\|x\|_{\mathcal{D}(A)} = \|x\|_{\mathfrak{D}} + \|Ax\|_{\mathfrak{B}}$; in particular this holds when $\mathfrak{D} = \mathfrak{B} = \mathfrak{X}$ and A is either A_0 or A_1.
- If $\mathfrak{D}$ is any Banach space, then $C([0,T];\mathfrak{D})$, $C^1([0,T];\mathfrak{D})$, $C^\theta([0,T];\mathfrak{D})$, $C^{1,\theta}([0,T];\mathfrak{D})$ are the spaces of the $\mathfrak{D}$-valued functions which are (respectively) continuous, continuously differentiable, Hölder continuous with exponent $\theta \in]0,1[$, differentiable with derivative in $C^\theta([0,T];\mathfrak{D})$, in each case with the usual norm; the subscript 0 will denote the closed subspace of the functions vanishing at 0. To be precise, we agree that the "usual norms" are the following:

$$\|u\|_{C([0,T];\mathfrak{D})} = \sup_{0\le t\le T} \|u(t)\|_{\mathfrak{D}}$$

$$\|u\|_{C^1([0,T];\mathfrak{D})} = \|u\|_{C([0,T];\mathfrak{D})} + \|u'\|_{C([0,T];\mathfrak{D})}$$

$$\|u\|_{C^\theta([0,T];\mathfrak{D})} = \|u\|_{C([0,T];\mathfrak{D})} + \sup_{t,s} |t-s|^{-\theta}\|u(t)-u(s)\|_{\mathfrak{D}}$$

$$\|u\|_{C^{1,\theta}([0,T];\mathfrak{D})} = \|u\|_{C([0,T];\mathfrak{D})} + \|u'\|_{C^\theta([0,T];\mathfrak{D})}$$

- When $\mathfrak{D}$ is a Banach space and A is the infinitesimal generator of either a C_0-semigroup or an analytic semigroup (not necessarily strongly continuous at 0), $(\mathfrak{D},\mathcal{D}(A))_{\theta,\infty} (\theta \in]0,1[)$ is the usual real interpolation space; having fixed $T > 0$, it can be characterized as the space of those $x \in \mathfrak{D}$ for which $\sup\limits_{0<t\le T} \frac{\|\exp(tA)x - x\|_{\mathfrak{D}}}{t^\theta} < +\infty$, and we can set $\|x\|_{(\mathfrak{D},\mathcal{D}(A))_{\theta,\infty}} = \|x\|_{\mathfrak{D}} + \sup\limits_{0<t\le T} \frac{\|\exp(tA)x - x\|_{\mathfrak{D}}}{t^\theta}$ (see [11], §1.13 for the case of a C_0-semigroup; the fact that this characterization remains true even when A generates an analytic semigroup which is not strongly continous at 0 can be easily deduced from Proposition 1.3 of [9]). In any case, since $\|\exp(tA)\|_{\mathcal{L}(\mathfrak{D})}$ is bounded on $[0,T]$ and for $0 \le s \le t \le T$, $\exp(tA)x - \exp(sA)x = \exp(sA)(\exp((t-s)A)x - x)$, we get that $x \in (\mathfrak{D},\mathcal{D}(A))_{\theta,\infty}$ if and only if the function $t \mapsto g(t) = \exp(tA)x$ belongs to $C^\theta([0,T];\mathfrak{D})$; moreover, $\|x\|_{(\mathfrak{D},\mathcal{D}(A))_{\theta,\infty}} \le \|g\|_{C^\theta(0,T];\mathfrak{D})} \le M\,\|x\|_{(\mathfrak{D},\mathcal{D}(A))_{\theta,\infty}}$, with $M = \sup\limits_{0<t\le T} \|\exp(tA)\|_{\mathcal{L}(\mathfrak{D})}$.

Concerning problem (1), we are interested in "strict" solutions, i.e. functions $u \in C([0,T_1];\mathcal{D}(A_0)) \cap C^1([0,T_1];\mathfrak{X})$ such that:

$$\begin{cases} u' - A_0 u \in C([0,T_1];\mathcal{D}(A_1)) \cap C^1([0,T_1];\mathfrak{X}) \\ \text{the differential equation is satisfied at every } \ t \in [0,T_1] \\ \text{the initial condition holds.} \end{cases}$$

Then it is obvious that a necessary condition for the existence of such a solution is that

$$u_0 \in \mathcal{D}(A_0), u_1 - A_0 u_0 \mathcal{D}(A_1)\ . \tag{2}$$

Actually, we look for solutions $u \in C^\theta([0,T_1];\mathcal{D}(A_0)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})$, with $u' - A_0 u \in C^\theta([0,T_1];\mathcal{D}(A_1)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})$ for some $\theta \in]0,1[$. To this purpose, we shall prove the following

Theorem 1. *Let $\mathfrak{X}$, A_0, A_1, f be as above. Let us fix $\theta \in]0,1[$ and suppose that*

(a) $u_0 \in \mathcal{D}(A_0)$, $u_1 - A_0 u_0 \in \mathcal{D}(A_1)$

(b) $u_1 \in (\mathfrak{D},\mathcal{D}(A_0))_{\theta,\infty}$, $A_1(u_1 - A_0 u_0) + f(0,u_0,u_1) \in (\mathfrak{D},\mathcal{D}(A_1))_{\theta,\infty}$

(c) $\forall\ (t,u,v) \in [0,T] \times \mathcal{D}(A_0) \times \mathfrak{X}$ *there exist the "partial derivatives"* $\partial_2 f(t,u,v) \in \mathcal{L}(\mathcal{D}(A_0),\mathfrak{X})$ *and* $\partial_3 f(t,u,v) \in \mathcal{L}(\mathfrak{X})$

(d) *there is an increasing function* $k : \mathbb{R}^+ \to \mathbb{R}^+$ *such that for arbitrary* $s,t \in [0,T]$, $u,w \in \mathcal{D}(A_0)$, $v,z \in \mathfrak{X}$, *with* $\max\{\|u\|_{\mathcal{D}(A_0)}, \|w\|_{\mathcal{D}(A_0)}, \|v\|_{\mathfrak{X}}, \|z\|_{\mathfrak{X}}\} \le r$ *we have:*

  (d_1) $\|f(s,u,v) - f(t,w,z)\|_{\mathfrak{X}} \le k(r)(|s-t|^\theta + \|u-w\|_{\mathcal{D}(A_0)} + \|v-z\|_{\mathfrak{X}})$

  (d_2) $\|\partial_2 f(s,u,v) - \partial_2 f(t,w,z)\|_{\mathcal{L}(\mathcal{D}(A_0))} \le k(r)(|s-t|^\theta + \|u-w\|_{\mathcal{D}(A_0)} + \|v-z\|_{\mathfrak{X}})$

  (d_3) $\|\partial_3 f(s,u,v) - \partial_3 f(t,w,z)\|_{\mathcal{L}(\mathfrak{X})} \le k(r)(|s-t|^\theta + \|u-w\|_{\mathcal{D}(A_0)} + \|v-z\|_{\mathfrak{X}})$.

Then $\exists\ r_0 > 0$ such that $\forall\ r \ge r_0\ \exists\ T_1 = T_1(r) \in]0,T]$ with the following property: the Cauchy problem (1) has a unique strict solution u on $[0,T_1]$ such that

$$u \in C^\theta([0,T_1];\mathcal{D}(A_0)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})\ , \tag{3}$$

$$u' - A_0 u \in C^\theta([0,T_1];\mathcal{D}(A_1)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})\ , \tag{4}$$

$$\|u\|_{C^\theta([0,T_1];\mathcal{D}(A_0))} \le r, \|u' - A_0 u\|_{C^\theta([0,T_1];\mathfrak{X})} \le r\ . \tag{5}$$

Moreover, such solution u belongs to $C^{2,\theta}([\varepsilon,T_1];\mathfrak{X})\ \forall\, \varepsilon \in]0,T_1[$. Finally, if $u_1 \in \mathcal{D}(A_0)$ and $A_0 u_1 + A_1(u_1 - A_0 u_0) + f(0,u_0,u_1) \in \overline{\mathcal{D}(A_0)}$, then $u \in C^2([0,T_1];\mathfrak{X})$.

3. Some preliminary remarks

We begin by recalling the following result (see [9], Theorem 4.5).

Lemma 2. *Let A be the infinitesimal generator of an analytic semigroup in the Banach space $\mathfrak{X}$ and $g \in C^\theta([0,T];\mathfrak{X})$ (with $0 < \theta < 1$). Let u be the mild solution on $[0,T]$ of the Cauchy problem*

$$\begin{cases} u'(t) = Au(t) + g(t) \\ u(0) = x_0 \end{cases} \tag{6}$$

(i.e. $u(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)g(s)\,ds$*). If* $x_0 \in \mathcal{D}(A)$ *and* $Ax_0 + g(0) \in (\mathfrak{X}, \mathcal{D}(A))_{\theta,\infty}$*, then* u *is a strict solution,* $u \in C^\theta([0,T];\mathcal{D}(A)) \cap C^{1,\theta}([0,T];\mathfrak{X})$ *and* $\|u\|_{C^\theta([0,T];\mathcal{D}(A))} + \|u\|_{C^{1,\theta}([0,T];\mathfrak{X})} \le M(\|x_0\|_{\mathfrak{X}} + \|Ax_0 + g(0)\|_{(\mathfrak{X},\mathcal{D}(A))_{\theta,\infty}} + \|g\|_{C^\theta([0,T];\mathfrak{X})})$ *for a suitable* $M > 0$*, independent of* g*.*

Hence we deduce the following

Corollary 3. *If* A *is the infinitesimal generator of an analytic semigroup in the Banach space* $\mathfrak{X}$*, then* $\phi \mapsto (\phi(0), \phi'(0))$ *is a bounded linear operator from* $C^\theta([0,T];\mathcal{D}(A)) \cap C^{1,\theta}([0,T];\mathfrak{X})$ *onto* $\mathcal{D}(A) \times (\mathfrak{X}, \mathcal{D}(A))_{\theta,\infty}$ *which has a continuous right inverse.*

Proof. The boundedness is obviously understood with respect to the intersection norm: $\|\phi\| = \|\phi\|_{C^\theta([0,T];\mathcal{D}(A))} + \|\phi\|_{C^{1,\theta}([0,T];\mathfrak{X})}$. Suppose that $\phi \in C^\theta([0,T];\mathcal{D}(A)) \cap C^{1,\theta}([0,T];\mathfrak{X})$. Then $\phi(0) \in \mathcal{D}(A)$ and $\|\phi(0)\|_{\mathcal{D}(A)} \le \|\phi\|_{C^\theta([0,T];\mathcal{D}(A))}$. Now we set $u(t) = \phi(t) - \int_0^t \exp(sA)\phi'(0)\,ds$. As $\phi'(0) \in \overline{\mathcal{D}(A)}$ from Proposition 1.2 of [9] we get $u \in C^1([0,T];\mathfrak{X})$ and $u(t) \in \mathcal{D}(A)$ $\forall t \in [0,T]$, with $u(0) = \phi(0)$ and $u'(t) - Au(t) = \phi'(t) - A\phi(t) - \phi'(0)$. Now the function $t \mapsto g(t) = \phi'(t) - A\phi(t) - \phi(0)$ belongs to $C^\theta([0,T];\mathfrak{X})$ and $A\phi(0) + g(0) = 0$. Hence, from Lemma 2, it follows that $u' \in C^\theta([0,T];\mathfrak{X})$, with $\|u'\|_{C^\theta([0,T];\mathfrak{X})} \le M(\|\phi(0)\|_{\mathfrak{X}} + \|\phi' - A\phi - \phi'(0)\|_{C^\theta([0,T];\mathfrak{X})}) \le M_1(\|\phi\|_{C^\theta([0,T];\mathcal{D}(A))} + \|\phi\|_{C^{1,\theta}([0,T];\mathfrak{X})})$. Therefore the function $t \mapsto \exp(tA)\phi'(0) = \phi'(t) - u'(t)$ belongs to $C^\theta([0,T];\mathfrak{X})$, so that $\phi'(0) \in (\mathfrak{X}, \mathcal{D}(A))_{\theta,\infty}$ and $\|\phi'(0)\|_{\mathfrak{X},\mathcal{D}(A))_{\theta,\infty}} \le \|\phi'\|_{C^\theta([0,T];\mathfrak{X})} + \|u'\|_{C^\theta([0,T];\mathfrak{X})} \le M_2(\|\phi\|_{C^\theta([0,T];\mathcal{D}(A))} + \|\phi\|_{C^{1,\theta}([0,T];\mathfrak{X})})$.

Conversely, suppose that $(x,y) \in \mathcal{D}(A) \times (\mathfrak{X}, \mathcal{D}(A))_{\theta,\infty}$, and consider the function $t \mapsto \psi(t) = x + \int_0^t \exp(sA)y\,ds$ (where the integral is meant in the norm of $\mathfrak{X}$). Since $y \in \overline{\mathcal{D}(A)}$, we have that $\psi \in C^1([0,T];\mathfrak{X})$, with $\psi'(t) = \exp(tA)y$, so that $\psi \in C^{1,\theta}([0,T];\mathfrak{X})$ and $\psi(0) = x$, $\psi'(0) = y$. On the other hand, since $x \in \mathcal{D}(A)$, we have that $\psi(t) \in \mathcal{D}(A)$ $\forall t$ and $A\psi(t) = Ax + \exp(tA)y - y$, which is again a Hölder continuous function with exponent θ. Hence the linear operator $(x,y) \mapsto \psi$ is a right inverse of $\phi \mapsto (\phi(0) < \phi'(0))$. In order to prove that this right inverse is bounded, we have to estimate $\|\psi\|_{C([0,T];\mathfrak{X})}$, $\|\psi'\|_{C^\theta([0,T];\mathfrak{X})}$ and $\|A\psi\|_{C^\theta([0,T];\mathfrak{X})}$. Now, for suitable constants, $\|\psi\|_{C([0,T];\mathfrak{X})} \le \|x\|_{\mathfrak{X}} + M_3\|y\|_{\mathfrak{X}} \le M_3(\|x\|_{\mathcal{D}(A)} + \|y\|_{\mathfrak{X},\mathcal{D}(A))_{\theta,\infty}})$, $\|\psi'\|_{C^\theta([0,T];\mathfrak{X})} \le M\|y\|_{(\mathfrak{X},\mathcal{D}(A))_{\theta,\infty}}$ (see §2) and $\|A\psi\|_{C^\theta([0,T];\mathfrak{X})} \le \|\psi'\|_{C^\theta([0,T];\mathfrak{X})} + \|Ax - y\|_{\mathfrak{X}} \le M_4(\|x\|_{\mathcal{D}(A)} + \|y\|_{(\mathfrak{X},\mathcal{D}(A))_{\theta,\infty}})$.■

Remark 4. From Corollary 3 it follows at once that the condition (b) of Theorem 1 is necessary for the existence of a solution u of problem (1) satisfying (3) and (4).

Whenever A is the infinitesimal generator of an analytic semigroup in $\mathfrak{X}$, we can define the operator $S_A : C([0,T];\mathfrak{X}) \to C_0([0,T];\mathfrak{X})$, $(S_A g)(t) = \int_0^t \exp((t-s)A)g(s)\,ds$: in other words, $S_A g$ is the mild solution of the Cauchy problem (6) with $x_0 = 0$. It is known from Lemma 2 that $\forall g \in C_0^\theta([0,T];\mathfrak{X})$ we have $S_A g \in C_0^\theta([0,T];\mathfrak{X})$, with $(S_A g)'$ and $A(S_A g)$ in the same space; moreover, the operators $g \mapsto (S_A g)'$ and $g \mapsto A(S_A g)$ are bounded from $C_0^\theta([0,T];\mathfrak{X})$ into itself. In particular, S_A is a bounded operator from $C_0^\theta([0,T];\mathfrak{X})$ into $C_0^\theta([0,T];\mathcal{D}(A)) \cap C^{1,\theta}([0,T];\mathfrak{X})$. Moreover, on the space $C([0,T];\mathfrak{X})$ (and *a* *fortiori* on its subspaces) S_A is one-to-one: indeed we have $g = (S_A g)' - A(S_A g)$ $\forall g \in C([0,T];\mathcal{D}(A))$ (see [9], Theorem 5.2), while $\forall g \in C([0,T];\mathfrak{X})$ we have $(\lambda - A)^{-1} g \in C([0,T];\mathcal{D}(A))$ and $(\lambda - A)^{-1} S_A g = S_A (\lambda - A)^{-1} g$.

We recall that for the semigroup $t \mapsto \exp(tA)$ we have an estimate of the kind $\|\exp(tA)\|_{\mathcal{L}(\mathfrak{X})} \le M e^{\omega t}$; this gives immediately the estimate $\|S_A\|_{\mathcal{L}(C([0,T];\mathfrak{X}))} \le \frac{M}{|\omega|}\left|e^{\omega T} - 1\right|$, and the same estimate holds in the space $\mathcal{L}(C_0^1([0,T];\mathfrak{X}))$ since for $g \in C_0^1([0,T];\mathfrak{X})$ we have $S_A(g') = (S_A g)'$ (see [9], Theorem 4.2). Now the real interpolation space $(C_0([0,T];\mathfrak{X}), C_0^1([0,T];\mathfrak{X}))_{\theta,\infty}$ is equal to $C_0^\theta([0,T];\mathfrak{X})$, with equivalent norms (see e.g. [3], p. 384), so that the same estimate (possibly with a different M) holds in $\mathcal{L}(C_0^\theta([0,T];\mathfrak{X}))$.

4. Proof of Theorem 1

Owing to the assumptions (a), (b) we deduce from Corollary 3 the existence of a function $\psi \in C^\theta([0,T];\mathcal{D}(A)) \cap C^{1,\theta}([0,T];\mathfrak{X})$ with $\psi(0) = u_1 - A_0 u_0, \psi'(0) = A_1(u_1 - A_0 u_0) + f(0, u_0, u_1)$, and from Lemma 2 the existence of a function $\phi \in C^\theta([0,T];\mathcal{D}(A_0)) \cap C^{1,\theta}([0,T];\mathfrak{X})$ such that $\phi' = A_0\phi + \psi$, with $\phi(0) = u_0$ and therefore $\phi'(0) = A_0 u_0 + \psi(0) = u_1$. For future reference, we give a look at the smoothness of ϕ. We have $\phi(t) = \exp(tA_0)u_0 + \int_0^t \exp((t-s)A_0)\psi(s)\,ds = u_0 + \int_0^1 \exp((t-s)A_0)(A_0 u_0 + \psi(s))\,ds$, whence, by formula (4.26) of [9], $\phi'(t) = \psi(t) + A_0\phi(t) = \int_0^t \exp((t-s)A_0)\psi'(s)\,ds + \exp(tA_0)u_1 - A_0 u_0$. Here the first summand belongs to $C^{1,\theta}([\varepsilon,T];\mathfrak{X})$ $\forall \varepsilon \in]0,T[$ as $\psi' \in C^\theta([0,T];\mathfrak{X})$ (see [9], formula (4.38)), while $t \mapsto \exp(tA_0)u_1$ is C^∞ away from 0. Thus $\phi' \in C^{1,\theta}([\varepsilon,T];\mathfrak{X})$, $\forall \varepsilon \in]0,T[$. If, moreover, $u_1 \in \mathcal{D}(A_0)$, with $A_0 u_1 + A_1(u_1 - A_0 u_0) + f(0,u_0,u_1) = A_0 u_1 + \psi'(0) \in \overline{\mathcal{D}(A_0)}$, then $t \mapsto \phi'(t) + A_0 u_0$ is the mild solution to the Cauchy problem $\begin{cases} z'(t) = A_0 z(t) + \psi'(t) \\ z(0) = u_1 \end{cases}$ and hence, by Theorem 4.5 of [9], ϕ' is continuously differentiable up to 0.

Now, with the aid of the functions ϕ and ψ, we define a function F on the space $C_0^\theta([0,T];\mathcal{D}(A_0)) \times C_0^\theta([0,T];\mathfrak{X})$ in the following way:

$$(F(w,z))(t) = f(t, w(t) + \phi(t), z(t) + \psi(t) + A_0 w(t) + A_0\phi(t))\ .$$

Our immediate goal is to prove that:

(i) $\forall (w,z) \in C_0^\theta([0,T];\mathcal{D}(A_0)) \times C_0^\theta([0,T];\mathfrak{X})$, we have $F(w,z) \in C^\theta([0,T];\mathfrak{X})$.

(ii) There is an increasing function $k_1 : \mathbb{R}^+ \to \mathbb{R}^+$ such that for arbitrary $(w,z), (\xi,\eta) \in C_0^\theta([0,T];\mathcal{D}(A_0)) \times C_0^\theta([0,T];\mathfrak{X})$, we have

$$\begin{aligned}&\|F(w,z) - F(\xi,\eta)\|_{C^\theta([0,T];\mathfrak{X})} \\ &\le k_1(r)\left(\|w-\xi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|z-\eta\|_{C^\theta([0,T];\mathfrak{X})}\right)\end{aligned}$$

as soon as

$$r \ge \|w\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\xi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|z\|_{C^\theta([0,T];\mathfrak{X})} + \|\eta\|_{C^\theta([0,T];\mathfrak{X})}.$$

To prove (i), we fix $(w,z) \in C_0^\theta([0,T];\mathcal{D}(A_0)) \times C_0^\theta([0,T];\mathfrak{X})$ and set $\tau = \|w + \phi\|_{C([0,T];\mathcal{D}(A_0))} + \|z + \psi\|_{C([0,T];\mathfrak{X})}$. Thus $\forall t \in [0,T]$, we have

$\|w(t) = \phi(t)\|_{\mathcal{D}(A_0)} \le \tau$, $\|z(t) + \psi(t) + A_0 w(t) + A_0\phi(t)\|_{\mathfrak{X}} \le \|z(t) + \psi(t)\|_{\mathfrak{X}} + \|w(t) + \phi(t)\|_{\mathcal{D}(A_0)} \le \tau$. By applying (d_1) we get

$$\begin{aligned}
&\|(F(w,z))(t) - (F(w,z))(s)\|_{\mathfrak{X}} \\
= \;& \|f(t, w(t) + \phi(t), z(t) = \psi(t) + A_0 w(t) + A_0\phi(t)) - \\
& \quad f(s, w(s) + \phi(s), z(s) + \psi(s) + A_0 w(s) + A_0\phi(s))\|_{\mathfrak{X}} \\
\le \;& k(\tau)(|t-s|^\theta + 2\|w(t) - w(s)\|_{\mathcal{D}(A_0)} + 2\|\phi(t) - \phi(s)\|_{\mathcal{D}(A_0)} + \\
& \quad \|z(t) - z(s)\|_{\mathfrak{X}} + \|\psi(t) - \psi(s)\|_{\mathfrak{X}})
\end{aligned}$$

thus proving (i).

In order to prove (ii), we begin by estimating $\|F(w,z) - F(\xi,\eta)\|_{C([0,T];\mathfrak{X})}$. Owing to the choice of r, we have $\forall\, t \in [0,T]$:

$$\begin{aligned}
\|w(t) + \phi(t)\|_{\mathcal{D}(A_0)} &\le r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}, \\
\|z(t) + \psi(t) + A_0 w(t) + A_0\phi(t)\|_{\mathfrak{X}} &\le r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\psi\|_{C^\theta([0,T];\mathfrak{X})} \\
\|\xi(t) + \phi(t)\|_{\mathcal{D}(A_0)} &\le r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}, \\
\|\eta(t) + \psi(t) + A_0\xi(t) + A_0\phi(t)\|_{\mathfrak{X}} &\le r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\psi\|_{C^\theta([0,T];\mathfrak{X})}.
\end{aligned}$$

Hence, by (d_1), we have:

$$\begin{aligned}
&\|(F(w,z))(t) - (F(\xi,\eta))(t)\|_{\mathfrak{X}} \\
= \;& \|f(t, w(t) + \phi(t), z(t) + \psi(t) + A_0 w(t) + A_0\phi(t)) - \\
& \quad f(t, \xi(t) + \phi(t), \eta(t) + \psi(t) + A_0\xi(t) + A_0\phi(t))\|_{\mathfrak{X}} \\
\le \;& k(r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\psi\|_{C^\theta([0,T];\mathfrak{X})}) \cdot \\
& \quad (2\|w(t) - \xi(t)\|_{\mathcal{D}(A_0)} + \|z(t) - \eta(t)\|_{\mathfrak{X}}),
\end{aligned}$$

so that

$$\begin{aligned}
&\|F(w,z) - F(\xi,\eta)\|_{C([0,T];\mathfrak{X})} \\
\le \;& 2k(r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\psi\|_{C^\theta([0,T];\mathfrak{X})}) \cdot \\
& \quad (\|w - \xi\|_{C([0,T];\mathcal{D}(A_0))} + \|z - \eta\|_{C([0,T];\mathfrak{X})}).
\end{aligned}$$

Now for $(\tau,\rho) \in [0,T] \times [0,1]$ we set

$$\begin{aligned}
\lambda(\tau,\rho) = (\tau, (1-\rho)\xi(\tau) + \rho w(\tau) + \phi(\tau), (1-\rho)(\eta(\tau) + A_0\xi(\tau)) \\
+ \rho(z(\tau) + A_0 w(\tau)) + \psi(\tau) + A_0\phi(\tau)),
\end{aligned}$$

so that $(F(w,z))(\tau) = f(\lambda(\tau,1))$, $(F(\xi,\eta))(\tau) = f(\lambda(\tau,0))$. Therefore

$$\begin{aligned}
&(F(w,z) - F(\xi,\eta))\,(t) - (F(w,z) - F(\xi,\eta))\,(s) \\
= \;& \int_0^1 \frac{\partial}{\partial\rho} f\,(\lambda(t,\rho))\; d\rho - \int_0^1 \frac{\partial}{\partial\rho} f\,(\lambda(s,\rho))\; d\rho
\end{aligned}$$

$$\begin{aligned}
= & \int_0^1 [\partial_2 f(\lambda(t,\rho))(w(t)-\xi(t)) + \partial_3 f(\lambda(t,\rho)) \\
& \quad \cdot (z(t)-\eta(t)+A_0 w(t) - A_0\xi(t))]\, d\rho - \\
& \int_0^1 [\partial_2 f(\lambda(s,\rho))(w(s)-\xi(s)) + \partial_3 f(\lambda(s,\rho)) \\
& \quad \cdot (z(s)-\eta(s)+A_0 w(s) - A_0\xi(s))]\, d\rho \\
= & \int_0^1 [\partial_2 f(\lambda(t,\rho)) - \partial_2 f(\lambda(s,\rho))](w(t)-\xi(t))\, d\rho + \\
& \int_0^1 \partial_2 f(\lambda(s,\rho))[(w-\xi)(t)-(w-\xi)(s)]\, d\rho + \\
& \int_0^1 [\partial_3 f(\lambda(t,\rho)) - \partial_3 f(\lambda(s,\rho))](z(t)-\eta(t)+A_0w(t)-A_0\xi(t))\, d\rho + \\
& \int_0^1 \partial_3 f(\lambda(s,\rho))[(z-\eta+A_0w-A_0\xi)(t)-(z-\eta+A_0w-A_0\xi)(s)]\, dp = \sum_{k=1}^4 I_k\ .
\end{aligned}$$

Remark that

$$\|(1-\rho)\xi(\tau)+\rho w(\tau)+\phi(\tau)\|_{\mathcal{D}(A_0)} = \|(1-\rho)(\xi(\tau)+\phi(\tau))+\rho(w(\tau)+\phi(\tau))\|_{\mathcal{D}(A_0)}$$
$$\leq \max\left\{\|\xi+\phi\|_{\mathrm{C}([0,\mathrm{T}];\mathcal{D}(\mathrm{A}_0))}, \|\mathrm{w}+\phi\|_{\mathrm{C}([0,\mathrm{T}];\mathcal{D}(\mathrm{A}_0))}\right\} \leq \mathrm{r} + \|\phi\|_{\mathrm{C}^\theta([0,\mathrm{T}];\mathcal{D}(\mathrm{A}_0))} \text{ and}$$

$$\|(1-\rho)(\eta(\tau)+A_0\xi(\tau))+\rho(z(\tau)+A_0w(\tau))+\psi(\tau)+A_0\phi(\tau)\|_{\mathfrak{X}} =$$
$$\|(1-\rho)(\eta(\tau)+\psi(\tau)+A_0\xi(\tau)+A_0\phi(\tau))+\rho(z(\tau)+\psi(\tau)+A_0w(\tau)+A_0\phi(\tau))\|_{\mathfrak{X}} \leq$$
$$\max\{\|\mathrm{w}+\phi\|_{\mathrm{C}([0,\mathrm{T}];\mathcal{D}(\mathrm{A}_0))}+\|\mathrm{z}+\psi\|_{\mathrm{C}([0,\mathrm{T}];\mathfrak{X})}, \|\xi+\phi\|_{\mathrm{C}([0,\mathrm{T}];\mathcal{D}(\mathrm{A}_0))}+\|\eta+\psi\|_{\mathrm{C}([0,\mathrm{T}];\mathfrak{X})}\}$$
$$\leq r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\psi\|_{C^\theta([0,T];\mathfrak{X})}\ .$$

Hence, by applying (d_2) and (d_3), we get:

$$\|I_2\|_{\mathfrak{X}} + \|I_3\|_{\mathfrak{X}} \leq$$
$$k(r+\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})(1+2r+2\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})\cdot$$
$$\cdot|t-s|^\theta(\|w(t)-\xi(t)\|_{\mathcal{D}(A_0)}+\|z(t)-\eta(t)+A_0w(t)-A_0\xi(t)\|_{\mathfrak{X}}) \leq$$
$$2k(r+\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})(1+2r+2\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+$$
$$\|\psi\|_{C^\theta([0,T];\mathfrak{X})})|t-s|^\theta(\|w-\xi\|_{C([0,T];\mathcal{D}(A_0))}+\|z-\eta\|_{C([0,T];\mathfrak{X})})\ .$$

Now we estimate I_2 and I_4. Remark that

$$\|\partial_2 f(\lambda(s,\rho))\|_{\mathcal{L}(\mathcal{D}(A_0)<\mathfrak{X})} \leq \|\partial_2 f(0,0,0)\|_{\mathcal{L}(\mathcal{D}(A_0),\mathfrak{X})}+$$
$$k(r+\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})[s^\theta+\max\{\|\xi(\mathrm{s})+\phi(\mathrm{s})\|_{\mathcal{D}(A_0)},$$
$$\|w(s)+\phi(s)\|_{\mathcal{D}(A_0)}\}+\max\{\|\eta(\mathrm{s})+\mathrm{A}_0\xi(\mathrm{s})+\psi(\mathrm{s})+\mathrm{A}_0\phi(\mathrm{s})\|_{\mathfrak{X}},$$
$$\|z(s)+A_0w(s)+\psi(s)+A_0\phi(s)\|_{\mathfrak{X}}\}] \leq$$
$$\|\partial_2 f(0,0,0)\|_{\mathcal{L}(\mathcal{D}(A_0),\mathfrak{X})}+k(r+\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})\cdot$$
$$\cdot(T^\theta+2r+2\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})\ ,$$

and analogously

$$\|\partial_3 f(\lambda(s,\rho))\|_{\mathcal{L}(\mathfrak{X})} \leq \|\partial_3 f(0,0,0)\|_{\mathcal{L}(\mathfrak{X})}+$$
$$k(r+\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})(T^\theta+2r+2\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}+$$
$$\|\psi\|_{C^\theta([0,T];\mathfrak{X})})\ .$$

Therefore,

$$\|I_2\|_{\mathfrak{X}} + \|I_4\|_{\mathfrak{X}} \le 2|t-s|^\theta(\|w-\xi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|z-\eta\|_{C^\theta([0,T];\mathfrak{X})})\cdot$$
$$\cdot[\|\partial_2 f(0,0,0)\|_{\mathcal{L}(\mathcal{D}(A_0),\mathfrak{X})} + \|\partial_3 f(0,0,0)\|_{\mathcal{L}(\mathfrak{X})} + k(r + \|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))}$$
$$+\|\psi\|_{C^\theta([0,T];\mathfrak{X})})$$
$$(T^\theta + 2r + 2\|\phi\|_{C^\theta([0,T];\mathcal{D}(A_0))} + \|\psi\|_{C^\theta([0,T];\mathfrak{X})})]\ .$$

So we have proved (ii).

It is an obvious remark that if $0 < T_1 \le T$ and (w,z), (ξ,η) belong to the space $C_0^\theta([0,T];\mathcal{D}(A_0)) \times C_0^\theta([0,T];\mathfrak{X})$, then (i) and (ii) still hold when we replace everywhere T with T_1.

Now we set $S_j = S_{A_j}$ and $\forall\, t \in]0,T]$ we call:

$q_0(t)$ the norm of S_0 in the space $\mathcal{L}(C^\theta([0,t];\mathfrak{X}), C^\theta([0,t];\mathcal{D}(A_0)))$
$q_1(t)$ the norm of S_1 in $\mathcal{L}(C^\theta([0,t];\mathfrak{X}))$
$C_0(t)$ the norm of $F(0,0) + A_1\psi - \psi$ in $C^\theta([0,t];\mathfrak{X})$
$C_1(t)$ the norm of ϕ in $C^\theta([0,t];\mathfrak{X})$
$C_2(t)$ the norm of ψ in $C^\theta([0,t];\mathfrak{X})$.

(When we omit the mention of t, it is understood that $t = T$.)

It is obvious that all these functions of $t \in]0,T]$ are increasing, moreover, by the concluding remarks of §3, $q_j(t) \to 0$ as $t \to 0$ ($j = 0,1$). Therefore,

we fix $t_0 \in]0,T]$ such that $q_1(t_0) < \frac{1}{1+q_0}$,
then we fix $r_0 > C_0$ such that $\frac{C_1}{r_0} \le 1 - q_0q_1(t_0)$, $\frac{C_2}{r_0} \le 1 - q_1(t_0)$,
then we fix $r \ge r_0$,
then we set $k_2(r) = \max\{\mathrm{C}_0 + (2\mathrm{r} + \mathrm{C}_1 + \mathrm{C}_2)\mathrm{k}_1(2\mathrm{r} + \mathrm{C}_1 + \mathrm{C}_2), \mathrm{r}\}$,
then we fix $t_1 \in]0,t_0]$ such that $q_1(t_1)(1+q_0)k_1(q_1(t_1)(1+q_0)k_2(r)) \le 1 - \frac{C_0}{r}$ and $q_1(t_1)(1+q_0)k_1(4r) < 1$.

Remark that all these inequalities concerning t_0 and t_1 hold a fortiori for$0 < t \le t_1$. Then we prove that:

(iii) if $T_1 \in]0,t_1]$, then a necessary and sufficient condition for $u : [0,T_1] \to \mathfrak{X}$ to be a solution of problem (1) satisfying conditions (3), (4) and (5) is that $u = \phi + S_0S_1h$, where h is a solution of the nonlinear equation

$$h = F(S_0S_1h, S_1h) + A_1\psi - \psi' \tag{7}$$

in the closed ball centered at 0 and with radius r of the Banach space $C_0^\theta([0,T_1];\mathfrak{X})$.

Indeed, suppose that $u = \phi + S_0S_1h$, where h is a solution of the equation (7) in $C_0^\theta([0,T_1];\mathfrak{X})$ with $\|h\|_{C^\theta([0,T_1];\mathfrak{X})} \le r$. Then $u(0) = \phi(0) = u_0$ and $u \in C^\theta([0,T_1];\mathcal{D}(A_0)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})$ (see the remarks at the end of §3). Moreover, $u' - A_0u = \phi' - A_0\phi + S_1h = \psi + S_1h \in C^\theta([0,T_1];\mathcal{D}(A_1)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})$.

Hence $u'(0) = A_0u_0 + \psi(0) = u_1$ and $(u' - A_0u)' - A_1(u' - A_0u) = \psi' - A_1\psi + h = F(S_0S_1h, S_1h) = F(u - \phi, u' - A_0u - \psi) = f(\cdot, u(\cdot), u'(\cdot))$. Thus u is a strict solution of (1) with the required regularity. Finally,

$$\|u\|_{C^\theta([0,T_1];\mathcal{D}(A_0))} \le C_1 + q_0q_1(t_0)r = \left(\frac{C_1}{r} + q_0q_1(t_0)\right)r$$
$$\le \left(\frac{C_1}{r_0} + q_0q_1(t_0)\right)r \le r;\ \|u' - A_0u\|_{C^\theta([0,T_1];\mathfrak{X})}$$
$$\le C_2 + q_1(t_0)r = \left(\frac{C_2}{r} + q_1(t_0)\right)r \le \left(\frac{C_2}{r_0} + q_1(t_0)\right)r \le r\ .$$

Conversely, if u is a solution of (1) with the properties (3)-(4)-(5), then $(u-\phi)(0)=0$ and $(u-\phi)(t)-A_0((u-\phi)(t))=u'(t)-A_0u(t)-\psi(t)$, so that $u-\phi=S_0(u'-A_0u-\psi)$. On the other hand, if we set $v=u'-A_0u-\psi$, then $v(0)=0$ and $v'(t)=(u'-A_0u)'(t)-\psi'(t)=A_1(u'(t)-A_0u(t))+f(t,u(t),u'(t))-\psi'(t)=A_1v(t)+A_1\psi(t)+f(t,u(t),v(t)+A_0u(t)+\psi(t))-\psi'(t)=A_1v(t)+A_1\psi(t)-\psi'(t)+F(u-\phi,v)(t)$. Thus $v=S_1(A_1\psi-\psi'+F(u-\phi,v))$, and setting $h=A_1\psi-\psi'+F(u-\phi,v)$ we get $h=A_1\psi-\psi'+F(S_0S_1h,S_1h)$. In order to estimate the norm of h in $C^\theta([0,T_1];\mathfrak{X})$, we remark that $\|u-\phi\|_{C^\theta([0,T_1];\mathcal{D}(A_0))}+\|v\|_{C^\theta([0,T_1];\mathfrak{X})}\le 2r+C_1+C_2$, and since $h=A_1\psi-\psi'+F(0,0)+F(u-\phi,v)-F(0,0)$, from (ii) it follows that a crude estimate of h is $\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\le C_0+k_1(2r+C_1+C_2)(2r+C_1+C_2)\le k_2(r)$. Hence $\|S_0S_1h\|\backslash s\backslash do5(C^\theta([0,T_1];\mathcal{D}(A_0)))\le q_0q_1(T_1)k_2(r)$, $\|S_1h\|_{C^\theta([0,T_1];\mathfrak{X})}\le q_1(t_1)k_2(r)$, and so from $h=A_1\psi-\psi'+F(0,0)+F(S_0S_1h,S_1h)-F(0,0)$ it follows that $\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\le C_0+\|h\|_{C^\theta([0,T_1];\mathfrak{X})}q_1(T_1)(1+q_0)k_1(k_2(r)q_1(T_1)\cdot\cdot(1+q_0))$, that is

$$\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\le\frac{C_0}{1-q_1(T_1)(1+q_0)k_1(k_2(r)q_1(T_1)(1+q_0))}\le r\,.$$

Owing to the statement (iii), the existence and uniqueness of the solution on $[0,T_1]$ of problem (1), with the conditions (3), (4), (5) is equivalent to the existence and uniqueness of the solution of the equation (7) in the closed ball centered in 0 and with radius r of the Banach space $C_0^\theta([0,T];\mathfrak{X})$: indeed, as already remarked, the operators S_j are one-to-one on $C^\theta([0,T];\mathfrak{X})$. Now, suppose that $\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\le r$. Then $\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\le k_2(r)$ and, as above, we see that the norm of $A_1\psi-\psi'+F(S_0S_1h,S_1h)$ in $C_0^\theta([0,T_1];\mathfrak{X})$ is $\le r$. On the other hand, as $\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\le r$, we have $\|S_0S_1h\|_{C^\theta([0,T_1];\mathcal{D}(A_0))}\le q_0q_1(t_1)r\le r$, $\|S_1h\|_{C^\theta([0,T_1];\mathfrak{X})}\le q_1(t_1)r\le r$ and hence, if $\max\{\|g\|_{C^\theta([0,T_1];\mathfrak{X})},$ $\|h\|_{C^\theta([0,T_1];\mathfrak{X})}\}\le r$, we get the inequality

$$\begin{aligned}&\|F(S_0S_1g,S_1g)-F(S_0S_1h,S_1h)\|_{C^\theta([0,T_1];\mathfrak{X})}\\&\qquad\le k_1(4r)(1+q_0)q_1(t_1)\|g-h\|_{C^\theta([0,T_1];\mathfrak{X})},\end{aligned}$$

, where $k_1(4r)(1+q_0)q_1(t_1)<1$. Thus an application of the Banach-Caccioppoli fixed point theorem gives the result of existence and uniqueness.

As for the last statement of the theorem, we have already seen the regularity of ϕ. Moreover, from $h\in C_0^\theta([0,T];\mathfrak{X})$ it follows that S_1h is continuously differentiable with $(S_1h)'\in C_0^\theta([0,T];\mathfrak{X})$ (see [9], Theorem 4.1), so that $(S_0S_1h)'=S_0((S_1h)')$ is again in $C^{1,\theta}([0,T_1];\mathfrak{X})$. ■

5. Examples and applications

Some second-order equations of non-factored type can be reduced to equations of the kind considered above.

Example 5. Let $B_0:\mathcal{D}(B_0)\to\mathfrak{X}$ be the infinitesimal generator of an analytic semigroup in the Banach space $\mathfrak{X}$, with $\overline{\mathcal{D}(B_0)}=\mathfrak{X}$. Let $(B_1(t))_{t\in[0,T]}$ be a family of closed linear operators acting in $\mathfrak{X}$, bounded with respect to B_0 uniformly in t; this means that $\mathcal{D}(B_0)\subseteq\bigcap_t\mathcal{D}(B_1(t))$ and that $\exists\ C>0$ such that $\forall\ t\in[0,T]\ \forall\ x\in\mathcal{D}(B_0)\|B_1(t)x\|_{\mathfrak{X}}\le C\|x\|_{\mathcal{D}(B_0)}$ (and in particular $B_1(t)_{|\mathcal{D}(B_0)}\in\mathcal{L}(\mathcal{D}(B_0),\mathfrak{X})\ \forall\ t$). We suppose that:

$$\|B_1(t)-B_1(s)\|_{\mathcal{L}(\mathcal{D}(B_0);\mathfrak{X})}\le C|t-s|^\theta\,,\tag{8}$$

Now, let $g : [0,T] \times \mathcal{D}(B_0) \times \mathfrak{X} \to \mathfrak{X}$ be a function satisfying conditions (c)-(d) of Theorem 1 (with B_0 instead of A_0). We consider the problem

$$(9) \qquad \begin{cases} u''(t) - B_0 u'(t) + B_1(t)u(t) = g(t, u(t), u'(t)) \\ u(0) = u_0 \\ u'(0) = u_1 \end{cases}$$

in some interval $[0, T_1] \subseteq [0, T]$.

An appropriate definition of a strict solution to the equation appearing in (9) will require that $u \in C^2([0,T]); \mathfrak{X}) \cap C([0,T]; \mathcal{D}(B_0))$, $u' \in C([0,T]; \mathcal{D}(B_0))$ and the equality holds in $[0, T_1]$. Remark that from the closedness of B_0, the continuity of $B_0 u'$ and the identity $u(t) - u(s) = \int_s^t u'(\tau)\, d\tau$, we get $B_0 u(t) - B_0 u(s) = \int_s^t B_0 u'(\tau)\, d\tau$, and hence $B_0 u \in C^1([0,T_1]; \mathfrak{X})$, so that $u \in C^1([0,T_1]; \mathcal{D}(B_0))$. Conversely, if $u \in C^1([0,T_1]; \mathcal{D}(B_0))$ then (because of the closedness of B_0) $(B_0 u') = (B_0 u)'$, so that $u' \in C([0,T_1]; \mathcal{D}(B_0))$. Thus the regularity of the solution could have been expressed by $u \in C^2([0,T]; \mathfrak{X}) \cap C^1([0,T]; \mathcal{D}(B_0))$.

Now we choose $\omega \in \mathbb{R}$ such that $(B + \omega)^{-1}$ exists and belongs to $\mathcal{L}(\mathfrak{X})$ and we consider the problem

$$(10) \qquad \begin{cases} \left(\frac{d}{dt} - B_1(0)(B_0 + \omega)^{-1}\right)\left(\frac{d}{dt} - (B_0 + \omega)\right) u = f(t, u(t), u'(t)) \\ u(0) = u_0 \\ u'(0) = u_1 \end{cases}$$

where $f(t,u,v) = g(t,u,v) - (\omega + B_1(0)(B_0 + \omega)^{-1})v + (B_1(0) - B_1(t))u$. Problems (9) and (10) are formally equivalent, while if we refer to problem (1) we have $A_1 = B_1(0)(B_0 + \omega)^{-1} \in \mathcal{L}(\mathfrak{X})$ and $A_0 = B_0 + \omega$, so that $\mathcal{D}(A_0) = \mathcal{D}(B_0)$ (with equivalent norms). A non formal equivalence between problems (9) and (10) can be proved after showing that f satisfies conditions (c)–(d) of Theorem 1. Once we have shown this, we have:

> u is a strict solution to (9) with $u'' - B_0 u' \in C^\theta([0,T_1]; \mathfrak{X})$ if and only if u is a strict solution to (10) such that (3) and (4) (of Theorem 1) hold.

Indeed, suppose that u is a strict solution to (9) with $u'' - B_0 u' \in C^\theta([0,T_1]; \mathfrak{X})$. Then $u \in C^2([0,T_1]; \mathfrak{X}) \cap C^1([0,T_1]; \mathcal{D}(B_0))$, and (3) follows from $\mathcal{D}(A_0) = \mathcal{D}(B_0)$. Moreover, $u' - A_0 u = u' - B_0 u - \omega u \in C^{1,\theta}([0,T_1]; \mathfrak{X})$, and since $\mathcal{D}(A_1) = \mathfrak{X}$, (4) holds.

Conversely, suppose that u is a solution to (10) satisfying (3) and (4). As $\overline{\mathcal{D}(B_0)} = \mathfrak{X}$, from Theorem 1 we get $u \in C^2([0,T_1]; \mathfrak{X})$. Hence, from (4) we get $B_0 u = (A_0 - \omega)u = u' - (u' - A_0 u) - \omega u \in C^1([0,T_1]; \mathfrak{X})$, so that $u \in C^1([0,T_1]; \mathcal{D}(B_0))$; finally $u'' - B_0 u' = \frac{d}{dt}(u' - A_0 u) + \omega u' \in C^\theta([0,T_1]; \mathfrak{X})$ by (4).

Thus it remains to prove the asserted properties of f or, what is the same, of the function $(t,u,v) \mapsto F(t,u,v) = (B_1(0) - B_1(t))u - (\omega + B_1(0)(B_0 + \omega)^{-1})v$. Now $\partial_3 F(t,u,v) = \omega + B_1(0)(B_0 + \omega)^{-1}$, so that (d_3) of Theorem 1 is trivial; $\partial_2 F(t,u,v) = B_1(0) - B_1(t)$, and (d_2) is an easy consequence of assumption (8). As for (d_1) we have: $F(s,u,v) - F(t,w,z) = (B_1(0) - B_1(s))u - (B_1(0) - B_1(t))w - B_1(0)(B_0 + \omega)^{-1}(v - z) = (B_1(0) - B_1(s))(u - w) + (B_1(t) - B_1(s))w - B_1(0)(B_0 + \omega)^{-1}(v - z)$; therefore $\|F(s,u,v) - F(t,w,z)\|_{\mathfrak{X}} \le C(\|u - v\|_{\mathcal{D}(A_0)} + |t - s|^\theta \|w\|_{\mathcal{D}(A_0)} + \|v - z\|_{\mathfrak{X}})$, and the statement is proved.

Remark that the conditions imposed on u_0 and u_1 are now $u_0, u_1 \in \mathcal{D}(B_0)$.

Application 6. Let Ω be a bounded open set of $\mathbb{R}^n$, with smooth boundary, and $T \in \mathbb{R}^+$. On $[0,T] \times \Omega$ we consider the following initial-boundary value problem, to which we aim to apply Example 5:

$$(11) \qquad \begin{cases} \frac{\partial^2 u}{\partial t^2} - \alpha\Delta \frac{\partial u}{\partial t} + A(t,x,\partial_x)u = \beta u|u|^{q-1} & \text{on} \quad [0,T] \times \Omega \\ u(t,x) = 0 & \text{on} \quad [0,T] \times \partial\Omega \\ u(0,x) = u_0(x) & \text{on} \quad \Omega \\ \frac{\partial u}{\partial t}(0,x) = u_1(x) & \text{on} \quad \Omega. \end{cases}$$

Here $\alpha \in \mathbb{R}^+$, Δ is the Laplace operator, $\beta \in L^\infty(\Omega)$, $q \in [2, +\infty[$ and A is the second order linear differential operator formally defined by $(A(t,x,\partial_x)v)(x) = \sum_{i,j=1}^{n} a_{ij}(t,x)\frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_{k=1}^{n} b_k(t,x)\frac{\partial v}{\partial x_k}(x) + c(t,x)v(x)$. We remark explicitly that no sign or ellipticity condition is imposed to the operator. We take $p \in [n(q-1)/2q, \infty[$ and set $\mathfrak{X} = L^p(\Omega;\mathbb{R}), \mathcal{D}(B_0) = W_0^{2,p}(\Omega,\mathbb{R}) \cap W_0^{1,p}(\Omega,\mathbb{R})$, $B_0 u = \alpha\Delta u$. Then B_0 is the (densely defined) infinitesimal generator of an analytic semigroup in $\mathfrak{X}$. Our assumptions on the coefficients of the differential operator A are the following (with $q' = \frac{q}{q-1}$):

$$(12) \qquad \begin{aligned} &\forall\, t \in [0,T]\ \forall\, i,j,k \in \{1,\dots,n\} \\ &a_{ij}(t,\cdot) \in L^\infty(\Omega,\mathbb{R}), b_k(t,\cdot) \in L^{2pq'}(\Omega,\mathbb{R}), c(t,\cdot) \in L^{pq'}(\Omega;\mathbb{R}) \end{aligned}$$

$$(13) \quad \begin{aligned} &\exists\, C > 0, \theta \in]0,1[\text{ such that for arbitrary } t,s \in [0,T], i,j,k \in \{1,\dots,n\} \\ &\|a_{ij}(t,\cdot) - a_{ij}(s,\cdot)\|_{L^\infty} \le C|t-s|^\theta \\ &\|b_k(t,\cdot) - b_k(s,\cdot)\|_{L^{2pq'}} \le C|t-s|^\theta \\ &\|c(t,\cdot) - c(s,\cdot)\|_{L^{pq'}} \le C|t-s|^\theta . \end{aligned}$$

Finally we suppose $u_0, u_1 \in W^{2,p}(\Omega,\mathbb{R})$.

Under these assumptions problem (11) has a unique local (in time) solution $u \in C^2([0,T_1];L^p(\Omega)) \cap C^1([0,T_1];W^{2,p}(\Omega))$ such that $\frac{\partial^2 u}{\partial t^2} - \alpha\Delta\frac{\partial u}{\partial t} \in C^\theta([0,T_1];L^p(\Omega))$.

Indeed the requirements of Example 5 concerning the linear part of the equation are verified in an elementary way by means of: assumptions (12) and (13), Sobolev's embedding theorems, Hölder's inequality and the equivalence between $\|u\|_{W^{2,p}}$ and $\|\Delta u\|_{L^p}$ when $u \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$. As for the nonlinear part, standard results (see e.g. [1]) show that $u \mapsto g(u) = \beta u|u|^{q-1}$ is a differentiable function from $L^{pq}(\Omega;\mathbb{R})$ to $L^p(\Omega,\mathbb{R})$, with $dg(u)(h) = q\beta h|u|^{q-1}$, so that $u \mapsto dg(u)$ is a locally lipschitz function from $L^{pq}(\Omega;\mathbb{R})$ to the space $\mathcal{L}(L^{pq}(\Omega;\mathbb{R}), L^p(\Omega;\mathbb{R}))$ (to be more precise, if $r = \max\{\|u\|_{L^{pq}}, \|v\|_{L^{pq}}\}$, then $\|dg(u) - dg(v)\|_{\mathcal{L}(L^{pq},L^p)} \le q(q-1)\|\beta\|_{L^\infty} r^{q-2}\|u-v\|_{L^{pq}}$, and here $q \ge 2$ is needed). Since $W^{2,p}(\Omega)$ is continuously embedded in $L^{pq}(\Omega)$ (by our choice of p and Sobolev's embedding theorems), we get the same results if we replace everywhere $L^{pq}(\Omega)$ by $W^{2,p}(\Omega)$.

Application 7. Analogous results hold for the problem

$$(14)\qquad \begin{cases} \frac{\partial^2 u}{\partial t^2} - \alpha\Delta\frac{\partial u}{\partial t} + \gamma(t,x)\Delta u = \beta u|u|^{q-1} & \text{on } [0,T]\times\Omega \\ u(t,x) = 0 & \text{on } [0,T]\times\partial\Omega \\ u(0,x) = u_0(x) & \text{on } \Omega \\ \frac{\partial u}{\partial t}(0,x) = u_1(x) & \text{on } \Omega\ . \end{cases}$$

when $\mathfrak{X}$ is the space $C_0(\overline{\Omega};\mathbb{R})$ of the continuous functions from Ω to $\mathbb{R}$ which vanish on $\partial\Omega$. Here $\alpha \in \mathbb{R}^+$, $q \geq 2$, $u_0, u_1 \in \mathcal{D}(B_0) = \{u \in C_0(\overline{\Omega};\mathbb{R});\ \forall\, r > n\ u \in W^{2,r}(\Omega;\mathbb{R}),\ \Delta u \in C_0(\overline{\Omega};\mathbb{R})\}$, $B_0 u = \Delta u$, and moreover, we suppose that $\beta \in C(\overline{\Omega};\mathbb{R})$; $\gamma(t,\cdot) \in C(\overline{\Omega};\mathbb{R})$ with $t \mapsto \gamma(t,\cdot)$ Hölder continuous of exponent θ (with values in the Banach space $C(\overline{\Omega};\mathbb{R})$).

Application 8. Such problems as

$$(15)\qquad \begin{cases} \frac{\partial^2 u}{\partial t^2} + (-1)^m\alpha\Delta^m\frac{\partial u}{\partial t} \pm \Delta^m u = f(t,u,u') & \text{on } [0,T]\times\Omega \\ u(t,x) = \ldots = \Delta^{m-1}u(t,x) = 0 & \text{on } [0,T]\times\partial\Omega \\ u(0,x) = u_0(x) & \text{on } \Omega \\ \frac{\partial u}{\partial t}(0,x) = u_1(x) & \text{on } \Omega\ . \end{cases}$$

can be dealt with by the technique of Example 5, taking into account Corollary 5 of [4], whence it follows that $(-1)^{m+1}\Delta^m$ (with the above boundary conditions) generates an analytic semigroup in $L^p(\Omega)$.

Example 9. For $j = 0,1$, let $B_j : \mathcal{D}(B_j) \to \mathfrak{X}$ be the infinitesimal generator of an analytic semigroup in the Banach space $\mathfrak{X}$, with $\overline{\mathcal{D}(B_j)} = \mathfrak{X}$. We suppose, moreover, that $\mathcal{D}(B_0) \subseteq \mathcal{D}(B_1^2)$, so that B_1^2 is bounded with respect to B_0. Let $g : [0,T]\times\mathcal{D}(B_0)\times\mathfrak{X} \to \mathfrak{X}$ be a function satisfying conditions (c)–(d) of Theorem 1 (with B_0 instead of A_0). We consider the problem

$$(16)\qquad \begin{cases} u''(t) - B_0 u'(t) + B_1 B_0 u(t) = g(t,u(t),u'(t)) \\ u(0) = u_0 \\ u'(0) = u_1 \end{cases}$$

in some interval $[0,T_1] \subseteq [0,T]$. This problem is formally equivalent to

$$(17)\qquad \begin{cases} \left(\frac{d}{dt} - B_1\right)\left(\frac{d}{dt} - (B_0 - B_1)\right)u = g(t,u(t),u'(t)) - B_1^2 u \\ u(0) = u_0 \\ u'(0) = u_1\ . \end{cases}$$

It is our aim to apply Theorem 1 to problem (17). To this end, we set $A_1 = B_1$, $A_0 = B_0 - B_1$ (with $\mathcal{D}(A_0) = \mathcal{D}(B_0)$), $f(t,u,v) = g(t,u,v) - B_1^2 u$. Since B_1^2 is bounded with respect to B_0, it is obvious that f satisfies again the conditions of Theorem 1. On the other hand, by the moment inequality (see e.g., [7], Chapter 1, Lemma 2.8) we have that for suitable $\omega \in \mathbb{R}$, $M \in \mathbb{R}^+$, $\|(B_1+\omega)u\| \leq C\|(B_1+\omega)^2 u\|^{1/2}\|u\|^{1/2}\forall\, u \in \mathcal{D}((B_1+\omega)^2) = \mathcal{D}(B_1^2)$; in particular

$\|(B_1+\omega)u\| \leq M'\|B_0u\|^{1/2}\|u\|^{1/2}$ $\forall\, u \in \mathcal{D}(B_0)$. Therefore $\forall\, \eta \in \mathbb{R}^+$ $\exists\, C_\eta \in \mathbb{R}^+$ such that $\forall\, u \in \mathcal{D}(B_0) \|(B_1+\omega)u\| \leq \eta\|B_0u\| + C_\eta\|u\|$. Thus the B_0-bound of B_1 is 0, and hence $B_0 - B_1$ is the generator of an infinitesimal semigroup (see [7], Chapter 3, Theorem 2.1).

By applying Theorem 1 to problem (17), we get local (in time) existence and uniqueness of the solution $u \in C^\theta([0,T_1];\mathcal{D}(B_0)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})$, with $u' - B_0u \in C^\theta([0,T_1];\mathcal{D}(B_0)) \cap C^{1,\theta}([0,T_1];\mathfrak{X})$. Now it is easy to see that u is a solution of (17) satisfying these conditions if and only if it is a solution of (16) such that $u \in C^2([0,T_1];\mathfrak{X}) \cap C^1([0,T^1];\mathcal{D}(B_0))$ with $u' - B_0u \in C^\theta([0,T_1];\mathcal{D}(B_1))$ and $u'' - B_0u' \in C^\theta([0,T_1];\mathfrak{X})$.

References

[1] Appell, J. and P. P. Zabrejko, "Nonlinear Superposition Operators", Cambridge University Press, Cambridge, 1990.

[2] Aviles, P. and J. Sandefur, *Nonlinear second order equations withapplications to partial differential equations*, J. Differential Equations **58** (1980), 404–427.

[3] Da Prato, G. and P. Grisvard, *Sommes d'opérateurs linéaires et équations différentielles opérationnelles*, J. Math. Pures Appl. **54** (1975), 305–387.

[4] DeLaubenfels, R., *Powers of generators of holomorphic semigroups*, Proc. Amer. Math. Soc. **99** (1987), 105–108.

[5] Engler, H., F. Neubrander and J. Sandefur, *Strongly damped semilinear second order equations*, in: T. L. Gill and W. W. Zachary (eds.) "Nonlinear Semigroups, Partial Differential Equations and Attractors", Lecture Notes in Math. **1248** (1986), 52–62.

[6] Neubrander, F., *Well-posedness of higher order abstract Cauchy problems*, Trans. Amer. Math. Soc. **295** (1986), 257–290.

[7] Pazy, A., "Semigroups of Linear Operators and Applications to Partial Differential Equations", Springer, New York Berlin Heidelberg Tokyo, 1983.

[8] Sandefur, J. T. Jr., *Existence and uniqueness of solutions of second order nonlinear differential equations*, SIAM J. Math. Anal. **14** (1983), 477–487.

[9] Sinestrari, E., *On the abstract Cauchy problem of parabolic type in spaces of continuous functions*, J. Math. Anal. Appl. **107** (1985), 16–66.

[10] Tanabe, H., "Equations of Evolution", Pitman, London San Francisco Melbourne, 1979.

[11] Triebel, H., "Interpolation Theory, Function Spaces, Differential Operators", North Holland, Amsterdam New York Oxford, 1978.

Dipartimento di Matematica
Piazza di Porta San Donato 5
I-40127 Bologna - Italia

Received March 2, 1990
and in final form March 13, 1991

Semigroup Forum Vol. 45 (1992) 148–165

RESEARCH ARTICLE

On Limits in Complete Semirings

Georg Karner

Communicated by K. Keimel

Abstract. Various special cases of complete semirings are presented in a systematic way. We give new characterizations and a lot of examples.

0. Introduction

The notion of a complete semiring ([1], [2]) came up when classical formal language and automata theory was generalized to formal power-series (in noncommuting variables, with coefficients in a semiring). In the subsequent research, most authors considered more general concepts of summability (see [13] for an overview). On the other hand, it turned out that it may be useful to have additional axioms for complete semirings. Some of these axioms are motivated by the interconnection with the algebraic limit concept of Kuich and Salomaa [10], and so is the new one we introduce to define ω-finitary semirings. The main part of this paper is concerned with an overview of these and some related notions. We give a complete discussion of their interconnections, including new characterizations and counterexamples. Then we deal with topological summability as was introduced by Krob [6]. We connect this concept with the notions considered previously. (This was partially done by Krob himself.) Some problems remain open here. Finally, we consider finite resp. idempotent semirings. In these special cases, some notions happen to coincide.

The paper is organized as follows. Section 1 contains the basic definitions. (Apart from the basic topological concepts used in Sections 6 and 7, the paper is self-contained.) The subsequent sections then introduce step by step the various notions and discuss their interconnections. A first summary is given in Section 5. Topological semirings are discussed in Section 6. The final section considers finite resp. idempotent semirings. The information is always summarized in diagrams showing all implications and counterexamples.

We finally remark that all our semirings have both an absorbing zero and a unity. This is justified by the fact that we have automata theory in mind. (For example, the very important notion of a normalized automaton cannot be defined without the presence of the two neutral elements, cf. [10].)

1. Preliminaries

In this paper, a *semiring* is defined as an algebra $(A, +, \cdot, 0, 1)$, where $(A, +, 0)$ is a commutative monoid, $(A, \cdot, 1)$ is a monoid, and the operations are connected by the following laws:

$$(a+b)c = ac+bc, c(a+b) = ca+cb\ , \tag{1.1}$$

$$0 \cdot a = a \cdot 0 = 0\ . \tag{1.2}$$

(Because of (1.2), the zero 0 is called absorbing.) In the following we only write $(A, +, \cdot)$ for such a semiring. A semiring is called *idempotent* if it satisfies $1+1=1$. A *complete* semiring $(A, +, \cdot, \Sigma)$ is a semiring $(A, +, \cdot)$ where, for every index set I and every family $(x_i,\ i \in I)$ of elements of A, the *sum* $\sum_{i \in I} x_i$ is defined and satisfies the following conditions:

$$\sum_{i \in \emptyset} x_i = 0\ , \qquad \sum_{i \in \{1\}} x_i = x_1\ , \qquad \sum_{i \in \{1,2\}} x_i = x_1 + x_2\ , \tag{1.3}$$

$$\sum_{j \in J} \Big(\sum_{i \in I_j} x_i \Big) = \sum_{i \in I} x_i\ , \quad \text{if } I = \bigcup_{j \in J} I_j \text{ is a partition}\ , \tag{1.4}$$

$$\sum_{i \in I} z x_i = z \Big(\sum_{i \in I} x_i \Big)\ , \qquad \sum_{i \in I} x_i z = \Big(\sum_{i \in I} x_i \Big) z\ . \tag{1.5}$$

Informally speaking, $\sum$ extends the addition and satisfies the infinite laws of associativity and distributivity. Observe that, by (1.3) and (1.4),

$$\sum_{i \in I} x_i = \sum_{j \in J} \sum_{i \in \{\varphi(j)\}} x_i = \sum_{j \in J} x_{\varphi(j)} \tag{1.6}$$

for every bijection $\varphi : J \to I$.

Important examples of complete semirings are $\mathbb{B}$, $\mathbb{N}^{(\infty)}$ and $\mathbb{R}_+^{(\infty)}$, where $\mathbb{B} = \{0, 1\}$ with $1+1=1$ is the Boolean semiring and $\mathbb{N}^{(\infty)}$ resp. $\mathbb{R}_+^{(\infty)}$ are obtained from $\mathbb{N}$ resp $\mathbb{R}_+$ (the nonnegative integers resp. reals) by adjoining an element ∞ satisfying $a + \infty = \infty + a = a \cdot \infty = \infty \cdot a = \infty$ for $a \neq 0$. (Observe that $\infty \cdot 0 = 0 \cdot \infty = 0$ by (1.2).) The infinite summation is then defined by

$$\sum_{i \in I} x_i = \sup_{F \subseteq I,\, F \text{ finite}} \sum_{i \in F} x_i\ ,$$

where sup is taken with respect to the usual total order (in $\mathbb{B}$, we have $0 \leq 1$).

In a complete semiring, the sums $a^* = \sum_{i \in \mathbb{N}} a^i$ resp. $a^+ = \sum_{i \in \mathbb{N} \setminus \{0\}} a^i$ are called the *star* resp. the *quasiinverse* of $a \in A$.

If A is a complete semiring and $\sum$ is a finite or countable alphabet, then the semiring $(A\langle\langle \sum^* \rangle\rangle, +, \cdot)$ of formal power-series over the free monoid generated by $\sum$ is turned into a complete semiring by transferring the infinite summation of A pointwise to $A\langle\langle \sum^* \rangle\rangle$. A similar statement holds for the semiring of matrices $A^{J \times J}$ for an arbitrary index set J.

For any A, $A^{\mathbb{N}}$ denotes the set of sequences of elements of A. We will write $\alpha = (\alpha(n))$ for $\alpha \in A^{\mathbb{N}}$. Moreover, we will use the notations $(a_0, a_1, \ldots)$, $(a_n,\ n \in \mathbb{N})$, or simply (a_n) for sequences in $A^{\mathbb{N}}$.

A set of *convergent sequences* is a set $D \subseteq A^{\mathbb{N}}$ satisfying the following closure properties for α, α_1, $\alpha_2 \in D$, $c \in A$ (cf. [10]):

(1) $\eta = (1, 1, \ldots) \in D$,
(2) $\alpha_1 + \alpha_2 \in D$,
(3) $c\alpha, \alpha c \in D$,
(4) $\alpha_c = (c, \alpha(0), \alpha(1), \ldots) \in D$.

A *limit function* is a mapping $\lim : D \to A$ satisfying
(1) $\lim \eta = 1$,
(2) $\lim(\alpha_1 + \alpha_2) = \lim \alpha_1 + \lim \alpha_2$,
(3) $\lim c\alpha = c \lim \alpha$, $\lim \alpha c = (\lim \alpha)c$,
(4) $\lim \alpha_c = \lim \alpha$.

The least set satisfying (1)–(4) is the set D_d of ultimately constant sequences. The only possible limit function is denoted by $\lim_d$. A *notion of convergence* is a pair $(D, \lim)$. $(D_d, \lim_d)$ is called the *discrete convergence.* If A is also a partially ordered set (with order $\leq$), then any lim is called *compatible* with $\leq$ if $a_n \leq b_n$ for all $n \in \mathbb{N}$ (and two convergent sequences (a_n) and (b_n)) implies $\lim a_n \leq \lim b_n$. The discrete convergence is compatible with any order.

2. Basic Notions

All notions in this section are defined for semirings $(A, +, \cdot)$. Zerosumfree semirings were introduced by Kuich and Salomaa [10]. As the name indicates, the notion is defined by the following condition:

$$x + y = 0 \quad \text{implies} \quad x = y = 0 \ . \tag{2.1}$$

The next result is easily obtained.

Proposition 2.1. *Every complete semiring is zerosumfree.*

Proof. Observe first that, for all I,

$$\sum_{i \in I} 0 = \sum_{i \in I} 0 \cdot 1 = 0 \cdot \sum_{i \in I} 1 = 0 \ . \tag{2.2}$$

Assume now that $x + y = 0$. Then by (1.6),

$$0 = \sum_{i \in \mathbb{N}} (x + y) = x + \sum_{i \in \mathbb{N} \setminus \{0\}} x + \sum_{i \in \mathbb{N}} y = x + \sum_{i \in \mathbb{N}} (x + y) = x \ . \qquad \blacksquare$$

Equation (2.2) will be used in the sequel without further mention.

A semiring $(A, +, \cdot)$ together with a partial order $\leq$ on A is called here a *partially ordered (p.o.)* semiring $(A, +, \cdot, \leq)$ if the following conditions hold for all $a, b, c \in A$.

$$0 \leq a \quad \text{and} \tag{2.3}$$

$$a \leq b \quad \text{implies} \quad a + c \leq b + c, ac \leq bc, ca \leq cb \ . \tag{2.4}$$

From a more general point of view, a p.o. semiring as defined here should be called *positively* partially ordered, which we leave out for brevity. (See [14] for an overview and further references.) In the following, we only consider partial orders on semirings that make them p.o. in the above sense. The next result is well-known and will be very useful in the sequel.

Lemma 2.2. *Let $(A,+,\cdot,\leq)$ be a p.o. semiring. Then $a+x+y=a$ implies $a+x=a$ for all $a,x,y\in A$. Conversely, a semiring $(A,+,\cdot)$ satisfying the latter is a p.o. semiring $(A,+,\cdot,\sqsubseteq)$ if one defines $a\sqsubseteq b$ by $a+c=b$ for some $c\in A$.*

Proof. The first statement follows from $a\leq a+x\leq a+x+y=a$. Conversely, due to the assumption, $\sqsubseteq$ is a partial order on A which clearly satisfies (2.3) and (2.4). ∎

The partial order defined in Lemma 2.2 is called the *natural* order on $(A,+,\cdot)$. We will always denote it by $\sqsubseteq$. It is the weakest partial order on $(A,+,\cdot)$, i.e., $a\sqsubseteq b$ implies $a\leq b$ for any partial order $\leq$ on A. (This is a direct consequence of (2.3) and the definition of $\sqsubseteq$.) Clearly, the usual order on the semirings $\mathbb{B}$, $\mathbb{N}^{(\infty)}$ and $\mathbb{R}_+^{(\infty)}$ is their natural order. As we show next, a semiring may actually have more than one partial order.

Example 2.1. Consider the (complete) semiring $A=\mathbb{R}_+^{(\infty)}$ and, for $r\geq 1$, the (complete) subsemiring $A_r=\mathbb{N}\cup[r,\infty]$. Then A_r has continuously many partial orders if $r>1$.

Proof. Clearly, A_r is (totally) ordered by the restriction $\leq$ of the usual order on A. Hence, it is also naturally ordered by Lemma 2.2. We denote the corresponding order by $\sqsubseteq_r$. Since, for $1\leq s\leq r$, A_r is also a subsemiring of A_s, $\sqsubseteq_s$ is a partial order on A_r. This proves the claim. ∎

Every partially ordered semiring is zerosumfree, since $x+y=0$ implies $0\leq x\leq x+y=0$. Despite Proposition 2.1, a complete semiring is not necessarily partially ordered.

Example 2.2 (cf. [7]). Consider $A=\mathbb{Z}_2=\{\overline{0},\overline{1}\}$, the field of the residue classes of the integers (modulo 2). Adjoin a new zero element 0 according to (1.2). Then adjoin an element ∞ as described in Section 1 and define $\sum'$ on $A_1=\mathbb{Z}_2\cup\{0,\infty\}$ by

$$\sum_{i\in I}{}' a_i=\begin{cases}\sum\limits_{a_i\neq 0} a_i & \text{if} \quad \{i\mid a_i\neq 0\} \text{ is finite,}\\ \infty & \text{otherwise}\,.\end{cases} \tag{2.5}$$

Then $(A_1,+,\cdot,\sum')$ is complete, but not partially ordered. We have $\overline{0}+\overline{1}+\overline{1}=\overline{0}$, but $\overline{0}+\overline{1}=\overline{1}\neq\overline{0}$, contradicting Lemma 2.2. (In fact, $\mathbb{Z}_2$ may be replaced by any ring with unity if the first equation is generalized to $\overline{0}+\overline{1}+(-\overline{1})=\overline{0}$.) ∎

We finally remark that an idempotent semiring always has exactly one partial order, defined by $a\leq b$ iff $a+b=b$ (cf. [10], Exercise 5.4). Clearly $\leq$ coincides with the natural order.

3. d-Complete Semirings

We now start the discussion of the interconnection between limit notions and infinite sums. Recall that the discrete convergence $(D_d,\lim_d)$ can be defined in any semiring and is the "smallest" convergence. In a complete semiring, one may then ask whether $\sum$ is compatible with $\lim_d$. We call a complete semiring *d-complete* (cf. Goldstern [4]), if

$$\sum_{0\leq i\leq n} a_i=a \quad \text{for all } n\geq n_0 \text{ implies } \sum_{i\in\mathbb{N}} a_i=a\,. \tag{3.1}$$

Thus A being d-complete means that if the sequence of finite partial sums of a family $(a_i, i \in \mathbb{N})$ is ultimately constant, then the sum of this family equals that constant value. Using the notation of Section 1, this can also be expressed as

$$\lim_d \left(\sum_{0 \le i \le n} a_i \right) = \sum_{i \in \mathbb{N}} a_i \ . \tag{3.2}$$

The next example shows that a complete semiring is not necessarily d-complete.

Example 3.1 (cf. [4]). Consider the semiring $\mathbb{B}^{(\infty)} = \mathbb{B} \cup \{\infty\}$ with the extensions of the operations as in Section 1. Define $\sum'$ by (2.5). Then $(\mathbb{B}^{(\infty)}, +, \cdot, \sum')$ is complete, but not d-complete. Choose $a_i = 1$, $i \in \mathbb{N}$. Then $\sum'_{0 \le i \le n} a_i = 1$ for all n, but $\sum'_{i \in \mathbb{N}} a_i = \infty$. ■

The definition (3.1) uses a special order of the elements a_i which contrasts the "infinite commutativity" of $\sum$. The following characterization is more symmetric. (Condition (iii) is the axiom (PO3) of [5, p. 61].)

Proposition 3.1. *Assume that A is a complete semiring. Then the following statements are equivalent:*

(i) *A is d-complete.*

(ii) *$a + x_i = a$ for all $i \in \mathbb{N}$ implies $a + \sum_{i \in \mathbb{N}} x_i = a$.*

(iii) *If $(a_i, i \in I)$ is an at most countable family and $\sum_{i \in F} a_i = \sum_{i \in E} a_i$ for some finite $E \subseteq I$ and all finite F with $E \subseteq F \subseteq I$, then $\sum_{i \in I} a_i = \sum_{i \in E} a_i$.*

(iv) *If $(a_i, i \in I)$ is an at most countable family and for every finite $E \subseteq I$ there is a finite $F(E)$ with $E \subseteq F(E) \subseteq I$ and $\sum_{i \in F(E)} a_i = a$ (for some fixed a), then $\sum_{i \in I} a_i = a$.*

Proof. (i)$\Longrightarrow$(iv). We may assume that $I = \mathbb{N}$. We define the sequence $(F_n, n \in \mathbb{N})$ of finite index sets by $F_0 = F(\{0\})$, $F_{n+1} = F(F_n \cup \{\min(\mathbb{N} \backslash F_n)\})$. Clearly, $F_n \subsetneq F_{n+1}$ for all $n \in \mathbb{N}$ and $\bigcup_{n \in \mathbb{N}} F_n = \mathbb{N}$. Next we define the family $(c_n, n \in \mathbb{N})$ by

$$c_0 = \sum_{i \in F_0} a_i, \quad c_{n+1} = \sum_{i \in F_{n+1} \backslash F_n} a_i, \quad n \ge 0 \ .$$

Then we have $\sum_{i \le n} c_i = \sum_{i \in F_n} a_i = a$. Thus we obtain

$$\sum_{i \in I} a_i = \sum_{i \in \bigcup_{n \in \mathbb{N}} F_n} a_i = \sum_{n \in \mathbb{N}} c_n = a$$

since A is d-complete.

The implication (iv) $\Longrightarrow$ (iii) is trivial.

(iii) $\Longrightarrow$ (ii). Suppose that $a + x_i = a$ for all $i \in \mathbb{N}$. We define the family $(a_i, i \in \mathbb{N})$ by $a_0 = a$, $a_{i+1} = x_i$, $i \geq 0$, and choose $E = \{0\}$. Then for all finite $F \supseteq E$, we have $\sum_{i \in F} a_i = \sum_{i \in E} a_i$ by induction on the cardinality of F. Thus we obtain

$$a + \sum_{i \in \mathbb{N}} x_i = \sum_{i \in \mathbb{N}} a_i = \sum_{i \in E} a_i = a \ .$$

(ii) $\Longrightarrow$ (i). Assume that $\sum_{i \leq n} a_i = a$ for $n \geq n_0$. We define $x_i = a_{n_0+i+1}$, $i \in \mathbb{N}$. Then $a + x_i = \sum_{j \leq n_0+i} a_j + a_{n_0+i+1} = \sum_{j \leq n_0+i+1} a_j = a$ for all i. Hence we obtain

$$\sum_{i \in \mathbb{N}} a_i = \sum_{i \leq n_0} a_i + \sum_{i > n_0} a_i = a + \sum_{i \in \mathbb{N}} x_i = a \ . \qquad \blacksquare$$

Finally, we establish the connection with the previous section.

Proposition 3.2 (Goldstern [4]). *Every d-complete semiring is partially ordered.*

Proof. Similar to that of Proposition 2.1, using Lemma 2.2. ■

As d-completeness is a weak assumption, it is rather "natural" to conjecture that a complete semiring is partially ordered. However, this is not the case in general, as has already been shown in Example 2.2. Moreover, the converse of Proposition 3.2 does not hold, as is shown by Example 3.1. The semiring $\mathbb{B}^{(\infty)}$ is idempotent and totally ordered by $0 \leq 1 \leq \infty$.

4. l-Complete Semirings

This section continues the discussion about the interrelation between infinite sums and limit notions. (As already pointed out in the introduction, we use the algebraic limit concept according to [10], for topological considerations see Section 6.) This time we want to *define* a notion of convergence $(D, \lim)$ by the use of the infinite summation $\sum$. (Note that this is somewhat converse to real or complex analysis, where (topological) limits are used to define (some) countably infinite sums.) A preliminary step in this direction has been taken in the previous section. In any d-complete semiring, the notion of discrete convergence may be introduced according to (3.2). Clearly, (3.2) makes sense also for families (a_i) that do not satisfy the precondition of (3.1). Hence we choose the set of convergent sequences to be

$$D = \Big\{ (b_n, n \in \mathbb{N}) \mid \text{there are } (a_i, i \in \mathbb{N}) \text{ and } n_0 \in \mathbb{N} \text{ with } b_n = \sum_{0 \leq i \leq n} a_i \text{ for } n \geq n_0 \Big\} \ . \tag{4.1}$$

The definition of $\lim : D \to A$ is now a direct generalization of (3.2):

$$\lim(b_n) = \sum_{i \in \mathbb{N}} a_i \ , \tag{4.2}$$

where the families (a_i) and (b_n) are from (4.1). As we want lim to be well-defined, we need an additional assumption on $\sum$. We call a complete semiring *l-complete* (cf. Kuich [8]), if

$$\sum_{0 \leq i \leq n} a_i = \sum_{0 \leq i \leq n} a'_i \text{ for all } n \geq n_0 \text{ implies } \sum_{i \in \mathbb{N}} a_i = \sum_{i \in \mathbb{N}} a'_i \ . \tag{4.3}$$

As (4.3) is a generalization of (3.1), every l-complete semiring is d-complete. (The converse is not true, see the end of this section.) Hence it is partially ordered by the natural order of $\sqsubseteq$. Then by definition of D and $\sqsubseteq$, D coincides with the set of all ultimately monotonically increasing sequences. This gives additional motivation for the definition of D. In any l-complete semiring, the pair $(D, \lim)$ defined above is called the *natural convergence.*

The interrelation between the limit function and the partial order will be investigated in more detail in the next section. Our next result gives a more symmetric characterization of l-complete semirings. Condition (ii) is a strengthened version of the axiom (PO4) in [5, p. 61].

Proposition 4.1. *The following two statements are equivalent for a complete semiring A.*

(i) *A is l-complete.*

(ii) *If $(a_i, i \in I)$ and $(b_i, i \in I)$ are at most countable families and for every finite $E \subseteq I$, there is a finite $F(E)$ with $E \subseteq F(E) \subseteq I$ and $\sum_{i\in F(E)} a_i = \sum_{i \in F(E)} b_i$, then $\sum_{i\in I} a_i = \sum_{i \in I} b_i$.*

Proof. Similar to that of Proposition 3.1. ■

Propositions 3.1 and 4.1 provide means to generalize the notions d-complete and l-complete to non-countable sums. The reader may want to verify that the semirings of Examples 4.1 and 5.2 satisfy Proposition 3.1 (ii)–(iv), resp. 4.1 (ii) for arbitrary index sets I.

We conclude this section with some examples. The semirings $\mathbb{B}$, $\mathbb{N}^{(\infty)}$ and $\mathbb{R}_+^{(\infty)}$ are l-complete. The semiring $\mathbb{B}^{(\infty)}$ of Example 3.1 is not l-complete, since it is not d-complete. The following is an example of a semiring that *is* d-complete, but not l-complete.

Example 4.1. Consider the complete semiring $A_1 = \mathbb{R}_+^{(\infty)} \times \mathbb{B}^{(\infty)}$ and its complete subsemiring $A_2 = A_1 \setminus \{(0,1), (0,\infty)\}$. Then a complete semiring A_3 is obtained from A_2 by identifying the elements $(\infty, 0)$, $(\infty, 1)$ and (∞, ∞). Observe that A_3 was constructed in such a way that the equation $(a,b)+(x,y) = (a,b)$ always has only the solution $x = y = 0$ for $a < \infty$, whereas each $(x,y) \in A_3$ is a solution for $a = \infty$. From this it follows that A_3 is d-complete by Proposition 3.1 (ii). Define now $c_0 = d_0 = (1,1)$, $c_i = (1/2^i, 0)$, $d_i = (1/2^i, 1)$, $i \geq 1$. Then

$$\sum_{0\leq i\leq n} c_i = \sum_{0 \leq i \leq n} d_i \text{ for all } n \in \mathbb{N}, \quad \text{but} \quad \sum_{i\in\mathbb{N}} c_i = (2,1) \neq (2,\infty) = \sum_{i \in \mathbb{N}} d_i \,.$$

This shows that A_3 is not l-complete. ■

5. (ω-)Finitary and (ω-)Continuous Semirings

As was already seen in Section 3, rather weak assumptions on the summation $\sum$ make a complete semiring partially ordered. In this section we investigate the interrelation between partial order and summation in more detail. *Finitary* semirings were introduced by Goldstern [4] as p.o. complete semirings $(A, +, \cdot, \leq, \sum)$ satisfying the condition

$$\sum_{i\in F} a_i \leq c \text{ for all finite } F \subseteq I \text{ implies } \sum_{i\in I} a_i \leq c \tag{5.1}$$

for families $(a_i, i \in I)$. Equivalently, every infinite sum is the supremum of all finite partial sums. (Thus $\sum$ is completely determined by $+$ and $\leq$.)

Examples of finitary semirings are $\mathbb{B}$, $\mathbb{N}^{(\infty)}$ and $\mathbb{R}_+^{(\infty)}$ (where $\leq$ is the usual total order). We want to point out that the notion depends heavily on the partial order under consideration. Hence we will always speak of finitary semirings *with respect to an order* $\leq$. Then we have the following

FACT 5.1. *Assume that A is finitary w.r.t. some partial order $\leq$. Then A need not be finitary w.r.t. the natural order $\sqsubseteq$.*

As an example, consider the finitary semiring $A = \mathbb{R}_+^{(\infty)} \times \mathbb{R}_+^{(\infty)}$ and its subsemiring $A_1 = A\backslash\{0\} \times (0,\infty]$. A_1 is complete under the restricted summation and hence finitary w.r.t. the pointwise partial order $\leq$. Consider now the family $a_i = (1/2^i, 1/2^i)$, $i \in \mathbb{N}$, and $c = (2,3)$. Then

$$\sum_{i \in F} a_i + \left(\sum_{i \in \mathbb{N}\backslash F} a_i + (0,1) \right) = c \ ,$$

implying $\sum\limits_{i \in F} a_i \sqsubseteq c$. But

$$\sum_{i \in \mathbb{N}} a_i = (2,2) \not\sqsubseteq (2,3) \ ,$$

since $(0,1)$ is not in A_1. Hence A_1 is not finitary w.r.t. $\sqsubseteq$. ∎

In view of the previous sections, it makes sense to restrict attention to countable sums. Thus we call a p.o. complete semiring $(A, +, \cdot, \leq, \sum)$ *ω-finitary* (with respect to $\leq$) if (5.1) holds for all at most countable index sets I. Again, the notion depends on the particular order. (The finitary semiring used to prove Fact 5.1 is not even ω-finitary with respect to $\sqsubseteq$). Moreover, as one may expect, an ω-finitary semiring is not necessarily finitary.

Example 5.1 (cf. [4]). We rename the elements 1 and ∞ of the semiring $\mathbb{B}^{(\infty)}$ (Example 3.1) as $\aleph_0$ and $\aleph_1$ to indicate that the elements will now resemble cardinal numbers. (Addition and multiplication remain unchanged.) We define

$$\sum_{i \in I} a_i = \begin{cases} \aleph_1 & \text{if } \{i \mid a_i \neq 0\} \text{ is not countable or some } a_i = \aleph_1 \\ \sum\limits_{a_i \neq 0} a_i & \text{if } \{i \mid a_i \neq 0\} \text{ is finite} \\ \aleph_0 & \text{otherwise} . \end{cases}$$

It is easily verified that $A = \{0, \aleph_0, \aleph_1\}$ is ω-finitary with respect to $0 \leq \aleph_0 \leq \aleph_1$. Consider now the family $(a_i, i \in \mathbb{R})$ with $a_i = \aleph_0$ for all i. Then $\sum\limits_{i \in F} a_i = \aleph_0$ for all finite $F \neq \emptyset$, but $\sum\limits_{i \in \mathbb{R}} a_i = \aleph_1$. Thus A is not finitary w.r.t. $\leq$ (which is the only order, since A is idempotent). ∎

If one considers limits in a partially ordered semiring, it is reasonable to require that the limit function is compatible with the partial order (cf. [10], Section 5). The next result shows that ω-finitary semirings have exactly this property. (This was the reason for introducing the notion.)

Proposition 5.2. *Assume that A is partially ordered (by $\leq$) and complete. Then A is ω-finitary (w.r.t. $\leq$) iff A is l-complete and the natural convergence is compatible with the order.*

Proof. Assume first that A is ω-finitary and $\sum_{0\leq i\leq n} a_i \leq \sum_{0\leq i\leq n} b_i$ for all $n \in \mathbb{N}$. Then $\sum_{0\leq i\leq n} a_i \leq \sum_{0\leq i\leq n} b_i + \sum_{i\geq n+1} b_i = \sum_{i\in\mathbb{N}} b_i$ and hence $\sum_{i\in\mathbb{N}} a_i \leq \sum_{i\in\mathbb{N}} b_i$. By symmetry, this implies that A is l-complete. Then by the last inequality, lim is compatible with $\leq$. Conversely, assume that $\sum_{i\in F} a_i \leq c$ for all finite $F \subseteq I$ (and I countable; we may assume $I = \mathbb{N}$). We define the sequences (b_n) and (c_n) by $b_n = \sum_{0\leq i\leq n} a_i$, $c_n = c$, $n \in \mathbb{N}$. Then $b_n \leq c_n$ for all n and hence $\sum_{i\in I} a_i = \lim b_n \leq \lim c_n = c$. ■

The next example shows that, in general, an l-complete semiring is not ω-finitary (w.r.t. any order).

Example 5.2. Consider the complete semiring $\mathbb{Q}_+^{(\infty)}$, the nonnegative rational numbers with an element ∞ as described in Section 1 and define $\sum'$ by (2.5). Clearly, A is l-complete. However, A is not ω-finitary w.r.t. the usual order $\leq$. For example, $\sum'_{i\in F} 1/2^i \leq 2$ for all finite $F \subseteq \mathbb{N}$, but $\sum'_{i\in\mathbb{N}} 1/2^i = \infty$. Observe that $\leq$ coincides with the natural order and recall that $\sqsubseteq$ is always the weakest order in a semiring. As $\sqsubseteq$ is already a total order in our case, there cannot be any other order on A. Thus A is not ω-finitary w.r.t. any order. ■

The next two results deal with an ω-finitary semiring A. The order $\leq$ is assumed to be fixed, but need *not* be the natural order.

Proposition 5.3. *a^*b is the least solution of $y = ay + b$.*

Proof. In [10], the limit concept was introduced as a means to define a "star" operation for matrices and power-series in arbitrary semirings. It is defined w.r.t. a given notion $(D, \lim)$ of convergence. Then $a^* = \lim\Big(\sum_{0\leq i\leq n} a^i\Big)$, provided this limit exists. Our definition of the star in a complete semiring is a special case of the above one if the semiring is l-complete and $(D, \lim)$ is chosen to be the natural convergence (cf. [8], p. 214). The statement now follows from [10], Theorem 5.11, by our Proposition 5.2 and the remark preceding it. ■

This statement need not be true for a p.o. l-complete semiring $(A, +, \cdot, \leq, \sum)$. Consider $\mathbb{Q}_+^{(\infty)}$ and the equation $y = \frac{1}{2}y+1$. Then $y = \left(\frac{1}{2}\right)^* = \infty$ is a solution, but $y = 2$ is the minimal solution.

Proposition 5.4. *Denote the natural convergence by $(D, \lim)$. Assume that there is another notion of convergence $(D', \lim')$ that is also compatible with $\leq$. Then $\lim\alpha = \lim'\alpha$ for all sequences $\alpha \in D \cap D'$.*

Proof. By definition of D, there is an $n_0 \in \mathbb{N}$ and a sequence $\overline{\alpha}$ satisfying

$$\alpha(n) = \sum_{0\leq i\leq n} \overline{\alpha}(i) \tag{1}$$

for $n \geq n_0$. Set $a = \lim\alpha$, $a' = \lim'\alpha$. For arbitrary $b \in A$ and $j \geq 0$, define the ultimately constant sequence

$$\beta^{(b,j)} = (\alpha(0), \alpha(1), \ldots, \alpha(j-1), b, b, \ldots) \in D \cap D'\ .$$

By (1) we have $\alpha \leq \beta^{(a,n_0)}$ and hence

$$a' = \lim{}' \alpha \leq \lim{}' \beta^{(a,n_0)} = a \ . \tag{2}$$

To establish the reverse inclusion, we show that

$$\alpha(n) \leq a' \ \text{ for } \ n \geq n_0 \ . \tag{3}$$

Consider, for $j \in \mathbb{N}$, the sequence $\beta_j = \beta^{(\alpha(j),j)}$. Then $\beta_j \leq \alpha$ for $j \geq n_0$. This implies

$$\alpha(j) = \lim{}' \beta_j \leq \lim{}' \alpha = a' \ ,$$

which is (3). (Observe that j remains fixed in the above computation.) Condition (3) implies $\alpha \leq \beta^{(a',n_0)}$. Hence

$$a = \lim \alpha \leq \lim \beta^{(a',n_0)} = a' \ . \tag{4}$$

Now $a = a'$ by (2) and (4). ∎

The remainder of this section deals with the natural order. This is motivated by the fact that all semirings that are important in applications are finitary w.r.t. $\sqsubseteq$ (e.g. $\mathbb{B}$, $\mathbb{N}^{(\infty)}$, $\mathbb{R}_+^{(\infty)}$ and the matrix and power-series semirings derived from them, but also semirings occurring in connection with optimization problems, e.g. "shortest path", cf. [11] and the references given there). According to Sakarovitch [12] and Krob [7] we call these semirings *continuous.* Similarly, ω-continuous semirings are ω-finitary w.r.t. $\sqsubseteq$, cf. Kuich, [9] and the "d-continuous" semirings of [7]. Again, this notion is strictly weaker than continuity, cf. Example 5.1. The former authors do not consider the more general notion "finitary" at all, but as the example in the proof of Fact 5.1 shows, finitary semirings may occur as complete subsemirings of continuous ones.

As the natural order is completely determined by the addition, the same holds for the infinite summation in a continuous semiring by the remark immediately after (5.1). Clearly, a similar statement is true for countable sums and ω-continuity. As a consequence, we have the next result, which is again important for applications in automata theory.

Theorem 5.5. *Assume that A is a complete semiring, J an index set and $\sum$ an alphabet. Then A is continuous iff $A^{J\times J}$ (resp. $A\langle\langle\sum^*\rangle\rangle$) is, and in this case, summation in $A^{J\times J}$ (resp. $A\langle\langle\sum^*\rangle\rangle$) is the pointwise extension of the summation in A.*

Proof. As addition is defined pointwise in $A^{J\times J}$ and $A\langle\langle\sum^*\rangle\rangle$, so is the natural order. Now everything follows by the above remark and the fact that A is isomorphic to a subsemiring of $A^{J\times J}$ (resp. $A\langle\langle\sum^*\rangle\rangle$). ∎

Again, a similar statement holds for ω-continuous semirings and countable summation.

We think that restriction to continuous (or at least ω-continuous) semirings makes the axiomatization of automata theory more satisfactory. We want to illustrate this by an example.

Example 5.3. Consider the matrix

$$M = \begin{pmatrix} \frac{1}{2}\varepsilon & x \\ 0 & x \end{pmatrix} \in \left(\mathbb{R}_+^{(\infty)}\langle\langle\{x\}^*\rangle\rangle\right)^{2\times 2} \ ,$$

which can be viewed as the transition matrix of a finite automaton, cf. [10]. Transfer the infinite summation from the continuous semiring $\left(\mathbb{R}_+^{(\infty)}, +, \cdot, \sqsubseteq, \sum\right)$ pointwise to $\mathbb{R}_+^{(\infty)}\langle\langle\{x\}^*\rangle\rangle$ and then in turn to $\left(\mathbb{R}_+^{(\infty)}\langle\langle\{x\}^*\rangle\rangle\right)^{2\times 2}$. Then the latter semiring is also continuous and induction on n (to obtain M^n) yields

$$M^* = \begin{pmatrix} 2\varepsilon & 2x^+ \\ 0 & x^* \end{pmatrix} .$$

On the other hand, Theorem 2.5 of [8] allows to compute the star of any matrix $M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ over an l-complete semiring as $M^* = \begin{pmatrix} a^* & a^*bc^* \\ 0 & c^* \end{pmatrix}$.

Observe now that the nonnegative reals may also be equipped with the infinite summation according to (2.5) (cf. [6]). It is easily seen that $\mathbb{R}_+^{(\infty)}$ is l-complete also with respect to $\sum'$. If now power-series and matrices are summed pointwise according to $\sum'$, then application of the above mentioned result yields

$$M^* = \begin{pmatrix} \infty \cdot \varepsilon & \infty \cdot x^+ \\ 0 & x^* \end{pmatrix} . \qquad \blacksquare$$

With respect to this example, it is meaningful to restrict our attention to continuous semirings. (And as most sums occurring in automata theory are countable, restriction to ω-continuous ones will suffice, cf. [9]).

Continuous semirings were defined by a condition connecting summation and natural order. Surprisingly, it turns out that a characterization can be given that does not mention the partial order. This is the last result of the section.

Proposition 5.6. *Assume that A is a complete semiring. Then A is continuous (resp. ω-continuous) iff the following condition holds for all I (resp. for all at most countable I) and all $c \in A$:*

If for all finite $F \subseteq I$ there is a b_F such that $\sum_{i\in F} a_i + b_F = c$,

then there is also some b satisfying $\sum_{i\in I} a_i + b = c$. (5.7)

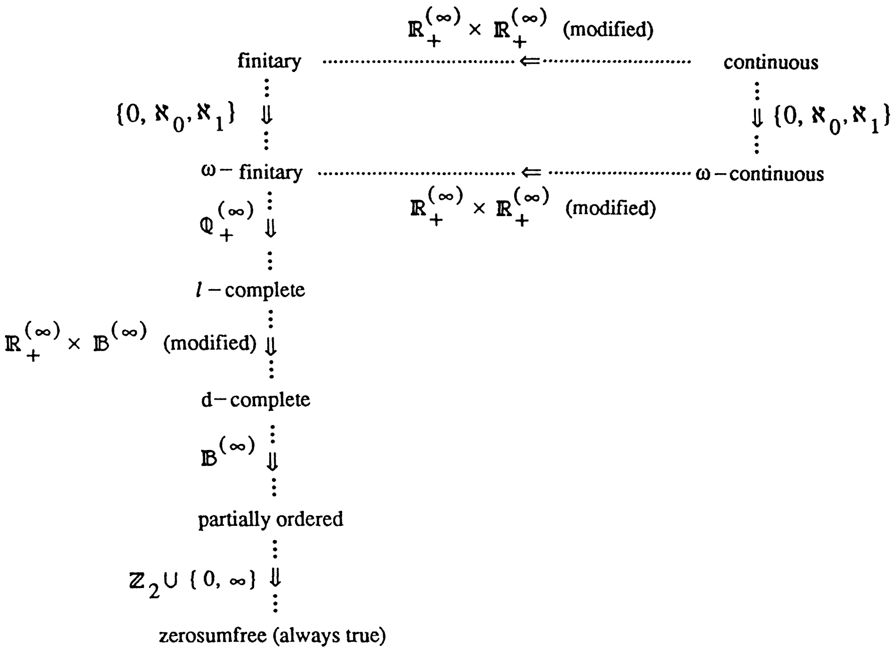
Proof. Observe that (5.7) is obtained from (5.1) by substituting the definition of the natural order. So we only have to show that every semiring satisfying (5.7) is naturally ordered. We use Lemma 2.2. Assume that $a + x + y = a$. Define $a_0 = a$, $a_{2i+1} = x$, $a_{2i+2} = y$, $i \geq 0$. Then the family $(a_i, i \in \mathbb{N})$ satisfies (5.7) with $c = a$, since by induction $a + n \cdot x + n \cdot y = a$ for all $n \in \mathbb{N}$. Hence

$$a = b + \sum_{i\in\mathbb{N}} a_i = b + \sum_{i\in\mathbb{N}} a_i + x = a + x ,$$

where the second equality follows as in Proposition 2.1. $\blacksquare$

The results of Sections 2–5 are summarized in Fig. 5.1. All implications are strict and are accompanied by a counterexample for the reverse direction.

Figure 5.1.



Properties of complete semirings

We invite the reader to verify that all our examples (except possibly the ones involving the reals) have minimal cardinality (see also Section 7).

6. Topological Complete Semirings

Every complete semiring A has a "natural" class of nets (in the topological sense). Consider an arbitrary family $(a_i, i \in I)$ of elements of A. Then $\{F \subseteq I \mid F \text{ is finite}\}$ is directed by set inclusion and thus $\left(\sum_{i \in F} a_i; F \subseteq I, F \text{ finite}\right)$ is a net. One may then ask whether $\sum_{i \in I} a_i$ is the limit of this net w.r.t. some T_2-topology. This was done by Krob [6] (cf. Def. II.4 and Def. III.2) who called a semiring $(A, +, \cdot, \mathcal{T})$ t-complete if $\mathcal{T}$ is a T_2-topology on A such that every net of the form described above has a limit and addition and multiplication are continuous. Every t-complete semiring is also a complete semiring $(A, +, \cdot, \sum)$ if $\sum_{i \in I} a_i$ is defined as the limit of $\left(\sum_{i \in F} a_i; F \subseteq I, F \text{ finite}\right)$, see [6], Proposition III.4. Thus for every neighbourhood V of $\sum_{i \in I} a_i$ there is a finite F_0 such that

$$\sum_{i \in F} a_i \in V \quad \text{for all finite } F \supseteq F_0 \,. \tag{6.1}$$

A reader familiar with topology may already have noticed that the characterizations of Propositions 3.1 and 4.1 reflect the above described view of sums as limits of nets. (For example, Proposition 3.1 (iii) is a statement on eventually constant nets, and Proposition 4.1 (ii) deals with nets that frequently coincide.) The following is now immediate by the uniqueness of limits in a T_2-space.

Proposition 6.1. *Every t-complete semiring is l-complete.* ∎

Thus every t-complete semiring is naturally ordered by Proposition 3.2, cf. also [6], Prop. II.9.

The above result cannot be strengthened, in general, see Example 6.1. The converse is also not true, as is seen by the following example.

Example 5.1 (continued). The semiring $A = \{0, \aleph_0, \aleph_1\}$ is ω-continuous, but not t-complete. This is again shown by the family $(a_i, i \in \mathbb{R})$ with $a_i = \aleph_0$ for all i. Assume that A is t-complete and consider an arbitrary neighbourhood V of $\aleph_1$. As $\sum_{i \in \mathbb{R}} a_i = \aleph_1$, but $\sum_{i \in F} a_i = \aleph_0$ for all non-empty, finite F, V must contain $\aleph_0$. This shows that the T_2-axiom is not satisfied. ∎

Krob [6] also considered "dt-complete" semirings, which are defined by restricting the nets defined above to at most countable index sets. (This means that only countable sums can be defined according to (6.1). Hence the notion does not really fit the framework of this paper.) Clearly, a dt-complete semiring also satisfies Proposition 4.1 (ii). By [7], Proposition V.1, the semiring of the above example *is* dt-complete. Seen from this point of view, the following example is more satisfactory.

Example 5.2 (continued). The l-complete semiring $\mathbb{Q}_+^{(\infty)}$ is not dt-complete. This can be seen by considering neighbourhoods of ∞. Assume that there is a T_2-topology on $\mathbb{Q}_+^{(\infty)}$ satisfying (6.1) for all at most countable index sets I. We show that addition is not continuous. Let a be an arbitrary fixed positive rational number. Then by (T_2), there is an open set V with $\infty \in V$, $a \notin V$. Recall that $\infty + \infty = \infty$. If the addition were continuous there would be open neighbourhoods U_1 and U_2 of ∞ with $U_1 + U_2 \subseteq V$. Assume furthermore that for all $b < a/2$, there is an $x(b)$ with $b < x(b) < a/2$ and $x(b) \notin U_1$. Then we define the sequences (c_n) and (d_n) by $c_0 = x(a/4)$, $c_{n+1} = x(c_n)$, $d_0 = c_0$, $d_{n+1} = c_{n+1} - c_n > 0$ for $n \geq 0$. Thus we have $\sum'_{i \in \mathbb{N}} d_n = \infty$, but $\sum'_{0 \leq i \leq n} d_i = c_n \notin U_1$ for all n, contradicting (6.1). Hence for some $b_0 < a/2$, $(b_0, a/2) \cap \mathbb{Q} \subseteq U_1$. Now if there is a $b_1 \in U_2$ with $a/2 < b_1 < a - b_0$, then $a = (a - b_1) + b_1 \in U_1 + U_2 \subseteq V$, contradicting $a \notin V$. But if there is no such b_1, then U_2 violates (6.1), as can be seen by an argument similar to the above one. ∎

The question arises if there are assumptions on a complete semiring that force it to be t-complete. The continuous semirings being promising candidates, only a partial result has been obtained so far. We call a semiring *naturally totally ordered* (n.t.o.), if the natural order is a total order. Observe that $\mathbb{B}$, $\mathbb{N}^{(\infty)}$ and $\mathbb{R}_+^{(\infty)}$ are n.t.o.

Theorem 6.2 (Krob, [7], Prop. V.1 and Cor. V.2). *Every naturally totally ordered continuous semiring is t-complete.* ∎

Again, the converse is not true. The following example combines ideas from [7] and [4]. It shows that in general a t-complete semiring does not satisfy any of the axioms of the previous sections that is stronger than "l-complete".

Example 6.1. Let $A = \mathbf{N} \cup \{\omega - n \mid n \in \mathbf{N}\}$ and denote $\omega - 0$ by ω. Addition and multiplication on $\mathbf{N}$ are defined as usual and are extended by

$$(\omega - n_1) + (\omega - n_2) = \omega, \quad (\omega - n_1) + n_2 = \begin{cases} \omega - (n_1 - n_2) & \text{if } n_1 > n_2 \\ \omega & \text{otherwise}, \end{cases}$$
$$(\omega - n_1) \cdot (\omega - n_2) = \omega, \quad (\omega - n_1) \cdot n_2 = n_2 \cdot (\omega - n_1) = \omega \text{ if } n_2 \geq 2 .$$

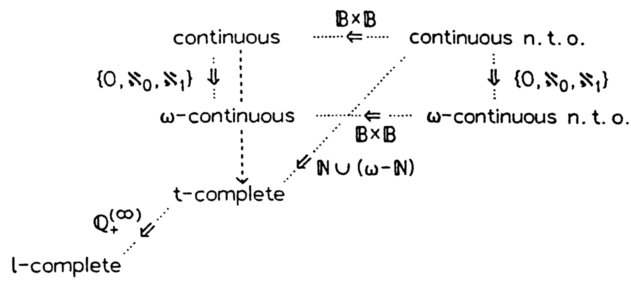
Here we have $n_1, n_2 \in \mathbf{N}$. We define infinite summation as in Example 2.2. Moreover, we introduce a topology on A by giving the base $\mathcal{B} = \{\{a\} \mid a \in A \backslash \{\omega\}\} \cup \{([n, \infty) \cap \mathbf{N}) \cap \{\omega\} \mid n \in \mathbf{N}\}$. Then A is t-complete, cf [7], p. 66 – the proof of the continuity of the addition easily transfers to the multiplication. Clearly, A is also totally ordered by the natural order. Observe furthermore that $\sum_{0 \leq i \leq n} i \sqsubseteq \omega - 1$ for all n, but $\sum_{i \in \mathbf{N}} i = \omega$. This shows that A is not ω-continuous. By the argument of Example 5.2, A is not ω-finitary w.r.t. any order. ∎

We remark that, as all our semirings, A is unitary. So the above example confirms a conjecture of Krob [7, p. 78].

Clearly, a continuous semiring is not necessarily n.t.o. A nice example is $A = \mathbf{B} \times \mathbf{B}$, which is isomorphic to the powerset of $\{x, y\}$ with union and intersection as operations. (This semiring is also t-complete, as will be seen in Section 7.)

We summarize the results of this section graphically. The main question (whether a continuous semiring is always topological) remains open. The broken arrow denotes the corresponding conjecture posed in [7, p. 61].

Figure 6.1.



Some more properties of complete semirings

7. Idempotent and Finite Semirings

Looking at the various counterexamples of Figures 5.1 and 6.1, a lot of them are "small" (i.e. contain three or four elements only) and, moreover, idempotent, whereas some others involve the real numbers. In this section we show that at least the latter cannot be replaced by "very simple" ones, i.e. some notions that are different in the general case turn out to be equivalent for finite or idempotent semirings. Moreover, considering idempotent semirings is motivated also by applications (e.g. formal language theory, cf. [10]).

The first result shows that d-completeness is a strong notion under our additional assumptions.

Proposition 7.1. *Let A be d-complete. If A is finite or idempotent, then A is ω-finitary w.r.t. every partial order.*

Proof. If A is finite, the statement follows from the observation that every ultimately monotonic sequence is ultimately constant. If A is idempotent, d-completeness implies that $\sum_{i\in\mathbb{N}} a = a$ for all $a \in A$. Recall that the only possible order is defined by $a \leq b$ iff $a + b = b$. If $\sum_{i\in F} a_i \leq c$ for all finite $F \subseteq I$ and I countable, then

$$\left(\sum_{i\in I} a_i\right) + c = \sum_{i\in I} a_i + \sum_{i\in I} c = \sum_{i\in I}(a_i + c) = \sum_{i\in I} c = c\,. \qquad \blacksquare$$

Remark. The condition $\sum_{i\in\mathbb{N}} a = a$ for all $a \in A$ is called "Axiom (ID)" by Hebisch [5]. It implies that A is idempotent ([5], Bemerkung 11.4). Together with the above proof this shows that (ID) implies A being ω-continuous, which extends [5], Satz 11.7.

Next, we deal with finitary semirings. As an idempotent semiring has exactly one partial order, the notions "finitary" and "continuous" coincide under this assumption. A similar result holds for finite semirings.

Proposition 7.2. *Assume that A is finite and finitary w.r.t. some partial order $\leq$. Then A is finitary w.r.t. any partial order (and hence continuous).*

Proof. Consider an arbitrary family $(a_i, i \in I)$. At first one can show that the set $X = \left\{\sum_{i\in F} a_i \mid F \subseteq I, F \text{ finite}\right\}$ has a maximum $\max_{\leq'} X$ w.r.t. each partial order $\leq'$ on A, and that each of these coincides with $\max_{\sqsubseteq} X$. Since A is finitary w.r.t. $\leq$, we have $\sum_{i\in I} a_i = \max X$, where the maximum is taken w.r.t. $\leq$ and hence w.r.t. any partial order $\leq'$ on A. $\blacksquare$

We finally consider t-complete semirings. This notion proves to be very strong under our assumptions, too. (Recall that every semiring of this kind is partially ordered.)

Proposition 7.3. *Let A be t-complete. If A is finite or idempotent, then A is finitary w.r.t. any partial order.*

Proof. If A is idempotent, the net $\left(\sum_{a\in F} a \mid F \subseteq I, F \text{ finite}\right)$ converges to a. This yields $\sum_{i\in I} a = a$ for all $a \in A$ and each I, and one can go on in the pattern of the proof of Proposition 7.1. If A is finite, recall that $X = \left\{\sum_{i\in F} a_i \mid F \subseteq I, F \text{ finite}\right\}$ has the same maximum $b = \max X$ w.r.t. any partial order. Since the net corresponding to X converges to b, we obtain $\sum_{i\in I} a_i = \max X$ which completes the proof. $\blacksquare$

The converse is also true for finite semirings, but the question remains open for idempotent ones (see also below, Theorem 7.5).

Proposition 7.4. *Every finite finitary semiring is t-complete.*

Proof. Consider the discrete topology. By the proof of Proposition 7.2, this topology satisfies (6.1). Clearly the semiring operations are continuous. ∎

Again we summarize our results in figures. All questions are settled for finite semirings. As regards the idempotent ones, it remains open, whether every continuous semiring is t-complete.

Figure 7.1.

$$
\begin{array}{ccc}
\text{finitary} \Leftrightarrow \text{continuous} \Leftrightarrow \text{t-complete} & \overset{\mathbb{B}\times\mathbb{B}}{\Leftarrow} & \text{continuous n.t.o.} \\
\{0,\aleph_0,\aleph_1\}\ \Downarrow & & \Downarrow\ \{0,\aleph_0,\aleph_1\} \\
\omega\text{-finitary} \Leftrightarrow \omega\text{-continuous} \Leftrightarrow \text{l-complete} \Leftrightarrow \text{d-complete} & \underset{\mathbb{B}\times\mathbb{B}}{\Leftarrow} & \omega\text{-continuous n.t.o.} \\
\mathbb{B}^{(\infty)}\ \Downarrow & & \\
\text{partially ordered} & & \\
\mathbb{Z}_2\cup\{0,\infty\}\ \Downarrow & & \\
\text{zerosumfree (always true)} & &
\end{array}
$$

Properties of finite, complete semirings

Figure 7.2.

$$
\begin{array}{ccc}
\text{finitary} \Leftrightarrow \text{continuous} \underset{\Leftarrow}{\dashrightarrow} \text{t-complete} & \overset{\mathbb{B}\times\mathbb{B}}{\Leftarrow} & \text{continuous n.t.o.} \\
\{0,\aleph_0,\aleph_1\}\ \Downarrow & & \Downarrow\ \{0,\aleph_0,\aleph_1\} \\
\omega\text{-finitary} \Leftrightarrow \omega\text{-continuous} \Leftrightarrow \text{l-complete} \Leftrightarrow \text{d-complete} & \underset{\mathbb{B}\times\mathbb{B}}{\Leftarrow} & \omega\text{-continuous n.t.o.} \\
\mathbb{B}^{(\infty)}\ \Downarrow & & \\
\text{partially ordered, zerosumfree (always true)} & &
\end{array}
$$

Properties of idempotent, complete semirings

Remark. The results about idempotent semirings are partially due to Hebisch [5]. Krob [7, p. 77], has a remark on finite complete monoids (cf. below). It claims that for such a monoid $(M,+,\sum)$ the properties "continuous", "d-continuous" (a notion similar to our "ω-continuous"), "t-complete", and "dt-complete" are equivalent. This would imply the same statement for finite complete semirings, which our Example 5.1 shows to be incorrect.

We conclude our paper with a partial result supporting Krob's conjecture mentioned above. All notions used in this paper transfer to (additively written) commutative monoids by omitting the axioms that involve multiplication. (This notion is called "complete monoid" in [6].) Idempotence of a monoid M means $a + a = a$ for all $a \in M$. With these definitions we can give the following characterization.

Theorem 7.5. *Let $(M, +, \sum)$ be a complete idempotent monoid. Then M is t-complete iff it is continuous.*

Proof. By Proposition 7.3, we only have to construct the topology for a continuous monoid. Define, for $c \in M$, the sets $U_c = \{x \in M \mid x \leq c\}$ and $V_c = M \setminus U_c$. Note that in general $V_c \neq \{x \in M \mid x > c\}$, since the order is only a partial one. Now a subbase for our topology is given by $\{U_c \mid c \in M\} \cup \{V_c \mid c \in M\}$. The topology satisfies (T_2), since if $a \neq b$, then for example $a \not\leq b$ and $a \in V_b$, $b \in U_b$. By continuity, (6.1) is satisfied for U_c and V_c. Clearly, (6.1) transfers then to arbitrary open sets. It only remains to show that addition is continuous. If $a + b \in U_c$, then $U_a + U_b \subseteq U_c$. If $a + b \in V_d$, then $a \in V_d$ or $b \in V_d$, since otherwise $a + b \leq d + d = d$. (The idempotence of M is used only here!) Now if $a \in V_d$, then $V_d + A \subseteq V_d$, and V_d resp. A are neighbourhoods of a and b, respectively. ∎

The topologies used by Krob ([7], Prop. III.3 and Cor. III.4) to prove Theorem 6.2 are finer than ours when transferred to *partially* ordered semirings. (This follows from [7], Proposition III.2 and the fact that M is continuous.) Thus it is possible to show the continuity of the multiplication. On the other hand, the topologies are too fine in the general case, since (6.1) may be violated.

The above characterization shows that as regards monoids, the broken arrow in Fig. 7.2 may be replaced by an equivalence sign. Thus the discussion is complete in this case.

We remark that some rather strange phenomena occur when dealing with complete monoids instead of semirings. For example, the very "natural" equation (2.2) does not hold in general, and also the zerosumfreeness in Fig. 7.2 then follows from the idempotence, but not from completeness axioms. See [6] for details and examples.

Conclusion

We gave a systematic overview of various axioms for complete semirings. It turned out that any of these forces the semiring to be partially ordered, which appears thus to be a basic feature of complete semirings. Moreover, in most cases, both a "symmetric" definition (reflecting the infinite commutativity of the summation) and a characterization using sequences were given. This shows the notions to be "sound". The main open question is whether a continuous semiring is always topologically complete.

References

[1] Conway, J. H., "Regular Algebra and Finite Machines", Chapman and Hall, 1971.

[2] Eilenberg, S., "Automata, Languages and Machines", Vol. A., Academic Press, 1974.

[3] Fuchs, L., "Partially Ordered Algebraic Systems", Pergammon Press, Oxford, 1963.

[4] Goldstern, M., Vervollständigung von Halbringen, Diplomarbeit, TU Wien, 1986.

[5] Hebisch, U., Zur algebraischen Theorie unendlicher Summen in Halbgruppen und Halbringen, Habilitationsschrift, TU Clausthal, 1990.

[6] Krob. D., *Monoides et semi-anneaux complets*, Semigroup Forum **36** (1987), 323–339.

[7] Krob. D., *Monoides et semi-anneaux continus*, Semigroup Forum **37** (1988), 59–78.

[8] Kuich, W., *The Kleene and the Parihk Theorem in Complete Semigrings*, Lecture Notes in Comp. Sci. **267** (1987), 212–225.

[9] Kuich, W., *Automata and Languages Generalized to ω-Continuous Semirings*, Theoretical Comput. Sci. **79** (1991), 137–150.

[10] Kuich, W. and S. Salomaa, "Semirings, Automata and Languages", Springer, 1986.

[11] Pan, V. and J. Reif, *Fast and Efficient Solution to Path Algebra Problems*, J. Comp. Sys. Sci. **38** (1989), 494–510.

[12] Sakarovitch, J., *Kleene's Theorem Revisited*, Lect. Notes in Comp. Sci. **281** (1987), 39–50.

[13] Weinert, H. J., *Generalized Semialgebras over Semirings*, Lect. Notes in Math. **1320** (1988), 380–416.

[13] Weinert, H. J., *Partially ordered semirings and semigroups*, in: Algebra and Order, Proceedings of the First International Symposium on Ordered Algebraic Structures, Luminy-Marseilles, June 1984, Heldermann, Berlin 1986.

Alcatel Austria-Elin
Forschungszentrum
Ruthnergasse 1–7
A-1210 Wien
Austria

Received May 8, 1990
and in final form December 5, 1991

Semigroup Forum Vol. 45 (1992) 166–182

RESEARCH ARTICLE

Bilateral Semidirect Products of Transformation Semigroups

M. Kunze

Communicated by B. M. Schein

Given a finite transformation semigroup (X,S), various types of decompositions may reveal different aspects of (X,S). The influence of algebraic properties of S on the action induced by S on X is often approached by using the wreath product and investigating the Krohn-Rhodes decomposition of S, cf. [4]. On the other hand, for studying the structure of the mappings $\alpha : X \to X$ in S, decompositions of α into a product of idempotents have been considered, cf. [5]. In this paper we mainly confine ourselves to aperiodic transformation semigroups (X,S) and we want to investigate a strong kind of decompositions of (X,S) interrelating both structure questions of S with possible compositions of its elements out of idempotents. More precisely, we are looking for a certain product (X,H) of suitable subsemilattice actions of S on X such that (X,S) is a homomorphic image of a transformation subsemigroup of (X,H). H is to be obtained by iterating the operation of forming bilateral semidirect products in general. The case where the bilateral semidirect product specializes to a semidirect or reverse semidirect product is, of course, most favorable. A comparison of the cardinalities $|H|$ and $|S|$ is significant for efficiency, an objective which makes semidirect products preferable over wreath products.

The semigroup $\mathrm{End}(X,\leq)$ of isotonic mappings of a chain (i.e., order-preserving transformations of a linear order) is probably one of the most natural examples of a general bilateral semidirect product. Its construction in Section 3 is the central part of this paper, and it also appears as a factor of the bilateral semidirect decomposition of the full transformation semigroup given in Section 4. For finite semigroup varieties $\mathbf{V}$ our approach suggests looking for such a set of generators that every semigroup in $\mathbf{V}$ is covered by some generator without having to form direct products of them again. Along these lines, results about definite semigroups, $\mathcal{R}$-trivial, $\mathcal{J}$-trivial, and locally testable semigroups are discussed.

1. Semidirect and Reverse Semidirect Products

Let $(S,\cdot)$, $(T,\cdot)$ be semigroups, $\varphi : S \to \mathcal{T}(T)$ be a homomorphism and $\delta : T \to \mathcal{T}(S)$ be an anti-homomorphism into the full transformation semigroup on T, resp. S. For $s \in S$ and $t \in T$, denote the operation of $\varphi(s)$ on T by $t \mapsto t^{\varphi(s)}$ and the operation of $\delta(t)$ on S by $s \mapsto \delta_t(s)$. Suppose that the following two conditions hold true for every s, s_1, $s_2 \in S$ and t, t_1, $t_2 \in T$:

$$(t_1 \cdot t_2)^{\varphi(s)} = t_1^{\varphi(\delta_{t_2}(s))} \cdot t_2^{\varphi(s)} \qquad \text{[Sequential Processing Rule]}$$

$$\delta_t(s_1 \cdot s_2) \;\; = \delta_t(s_1) \cdot \delta_{t^{\varphi(s_1)}}(s_2) \qquad \text{[Serial Composition Rule].}$$

Then $(S \times T, \circ)$ is a semigroup [9] with respect to the following multiplication:

$$(s_1,t_1) \circ (s_2,t_2) = \left(s_1 \cdot \delta_{t_1}(s_2), t_1^{\varphi(s_2)} \cdot t_2\right) \; .$$

Definition. This semigroup $(S\ {}_{\delta}\times_{\varphi} T, \circ)$ is called the *bilateral semidirect product* of S, T with respect to δ, φ. If $\varphi(S) = \{id_T\}$ we have a *semidirect product* $S \times_{\delta} T$, if $\delta(T) = \{id_S\}$ we have a *reverse semidirect product* $S\ {}_{\circledR}\!\times_{\varphi} T$, and if both operations are trivial we get a direct product of S and T.

Unless semigroups S, T come already equipped with a standard notation for their respective multiplications, we are free to use module notation like in Eilenberg's book [4]. This way the notation for the bilateral semidirect product becomes much less cumbersome:

$$(s_1, t_1) \circ (s_2, t_2) = (s_1 + t_1 * s_2, t_1 s_2 + t_2)$$

where the multiplication in S and T are written additively, juxtaposition denotes the right action of S on T, and $*$ denotes the left action of T on S.

For applications of bilateral semidirect products in semigroups and automata theory see [7], [9]. To get a feeling for this kind of product, we just mention that a group G is a bilateral semidirect product of two subgroups U_1, U_2 iff $G = U_1 U_2$ and $U_1 \cap U_2 = \{1\}$. This product is semidirect iff one of the subgroups is normal. In the case of groups, however, the notions of semidirect and reverse semidirect products coincide. The symmetric group on X is a bilateral semidirect product of the stabilizer of a given element and any cyclic subgroup generated by a cyclic permutation of order $|X|$.

Put another way, the above remarks say that in the context of groups bilateral semidirect products are nothing new: it is just the study of complementary subgroups. Classical examples from elementary geometry like the semidirect decomposition of the motion group of the Euclidean plane illustrate our basic idea, namely to decompose a given transformation semigroup into two subsemigroups. This is in contrast to decomposing into one substructure and one homomorphic image, as is done in ordinary (Schreier type) extension theory (cf. [9]) and more recently by Rhodes and Tilson in [11]. Consequently, the block product of Rhodes/Tilson relates to the kernel of a homomorphism as expressed by the Kernel Theorem [11] and Rhodes/Weil investigate maximal proper surmorphisms in [12]. In the associated double semidirect product [11], the second components multiply as in the direct product.

Definition. A transformation semigroup (X, H) is called a *bilateral semidirect [semidirect, reverse semidirect] product* of transformation semigroups (X, S) and (X, T), if

$$\mu : S\ {}_{\delta}\times_{\varphi} T \to \mathcal{T}(X), x\mu(s,t) = xst$$

is a homomorphism and $H \cong \mu(S\ {}_{\delta}\times_{\varphi} T)$ $\quad [\mu(S \times_{\delta} T), \mu(S\ {}_{\circledR}\!\times_{\varphi} T)]$.

Note that the last definition is a compromise, in order to avoid talking about non-faithful actions as a technical tool for decompositions of transformation semigroups. As a closure operation on classes of transformation semigroups the semidirect product is roughly speaking equivalent to the wreath product [8].

Our first example is the transformation semigroup of a shift register: Given an alphabet Σ, the set of states of a shift register of length k is Σ^k and its transition semigroup $D_{\Sigma,k}$ consists of the mappings induced by feeding input strings of lengths up to k.

Proposition 1. *The transformation semigroup $(\Sigma^k, D_{\Sigma,k})$ is covered by an iterated semidirect product (Σ^k, S_k) of k right zero semigroup actions [8]. Every definite semigroup [4, Ex. 3.7] is a homomorphic image of a subsemigroup of S_k for suitable k.*

As usual, (X, S) is covered by (Z, H) iff (X, S) is a homomorphic image of a transformation subsemigroup of (Z, H). For $k = 4$ and $|\Sigma| = 2$ we have $|D_{\Sigma,k}| = 30$ and $|S_k| = 120$ while a corresponding wreath product decomposition would involve 2^{15} elements. Since the action of $D_{\Sigma,k}$ on Σ^k is transitive, a decomposition into an iterated semidirect product of semilattice actions is impossible: reverse semidirect products have to be applied, too, e.g. to get right zero semigroups from semilattices.

A general procedure to obtain a decomposition of (X, S) into an iterated semidirect product of semilattice actions is based on the transitivity preorder $\mathrm{TR(S)} = (X, \leq)$ given by [13]:

$$x \leq y \iff x = y \text{ or } y = x\alpha \text{ for some } \alpha \in S .$$

Proposition 2. *The following statements are equivalent (cf. [4], [8]):*

(a) $\mathrm{TR(S)}$ *is a partial order.*

(b) S *is* $\mathcal{R}$*-trivial.*

(c) (X, S) *is covered by an iterated semidirect product of semilattice actions on* X.

In general, when the transitivity preorder is not anti-symmetric, (X, S) can still be covered by an iterated semidirect product $(X \cup \{0\}, H)$ of reverse semidirect products $(X \cup \{0\}, T_i \ {}_{\Theta}\!\times_{\varphi} S_i)$. The S_i are semilattices and the T_i are direct products of homomorphic images of S which induce transitive actions on equivalence classes of $\mathrm{TR(S)}$, cf. [8]. This is as much as we can expect to come out of investigations of the transitivity preorder of (X, S). In particular it remains to study transitive aperiodic transformation semigroups other than $(\Sigma^k, D_{\Sigma,k})$ mentioned above.

Examples. Let $(X, \leq)$ be a finite distributive lattice and

$$S_{\leq}^{-} = \{\alpha_{\langle a \rangle} : x \mapsto \inf(x, a) \mid a \in X\} \subseteq \mathcal{T}(X) ,$$
$$S_{\leq}^{+} = \{\alpha_{[a]} : x \mapsto \sup(x, a) \mid a \in X\} \subseteq \mathcal{T}(X) .$$

Both $S_{\leq}^{-}$ and $S_{\leq}^{+}$ are semilattices isomorphic to $(X, \leq)$. Because of the distributivity, the set theoretic product $S_{\leq}^{-} \cdot S_{\leq}^{+}$ is a subsemigroup of $\mathcal{T}(X)$. Furthermore it is a canonical homomorphic image of the the reverse semidirect product $S_{\leq}^{-} \ {}_{\Theta}\!\times_{\varphi} S_{\leq}^{+}$ with respect to

$$\varphi : S_{\leq}^{-} \to \mathrm{End}(S_{\leq}^{+}), \alpha_{[x]}^{\varphi(\alpha_{\langle a \rangle})} = \alpha_{[\inf(x,a)]} .$$

So the transformation semigroup $(X, S_{\leq}^{-} \cdot S_{\leq}^{+})$ is a reverse semidirect product of two semilattice actions. $S_{\leq}^{-} \cdot S_{\leq}^{+}$ is an $\mathcal{L}$-trivial band, in fact a right regular band [6]. It consists of the projections

$$\pi_A : x \mapsto \sup(\inf(x, a), b) = x\alpha_{\langle a \rangle}\alpha_{[b]}$$

where $A = [a, b]$ is an interval of $(X, \leq)$. We have $\mathrm{Fix}(\pi_A) = A$ and $\pi_A \circ \pi_B = \pi_{A\pi_B}$. The right zero semigroup consisting of the constant mappings on X is an ideal of $S_{\leq}^{-} \cdot S_{\leq}^{+}$. In case $(X, \leq)$ is a chain, i.e., a total order, then $S_{\leq}^{-} \ {}_{\Theta}\!\times_{\varphi} S_{\leq}^{+}$ is a subsemigroup of the bilateral semidirect product $S^{-} \ {}_{\delta}\!\times_{\varphi} S^{+}$ which will be constructed in Section 3. In the general case one has to consider direct products of those, as may be expected, because every distributive lattice is a sublattice of a chain product [1].

2. Monotone Order-Preserving Mappings of a Finite Chain

Let $(X,\leq)$ be a finite total order, i.e. a linear order or chain. $\mathrm{End}(X,\leq) = \{\alpha : X \to X \mid x \leq y \text{ implies } x\alpha \leq y\alpha \text{ for every } x,y \in X\}$ is the semigroup of order-preserving (or isotonic) mappings of $(X,\leq)$. We first look at its subsemigroup $\mathrm{End}(X,\leq)^- = \{\alpha \in \mathrm{End}(X,\leq) \mid x\alpha \leq x \text{ for every } x \in X\}$ of monotone decreasing mappings. $\mathrm{End}(X,\leq)^+$ is defined dually.

For an interval A_i of $(X,\leq)$ with least element i we consider the transformation

$$A_i^- : X \longrightarrow X, x \longmapsto \begin{cases} i & \text{if } x \in A_i \\ x & \text{otherwise} \end{cases}$$

and dually we define B_j^+ where j is the greatest element in the interval B_j. For every $i \in X$ we have subsemilattices

$$S_i^- = \{A_i^- \mid A_i = [i,a] \subseteq X\} \subseteq \mathrm{End}(X,\leq)^- \ ,$$

$$S_i^+ = \{B_i^+ \mid B_i = [b,i] \subseteq X\} \subseteq \mathrm{End}(X,\leq)^+ \ .$$

For notational convenience let us assume $X = \{0,1,\ldots,n\}$ with the usual order $0 < 1 < \ldots < n$. Then

$$\mathrm{End}(X,\leq)^- = S_0^- \circ S_1^- \circ \ \ldots \ \circ S_{n-1}^- \ ,$$

$$\mathrm{End}(X,\leq)^+ = S_n^+ \circ S_{n-1}^- \circ \ \ldots \ \circ S_1^+ \ ,$$

because any $\alpha \in \mathrm{End}(X,\leq)^-$ can be written as a composition of mappings

$$\alpha = A_0^- \circ \ \ldots \ \circ A_{n-1}^- \text{ where } A_i = [i, \max([0,i]\alpha^{-1})] \ ,$$

and similarly for $\alpha \in \mathrm{End}(X,\leq)^+$. The unique minimal set of idempotent generators of $\mathrm{End}(X,\leq)\backslash\{id\}$ given in [2] is $[0,1]^-, [1,2]^-, \ldots, [n-1,n]^-$, $[n-1,n]^+, \ldots, [0,1]^+$. For an n-tuple $A^- = (A_0^-, \ldots, A_{n-1}^-)$ and $a_i = \max(A_i)$, $i = 0,\ldots,n-1$, we often use the sequence notation $A^- = (a_0,\ldots,a_{n-1})^-$ or the function notation $A^-(i) = a_i$.

Proposition 3. *$\mathrm{End}(X,\leq)^-$ is covered by the following iterated semidirect product of $n = |X|-1$ semilattices:*

$$S(X,\leq)^- = \left((S_0^- \times_{\delta_1} S_1^-) \times_{\delta_2} S_2^- \ldots\right) \times_{\delta_{n-1}} S_{n-1}^- \ .$$

More precisely, $\mathrm{End}(X,\leq)^-$ *is the homomorphic image of* $S(X,\leq)^-$ *under the canonical homomorphism* $\mu : (A_0^-, \ldots, A_{n-1}^-) \mapsto A_0^- \circ \ldots \circ A_{n-1}^-$. *Alternatively,* $\mathrm{End}(X,\leq)^-$ *is also isomorphic to the subsemigroup*

$$S^- = \{A^- \in S(X,\leq)^- \mid A^-(0) \leq A^-(1) \leq \ldots \leq A^-(n-1)\}$$

of monotone sequences of $S(X,\leq)^-$.

Proof. For $i < j$ we form a semidirect product $S_i^- \times_{\delta_{ji}} S_j^-$ with respect to $\delta_{ji} : S_j^- \to \mathrm{End}(S_i^-)$ where

$$\delta_{ji}(B_j^-)[A_i^-] = \begin{cases} (A_i \cup B_j)_i^- & \text{if } A_i \cap B_j \neq \emptyset \ , \\ A_i^- & \text{otherwise} \ . \end{cases}$$

One observes $B_j^- \circ A_i^- = \delta_{ji}(B_j^-)[A_i^-] \circ B_j^-$ and verifies that δ_{ji} is a homomorphism and $\delta_{ji}(B_j^-) \in \mathrm{End}(S_i^-)$. In order to obtain an iterated semidirect product $S(X, \leq)^- = \left((S_0^- \times_{\delta_1} S_1^-) \times_{\delta_2} S_2^- \ldots\right) \times_{\delta_{n-1}} S_{n-1}^-$, we want to apply Lemma 1 of [8]. So the proof will be complete by showing

$$\delta_{ki}(C_k^-)\left[\delta_{ji}(B_j^-)[A_i^-]\right] = \delta_{ji}\left(\delta_{kj}(C_k^-)[B_j^-]\right)\left[\delta_{ki}(C_k^-)[A_i^-]\right]$$

for $i < j < k$ and $A_i^- \in S_i^-$, $B_j^- \in S_j^-$, $C_k^- \in S_k^-$.

<u>Case</u> $\max(A_i) < j$. Then both sides equal A_i^-.
<u>Case</u> $k \leq \max(A_i)$. Then both sides equal $(A_i \cup B_j \cup C_k)_i^-$.
<u>Case</u> $j \leq \max(A_i) < k$. If $\max(B_j) < k$, then $\max(A_i \cup B_j) < k$ and both sides equal $(A_i \cup B_j)_i^-$. If $k \leq \max(B_j)$, then both sides equal $(A_i \cup B_j \cup C_k)_i^-$.

Corollary. *Every $\mathcal{J}$-trivial finite transformation semigroup is covered by a direct product of transformation semigroups of the form $S(X, \leq)^-$ which are iterated semidirect products of semilattice actions on linear orders.*

3. Isotonic Mappings of a Chain

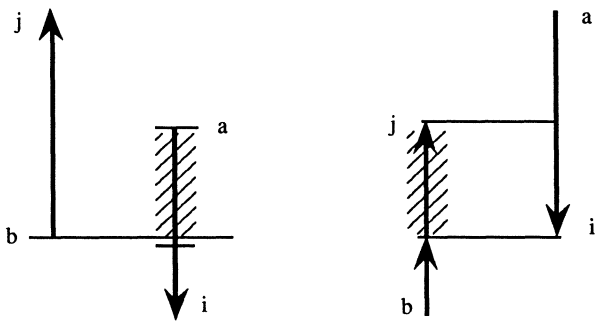
Again, let $(X, \leq)$ be a finite linear order. Without loss of generality assume $X = \{0, 1, \ldots, n\}$. Fundamental properties of $\mathrm{End}(X, \leq)$ are investigated by Aizenstat in [2] and related papers. Here we show:

Proposition 4. $\mathrm{End}(X, \leq)$ *is canonical homomorphic image of a bilateral semidirect product $S^- \,{}_\delta\!\times_\varphi S^+$ where S^-, S^+ are the subsemigroups of iterated semidirect products of semilattices given in Proposition 3.*

Proof. We are going to construct this bilateral semidirect product step by step. First we define δ and φ for single components $[i, a]^- \in S_i^-$ and $[b, j]^+ \in S_j^+$:

$$\delta_{[b,j]^+}([i, a]^-) = \begin{cases} [i, \max(i, b-1)]^- & \text{if } b \leq a < j \ , \\ [i, a]^- & \text{otherwise} \ , \end{cases}$$

$$[b, j]^{+\varphi([i,a]^-)} = \begin{cases} [\min(i, b), i]^+ & \text{if } j \in [i, a] \ , \\ [b, j]^+ & \text{otherwise} \ . \end{cases}$$



expansion: δ-transition takes place

contraction: φ-operation takes place

Observe

$$[b,j]^+ \circ [i,a]^- = \delta_{[b,j]^+}([i,a]^-) \circ [b,j]^{+\varphi([i,a]^-)} \ .$$

The case where the δ-transition is nontrivial describes an expansion type mapping unless the downwards action is void (i.e., $b \leq i$), the case where the φ-operation is nontrivial describes a contraction type mapping unless the upwards action is void (i.e., $i \leq b$).

For $A^- = (a_0, \ldots, a_{n-1})^- \in S^-$ and $B^+ = (b_n, \ldots, b_1)^+ \in S^+$ define $\delta_{B^+}(A^-)$ by giving its i-th component

$$\delta_{B^+}(A^-)[i] = \max(i, \min(a_i, b_{a_i+1} - 1))$$

where $b_{n+1} = n+1$ is assumed for the case $a_i = n$. Since $(a_0, \ldots, a_{n-1})^-$ is monotone increasing, we have

$$\delta_{B^+}(A^-)[i] = \delta_{[b_n,n]^+}\left(\ldots\ \left(\delta_{[b_1,1]^+}([i,a_i]^-)\right)\right) \ .$$

In particular, $\delta_{B^+}(A^-) \in S^-$. A good way of looking at it is to imagine the $\uparrow$-arrows and $\downarrow$-arrows as symbols of a semi-Thue system with alphabet $S_0^- \cup \ \ldots \ \cup S_{n-1}^- \cup S_n^+ \cup \ \ldots \ \cup S_1^+$ and derivation rules

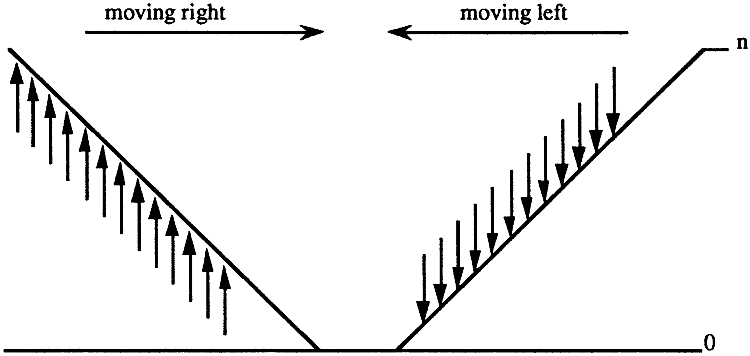
$$[b,j]^+[i,a]^- \longrightarrow \delta_{[b,j]^+}([i,a]^-)[b,j]^{+\varphi([i,a]^-)} \ .$$

Since $(b_n, \ldots, b_1)$ is sorted in decreasing order and $(a_0, \ldots, a_{n-1})$ is sorted in increasing order, the $\varphi([i,a]^-)$-operation on $[b,j]^+$ takes place before $[i,a]^-$ is affected by nontrivial $\delta_{[d,k]^+}$-transitions for $k \neq j$, and similarly for the $\delta_{[b,j]^+}$-transition of $[i,a]^-$. Therefore, the final result of the derivation is of the form

$$\delta_{B^+}(A^-)B^{+\varphi([0,a_0]^-)\ \ldots\ \varphi([n-1,a_{n-1}]^-)} \ .$$

The point is that the mapping induced on $(X, \leq)$ by B^+A^- is invariant under this derivation, so that we will get a homomorphism $\mu : S^- {}_\delta\!\times_\varphi S^+ \to \mathrm{End}(X, \leq)$,

$$\mu(A^-, B^+) = [0,a_0]^- \circ \ \ldots \ \circ [n-1, a_{n-1}]^- \circ [b_n, n]^+ \circ \ \ldots \ \circ [b_1, 1]^+ \ .$$



To see that μ is surjective, pick any $\alpha \in \mathrm{End}(X, \leq)$ and set

$$a_i := \begin{cases} \max([0,i]\alpha^{-1}) & \text{if } [0,i]\alpha^{-1} \cap [i,n] \neq \emptyset, \\ i & \text{otherwise}, \end{cases}$$

$$b_j := \begin{cases} \min([j,n]\alpha^{-1}) & \text{if } [j,n]\alpha^{-1} \cap [0,j] \neq \emptyset, \\ j & \text{otherwise}. \end{cases}$$

This representation is almost precisely the canonical form given in [2]. In terms of [2] we are working on the problem how to find the canonical form of the composition of two mappings.

A technical problem with the φ-operation is that the arrowheads are altered, so that the result $B^{+\varphi([0,a_0]^-) \,\cdots\, \varphi([n-1,a_{n-1}]^-)}$ of the derivation is not an element of S^+, even after applying additional rules to compute the products in $S_n^+, \ldots, S_1^+$ which amounts to merging arrows with identical heads. For a partial monotone function $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ satisfying $f(i) \leq i$ whenever $i \in \mathrm{Dom}(f)$, we define a monotone extension $\mathrm{Monex}(f) : \{1, \ldots, n\} \to \{1, \ldots, n\}$ by

$$\mathrm{Monex}(f)[n] = \begin{cases} f(n) & \text{if } n \in \mathrm{Dom}(f), \\ n & \text{otherwise}, \end{cases}$$

$$\mathrm{Monex}(f)[i] = \begin{cases} f(i) & \text{if } i \in \mathrm{Dom}(f), \\ \min(i, \mathrm{Monex}(f)[i+1]) & \text{otherwise}. \end{cases}$$

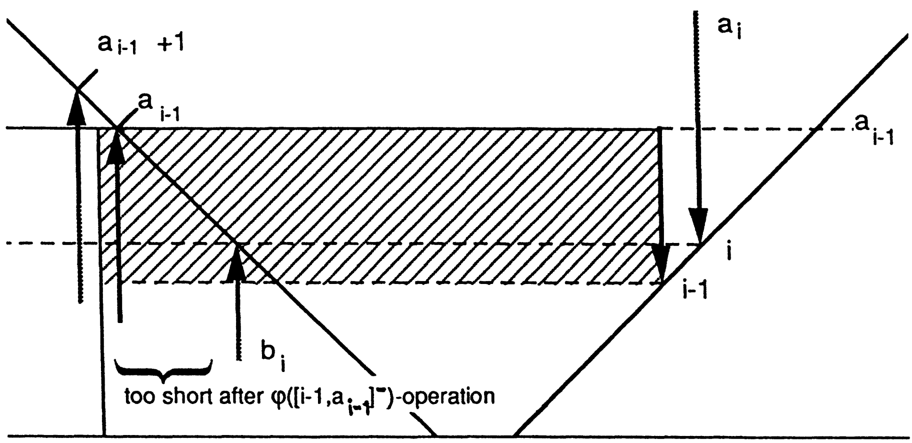
With that concept in mind we are ready to define

$$B^{+\varphi(A^-)}(n) = \begin{cases} b_n & \text{if } a_{n-1} = n-1, \\ n & \text{otherwise}, \end{cases}$$

$$B^{+\varphi(A^-)}(i) = \begin{cases} b_i & \text{if } a_{i-1} = i-1, \\ \min(i, b_{a_{i-1}+1}) & \text{if } a_i > a_{i-1} \geq i, \\ \min(i, B^{+\varphi(A^-)}(i+1)) & \text{if } a_i = a_{i-1}. \end{cases}$$

Lemma 1. $B^{+\varphi(A^-)} = \text{Monex}\left(\textit{reduced}\left(B^{+\varphi([0,a_0]^-)\ \cdots\ \varphi([n-1,a_{n-1}]^-)}\right)\right)$ *where the reduced sequence is obtained by computing the products in* $S_n^+, \ldots, S_1^+$, *i.e., applying rules* $[c,j]^+[d,j]^+ \to [\min(c,d),j]^+$.

Proof. Case $a_{i-1} = i-1$. Then $a_k \leq i-1$ for every $k < i$. Hence no $\varphi([k,a_k]^-)$ for $k < i$ can affect $[b_i,i]^+$ and $\varphi([k,a_k]^-)$ for $k \geq i$ won't affect it anyway. So $[b_i,i]^+$ remains as it is. Also note that $\uparrow$-arrows only can get shortened from the top. Since $(b_n,\ldots,b_1)$ is monotone, no shortened $[b_j,j]^+$ for $j < i$ has a longer tail than $[b_i,i]^+$.



Otherwise we have $a_{i-1} \geq i$. Then $[b_i,i]^+$ gets its head cut by $\varphi([i-1,a_{i-1}]^-)$ and does not make a contribution to the i-th component of $B^{+\varphi([0,a_0]^-)\ \cdots\ \varphi([n-1,a_{n-1}]^-)}$. The same is true for $[b_j,j]^+$ satisfying $j \leq \alpha_{i-1}$. Hence the first $[b_j,j]^+$ that may contribute is $[b_{a_{i-1}+1}, a_{i-1}+1]^+$, and $[b_j,j]^+$ for $j > \alpha_{i-1}+1$ won't contribute any more.

Case $a_i > a_{i-1}$. Then $a_i \geq a_{i-1}+1$ and we actually get $[b_{a_{i-1}+1}, a_{i-1}+1]^{+\varphi([i,a_i]^-)} = [b_{a_{i-1}+1}, i]^+$ as long as $b_{a_{i-1}+1} \leq i$, and we obtain $[i,i]^+$ otherwise.

If finally $a_i = a_{i-1}$, then $B^{+\varphi([0,a_0]^-)\ \cdots\ \varphi([n-1,a_{n-1}]^-)}$ does not contain an arrow with head i, and some value at i is provided from values at $i+1$, $i+2,\ldots,n$ by monotone extension.

The following two Lemmata will complete the proof of Proposition 4.

Lemma 2. *For* $A^- = (a_0,\ldots,a_{n-1})^- \in S^-$, $C^- = (c_0,\ldots,c_{n-1})^- \in S^-$, $B^+ = (b_n,\ldots,b_1)^+ \in S^+$:

(a) $$B^{+\varphi(A^- \circ C^-)} = \left(B^{+\varphi(A^-)}\right)^{\varphi(C^-)}$$

(b) $$\delta_{B^+}(A^- \circ C^-) = \delta_{B^+}(A^-) \circ \delta_{B^{+\varphi(A^-)}}(C^-)$$

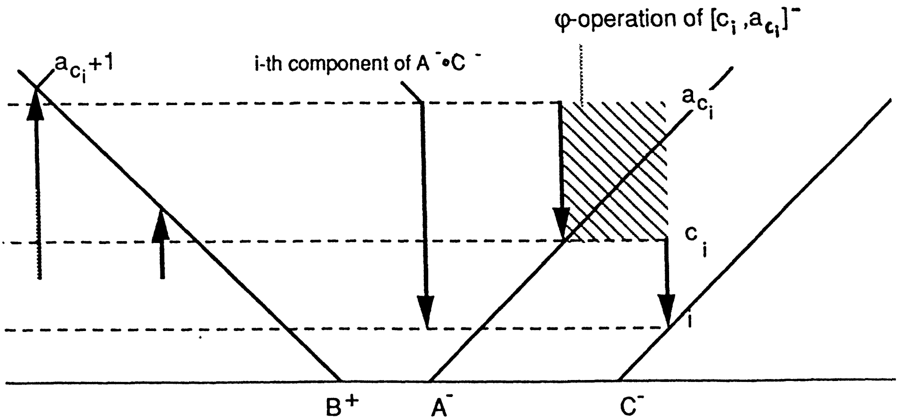
Proof. First note $A^- \circ C^- = (a_{c_0},\ldots,a_{c_{n-1}})^- \in S^-$.

(a) $\varphi([0,a_0]^-),\ldots,\varphi([n-1,a_{n-1}]^-)$ act on the heads of B^+ like $\mu(A^-)$. Thereafter, $\varphi([0,c_0]^-),\ldots,\varphi([n-1,c_{n-1}]^-)$ act on the heads of the result, joint by the monotone extension arrows, like $\mu(C^-)$. Finally monotone extension arrows are added again. – The same effect on the heads of B^+ is achieved by applying the composition $\mu(A^-)\circ\mu(C^-)=\mu(A^-\circ C^-)$ and filling the result with monotone extension arrows afterwards: Arrows resulting from the intermediate monotone extension in $B^{+\varphi(A^-)}$ do have identical base as some original arrow to the left. After applying $\varphi([0,c_0]^-),\ldots,\varphi([n-1,c_{n-1}]^-)$ this original arrow causes monotone extension arrows in the final step to be added that supersede those arrows from the intermediate extension in $B^{+\varphi(A^-)}$.

(b) We want to show that the i-th component of the right side equals that of the left side, namely

$$\delta_{B^+}(A^-\circ C^-)[i]=\delta_{[b_n,n]^+}\left(\cdots\ \left(\delta_{[b_1,1]^+}([i,a_{c_i}]^-)\right)\right)\ .$$

Those $[b_j,j]^+$ with $j\le c_i$ do not make any contribution to either side by definition of δ, because the $\varphi(A^-)$-operation can only shorten the $\uparrow$-arrows by cutting their heads. Those $[b_j,j]^+$ with j in the range $c_i<j\le a_{c_i}$ do not affect the left side, nor the c_i-th component of A^-. But they would act on $[i,c_i]^-$ in case $b_j\le c_i$, if that would not be taken care of by the φ-operation: $[b_j,j]^{+\varphi([c_i,a_{c_i}]^-)}=[b_j,c_i]^+$ which is too short to be able to act on $[i,c_i]^-$. Also, $\uparrow$-arrows added by monotone extension in $B^{+\varphi(A^-)}$ may be disregarded for further δ-transition anyway. It remains to consider $[b_j,j]^+$ with $j>a_{c_i}$ and, in order to have any influence, $b_j\le a_{c_i}$. These $[b_j,j]^+$ are not altered by the φ-operations $\varphi([0,a_0]^-),\ldots,\varphi([c_i,a_{c_i}]^-)$, because the corresponding $\downarrow$-arrows are too short. Although these $[b_j,j]^+$ may possibly have cut their heads by $\varphi([c_i+1,a_{c_i+1}]^-),\ldots,\varphi([n-1,a_{n-1}]^-)$, they still remain tall enough (i.e., with head $\ge c_i$) to be able to act on $[0,c_0]^-,\ldots,[i,c_i]^-$ the same way as $[b_j,j]^+$ would. Because of the monotone order in B^+ the final result is determined by $[b_j,j]^+$ with $j=a_{c_i}+1$. So we may assume $j=a_{c_i}+1$.



Case $b_j \leq i$. Then $\delta_{B^+}(A^- \circ C^-)$, $\delta_{B^+}(A^-)$, $\delta_{B^{+\varphi(A^-)}}(C^-)$ all have trivial i-th component.

Case $i < b_j \leq c_i$. Then $\delta_{B^+}(A^- \circ C^-)[i] = \delta_{B^{+\varphi(A^-)}}(C^-)[i] = b_j - 1$ and $\delta_{B^+}(A^-)[b_j - 1] = b_j - 1$, because either $b_j - 1 = a_{b_j - 1}$ to start with or $b_j \leq a_{b_j-1} \leq a_{c_i} < j$ and δ-transition takes place.

Case $c_i < b_j \leq a_{c_i}$. Then $\delta_{B^+}(A^- \circ C^-)[i] = b_j - 1 = \delta_{B^+}(A^-)[c_i]$ and $\delta_{B^{+\varphi(A^-)}}(C^-)[i] = c_i$.

Lemma 3. *For $A^- = (a_0, \ldots, a_{n-1})^- \in S^-$, $B^+ = (b_n, \ldots, b_1)^+ \in S^+$, $D^+ = (d_n, \ldots, d_1)^+ \in S^+$:*

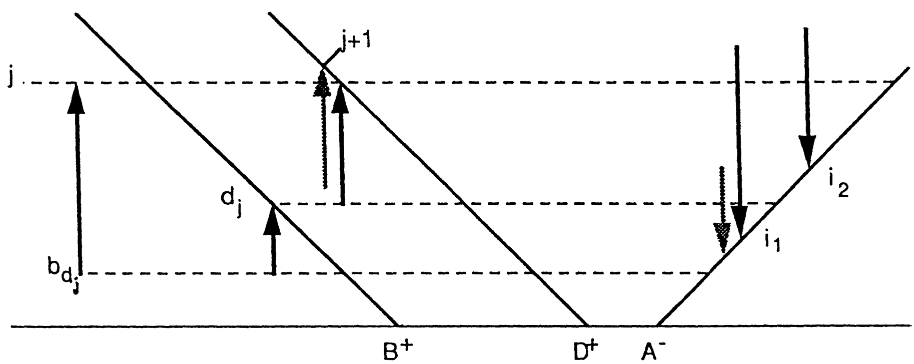
$$(a) \qquad \delta_{B^+ \circ D^+}(A^-) = \delta_{B^+}\left(\delta_{D^+}(A^-)\right)$$

$$(b) \qquad (B^+ \circ D^+)^{\varphi(A^-)} = B^{+\varphi(\delta_{D^+}(A^-))} \circ D^{+\varphi(A^-)} .$$

Proof. First note $B^+ \circ D^+ = (b_{d_n}, \ldots, b_{d_1})^+ \in S^+$.

(a) is obvious from the definition of δ.

(b) We apply Lemma 1 and study the operation $\varphi([0, a_0]^-), \ldots, \varphi([n-1, a_{n-1}]^-)$ on $B^+ \circ D^+$. Consider the j-th component $[b_{d_j}, j]^+$ of $B^+ \circ D^+$. Since δ_{D^+} shortens the $\downarrow$-arrows of A^- at their tails, only those $\varphi([i, a_i]^-)$ satisfying $b_{d_j} \leq i < j$ may act nontrivially on $[b_{d_j}, j]^+$, $[b_{d_j}, d_j]^+$, or $[d_j, j]^+$.



Case $a_i < j$. Then the φ-operation of $[i, a_i]^-$ on $[b_{d_j}, j]^+$ and $[d_j, j]^+$ is trivial. For the φ-operation of $[i, a_i]^-$ on $[b_{d_j}, d_j]^+$ to be nontrivial, we must have $i \leq d_j \leq a_i$. But in that case $\delta_{D^+}(A^-)[i] \leq \delta^+_{[d_j, j]}(A^-)[i] < d_j$, so that the d_j-th component of B^+ is not affected either by the φ-operation of the $\downarrow$-arrow of $\delta_{D^+}(A^-)$ that results from $[i, a_i]^-$. Thus, only the following case is of interest:

Case $j \leq a_i$. Define $I_1 = \{i \mid b_{d_j} \leq i < d_j \text{ and } j \leq a_i\}$, $I_2 = \{i \mid d_j \leq i < j \text{ and } j \leq a_i\}$.

Subcase $I_1 = \emptyset$. If $I_2 = \emptyset$, too, then no action takes place. Otherwise let $i_2 = \min(I_2)$. Then $[b_{d_j}, j]^{+\varphi(A^-)} = [b_{d_j}, j]^{+\varphi([i_2, a_{i_2}]^-)} = [b_{d_j}, i_2]^+$, $[d_j, j]^{+\varphi(A^-)} = [d_j, j]^{+\varphi([i_2, a_{i_2}]^-)} = [d_j, i_2]^+$, and $[b_{d_j}, d_j]^{+\varphi(\delta_{D^+}(A^-))} = [b_{d_j}, d_j]^+$ because of the discussion above.

Subcase $I_1 \neq \emptyset$. Let $i_1 = \min(I_1)$. Then

$$[b_{d_j}, j]^{+\varphi(A^-)} = [b_{d_j}, j]^{+\varphi([i_1, a_{i_1}]^-)} = [b_{d_j}, i_1]^+ .$$

It may happen that $b_{d_j} = b_{d_{j+1}}$. However, in that case we may drop $[b_{d_j}, j]^+$ from its consideration as a component of $B^+ \circ D^+$ altogether, because its contribution to $(B^+ \circ D^+)^{\varphi(A^-)}$ will be superseded by monotone extension arrows anyway. Now assuming $b_{d_j} \neq b_{d_{j+1}}$, $d_{j+1} > d_j$ follows. Since for $k \leq j$ the δ-transition of $[i_1, a_{i_1}]^-$ induced by $[d_k, k]^+$ is trivial and $d_{j+1} > d_j$, $\delta_{D^+}([i_1, a_{i_1}]^-)$ has a tail at least as long as d_j. Therefore the φ-operation of $\delta_{D^+}([i_1, a_{i_1}]^-)$ on $[b_{d_j}, d_j]^+$ yields $[b_{d_j}, i_1]^+$. Regarding $D^{+\varphi(A^-)}$ we note in this case: For $k \leq j$ the $\uparrow$-arrows of $D^{+\varphi(A^-)}$ which are based at i_1 or below cannot extend their heads beyond i_1 (even after monotone extension). For $k > j$ the $\uparrow$-arrows of $D^{+\varphi(A^-)}$ are already based at values $> d_j > i_1$. Thus, forming the product on the right hand side of (b) still gives i_1-st component $[b_{d_j}, i_1]^+$. Monotone extension does not make a difference for the value of the product.

4. The Full Transformation Semigroup

Again, consider linear order relations on a finite set X. As mentioned above, the semigroup $\mathrm{End}(X, \leq)$ of isotonic mappings is extensively studied in the context of semigroups acting on graphs [10]. But independent of graphs and order, together with the symmetric group, it serves as the other pillar of the full transformation semigroup in a bilateral semidirect decomposition:

Proposition 5. *For any linear order $\leq$ on X, the full transformation semigroup $\mathcal{T}(X)$ is a canonical homomorphic image of a bilateral semidirect product*

$$S_X \;{}_\delta\!\times_\varphi \mathrm{End}(X, \leq)$$

of the symmetric group S_X on X and the semigroup $\mathrm{End}(X, \leq)$ *of isotonic mappings.*

Proof. For $\alpha \in \mathcal{T}(X)$ define a derived linear order $(X, \leq_\alpha)$ by

$$x \leq_\alpha y \iff \begin{cases} x\alpha \leq y\alpha & \text{if } x\alpha \neq y\alpha , \\ x \leq y & \text{otherwise} . \end{cases}$$

Let σ_α be the permutation of X that rearranges the order $(X, \leq)$ into $(X, \leq_\alpha)$ by sorting with respect to $\leq_\alpha$. This means

$$x \leq_\alpha y \iff x\sigma_\alpha \leq y\sigma_\alpha . \tag{$*$}$$

As a consequence $\sigma_\alpha^{-1}\alpha$ is isotonic with respect to the original order $(X, \leq)$:

$$x \leq y \Leftrightarrow (x\sigma_\alpha^{-1})\sigma_\alpha \leq (y\sigma_\alpha^{-1})\sigma_\alpha \Leftrightarrow x\sigma_\alpha^{-1} \leq_\alpha y\sigma_\alpha^{-1} \Rightarrow x\sigma_\alpha^{-1}\alpha \leq y\sigma_\alpha^{-1}\alpha .$$

Therefore $\alpha = \sigma_\alpha \cdot \sigma_\alpha^{-1}\alpha \in S_X \circ \mathrm{End}(X, \leq)$. So $\mathcal{T}(X) = S_X \circ \mathrm{End}(X, \leq)$. Now define $\delta : \mathrm{End}(X, \leq) \to \mathcal{T}(S_X)$ and $\varphi : S_X \to \mathcal{T}(\mathrm{End}(X, \leq))$ by

$$\delta_\tau(\pi) = \sigma_{\tau\pi} \qquad \text{and} \qquad \tau^{\varphi(\pi)} = \sigma_{\tau\pi}^{-1}\tau\pi$$

for $\tau \in \mathrm{End}(X,\leq)$ and $\pi \in S_X$. We want to prove that δ, φ give rise to a bilateral semidirect product $S_X\ {}_\delta\!\times_\varphi \mathrm{End}(X,\leq)$. Obviously,

$$\mu : S_X\ {}_\delta\!\times_\varphi \mathrm{End}(X,\leq) \to \mathcal{T}(X), \quad \mu(\pi,\tau) = \pi\circ\tau$$

will be a homomorphism. First note

$$\delta_\tau(id) = \sigma_\tau = id, \delta_{id}(\pi) = \sigma_\pi = \pi, (id)^{\varphi(\pi)} = \pi^{-1}\pi = id, \tau^{\varphi(id)} = \sigma_\tau^{-1}\tau = \tau\ .$$

Claim. δ is an anti-homomorphism.

Proof. Let $\tau_1,\tau_2 \in \mathrm{End}(X,\leq)$ and $\pi \in S_X$. We have to compare $\delta_{\tau_1\tau_2}(\pi) = \sigma_{\tau_1\tau_2\pi}$ with $\delta_{\tau_1}(\delta_{\tau_2}(\pi)) = \sigma_{\tau_1\sigma_{\tau_2\pi}}$. Translated into order relations, we have to show that $\leq_{\tau_1\tau_2\pi}$ coincides with $\leq_{\tau_1\sigma_{\tau_2\pi}}$. If $x\tau_1\tau_2\pi \neq y\tau_1\tau_2\pi$, then also $x\tau_1\sigma_{\tau_2\pi} \neq y\tau_1\sigma_{\tau_2\pi}$ and in this case we have

$$\begin{array}{lllll}
x \leq_{\tau_1\tau_2\pi} y & \Longleftrightarrow & x\tau_1\tau_2\ \pi & \leq & y\tau_1\tau_2\pi & \\
& \Longleftrightarrow & x\tau_1 & \leq_{\tau_2\pi} & y\tau_1 & \text{in this case} \\
& \Longleftrightarrow & x\tau_1\sigma_{\tau_2\pi} & \leq & y\tau_1\sigma_{\tau_2\pi} & \text{by } (*) \\
& \Longleftrightarrow & x & \leq_{\tau_1\sigma_{\tau_2\pi}} & y & \text{in this case}
\end{array}$$

If $x\tau_1\tau_2\pi = y\tau_1\tau_2\pi$ but $x\tau_1\sigma_{\tau_2\pi} \neq y\tau_1\sigma_{\tau_2\pi}$, then

$$\begin{array}{lllll}
x \leq_{\tau_1\tau_2\pi} y & \Longleftrightarrow & x & \leq & y & \text{in this case} \\
& \Longleftrightarrow & x\tau_1 & \leq & y\tau_1 & \text{because } x\tau_1 \neq y\tau_1 \\
& \Longleftrightarrow & x\tau_1 & \leq_{\tau_2\pi} & y\tau_1 & \text{by definition} \\
& \Longleftrightarrow & x\tau_1\sigma_{\tau_2\pi} & \leq & y\tau_1\sigma_{\tau_2\pi} & \text{by } (*) \\
& \Longleftrightarrow & x & \leq_{\tau_1\sigma_{\tau_2\pi}} & y & \text{in this case}
\end{array}$$

The remaining case where $x\tau_1 = y\tau_1$ is trivial.

Claim. $\delta_\tau(\pi_1\pi_2) = \delta_\tau(\pi_1)\cdot\delta_{\tau^{\varphi(\pi_1)}}(\pi_2)$, or equivalently

$$\sigma_{\tau\pi_1\pi_2} = \sigma_{\tau\pi_1}\cdot\sigma_{\sigma_{\tau\pi_1}^{-1}\tau\pi_1\pi_2}\ .$$

Proof. With the claim translated into order relations we compute

$$\begin{array}{l}
x\,\sigma_{\tau\pi_1}\cdot\sigma_{\sigma_{\tau\pi_1}^{-1}\tau\pi_1\pi_2} \leq y\sigma_{\tau\pi_1}\cdot\sigma_{\sigma_{\tau\pi_1}^{-1}\tau\pi_1\pi_2} \\
\qquad \Longleftrightarrow\ x\,\sigma_{\tau\pi_1} \leq_{\sigma_{\tau\pi_1}^{-1}\tau\pi_1\pi_2} \quad y\sigma_{\tau\pi_1}\ .
\end{array}$$

In case $x\tau\pi_1\pi_2 \neq y\tau\pi_1\pi_2$ we continue:

$$\begin{array}{llll}
\Longleftrightarrow & x\,(\sigma_{\tau\pi_1}\sigma_{\tau\pi_1}^{-1})\tau\pi_1\pi_2 & \leq & y\,(\sigma_{\tau\pi_1}\sigma_{\tau\pi_1}^{-1})\tau\pi_1\pi_2 \\
\Longleftrightarrow & x & \leq_{\tau\pi_1\pi_2} & y \\
\Longleftrightarrow & x\,\sigma_{\tau\pi_1\pi_2} & \leq & y\,\sigma_{\tau\pi_1\pi_2}
\end{array}$$

Otherwise we have $x\tau\pi_1\pi_2 = y\tau\pi_1\pi_2$ which implies $x\tau = y\tau$, and the above chain of equivalences may be continued as follows:

$$\begin{array}{llll}
\Longleftrightarrow & x\,\sigma_{\tau\pi_1} & \leq & y\,\sigma_{\tau\pi_1} \\
\Longleftrightarrow & x & \leq & y \\
\Longleftrightarrow & x\,\sigma_{\tau\pi_1\pi_2} & \leq & y\,\sigma_{\tau\pi_1\pi_2}
\end{array}$$

A straightforward calculation showing that the required properties of φ follow from those of δ concludes the proof of Proposition 5.

Corollary. *Every finite semigroup divides a bilateral semidirect product of a group and an aperiodic semigroup.*

Hence, with respect to bilateral semidirect products a notion of group complexity of a semigroup, similar to the one in the sense of Rhodes with respect to the wreath product, does not arise. An open problem now is to compare the number of factors in decompositions based on bilateral semidirect products and based on block products. In particular, this is also interesting for aperiodic semigroups which can be obtained from semilattices either way. About the variety setting we only mention

Proposition 6. *Every semigroup in the variety* $\mathbf{LJ_1}$ *of locally testable semigroups is covered by a semidirect product of the form* $2^{D_{\Sigma,k}} \times_\delta D_{\Sigma,k}$ *where* $D_{\Sigma,k}$ *is the semigroup of Proposition 1 and* $2^{D_{\Sigma,k}}$ *denotes the power set of* $D_{\Sigma,k}$ *together with the union of sets as semigroup multiplication.*

Appendix

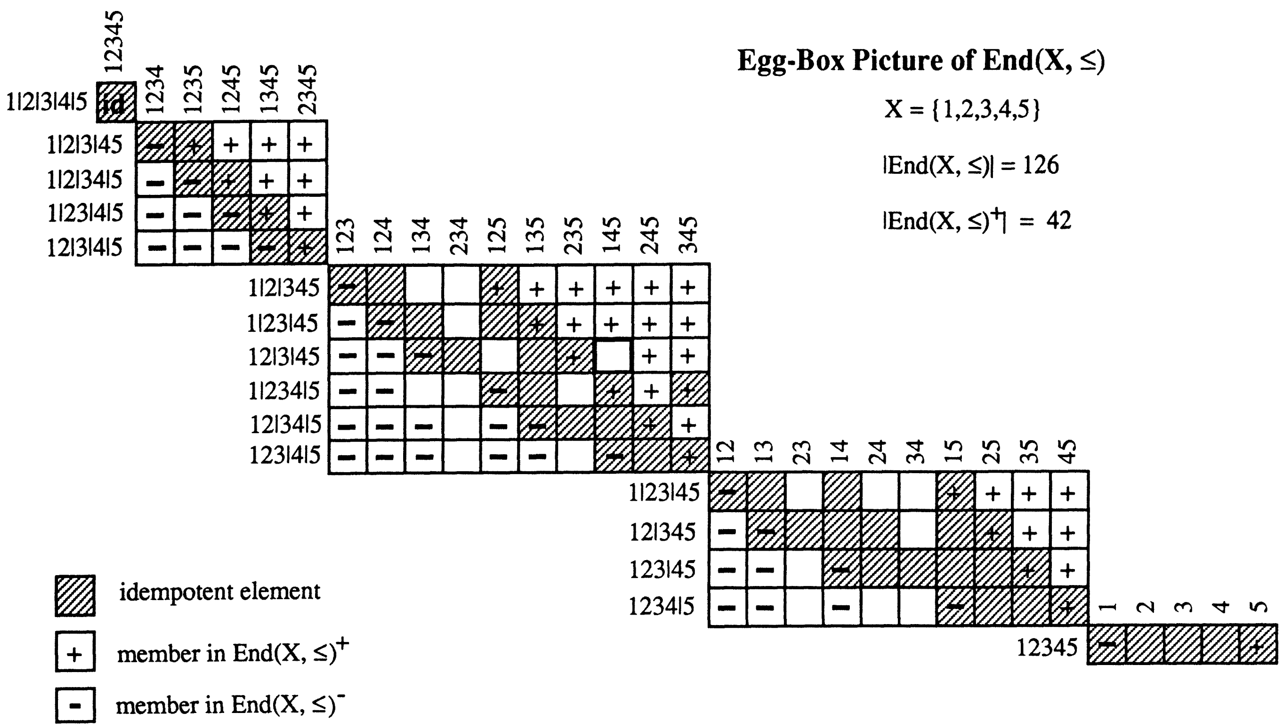
When investigating the structure of a finite semigroup S, a usual approach is to study the idempotent elements of S. Our idea is to elaborate that approach by involving part of the semigroup structure and to study subsemilattices and their products. While a complete survey of all subsemilattices and their products is most likely an unrealistic goal (e.g., just think of the full transformation semigroup $\mathcal{T}(X)$), it is probably sensitive to single out certain subsemilattices that provide some insight, such as a bilateral semidirect decomposition. This is what we did for $\mathrm{End}(X, \leq)$ in the diagram shown on the page following the egg-box picture. The upper part illustrates Propositions 3 and 4. The part below $S^-_\leq S^+_\leq$ is merely included to exhibit the relationship to the Example in Section 1. Since the general bilateral semidirect product is a very powerful operation, it may be interesting to observe that both Propositions 4 and 5 require only one application of that product, independent of the size of X. This is in contrast to the bilateral semidirect decomposition of the symmetric group on X where nested application is necessary, indicating for another time that a complex semigroup structure is often related to nontrivial subgroups.

Egg-Box Picture of End(X, ≤)

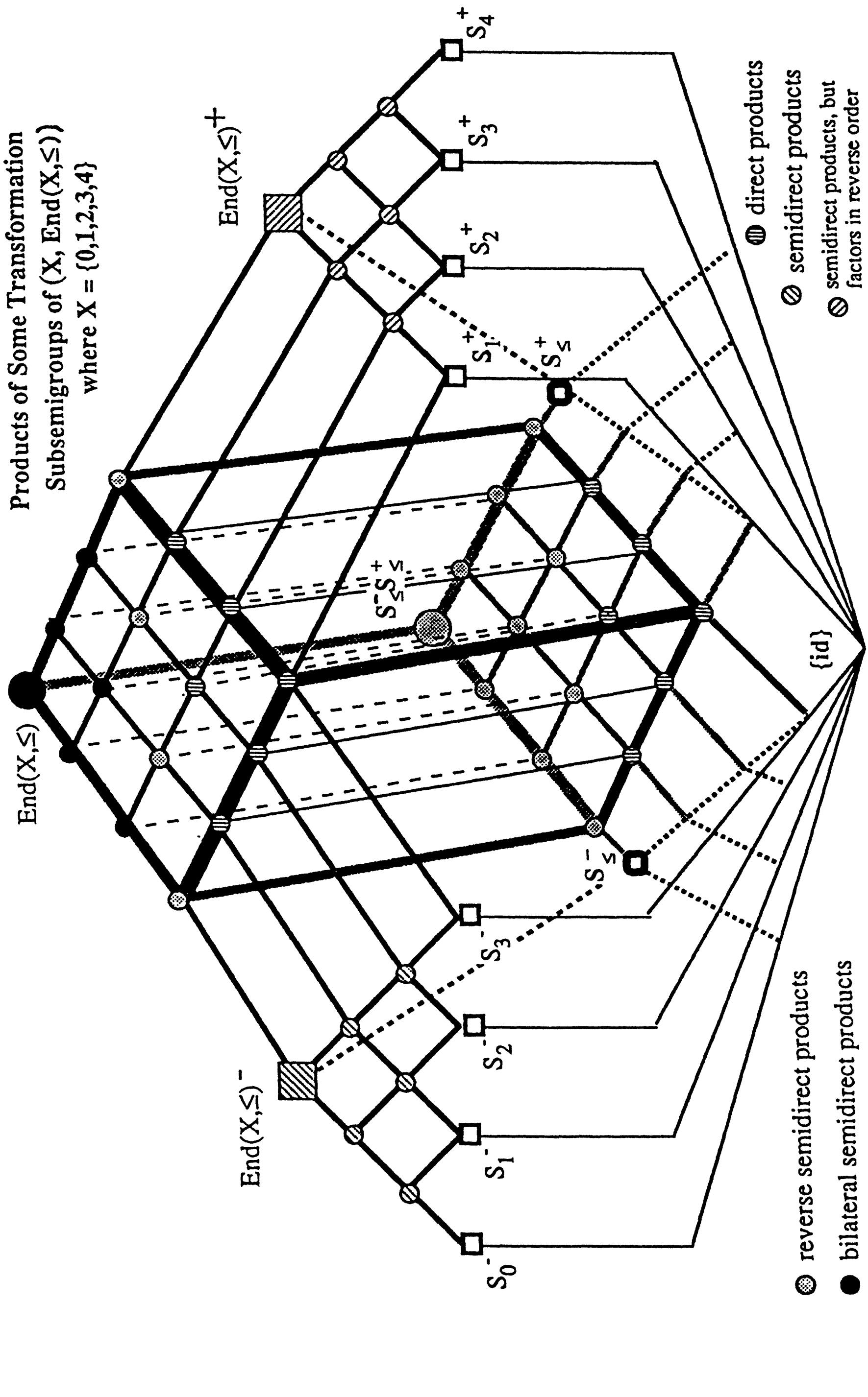
X = {1,2,3,4,5}

|End(X, ≤)| = 126

|End(X, ≤)$^+$| = 42



For $X = \{1,2,3,\ldots,n\}$ we have $|\mathrm{End}(X,\leq)| = \sum_{k=0}^{n-1} \binom{n-1}{k}\binom{n}{k+1} = \binom{2n-1}{n}$ and $|\mathrm{End}(X,\leq)^+| = \binom{2n}{n}$ = the n-th Catalan number. Therefore $|\mathrm{End}(X,\leq)|/|\mathrm{End}(X,\leq)^+| = \frac{n+1}{2}$.


Products of Some Transformation
Subsemigroups of (X, End(X,≤))
where X = {0,1,2,3,4}
End(X,≤)
End(X,≤)+
End(X,≤)−
S0−
S1−
S2−
S3−
S1+
S2+
S3+
S4+
S≤+
S≤−
S≤−S≤+
{id}
direct products
semidirect products
semidirect products, but factors in reverse order
reverse semidirect products
bilateral semidirect products

The following table indicates the types of products of transformation subsemigroups of $\mathrm{End}(X, \leq)$, $X = \{0, 1, 2, 3, 4\}$, of the form

$$(S^-_{i_1} \cdots S^-_{i_k}) \cdot (S^+_{j_m} \cdots S^+_{j_1})$$

where the indices $i_1, \ldots, i_k$ are given as row headings and the indices $j_1, \ldots, j_m$ are given as column headings. The heavy outlined cells correspond to products shown in the diagram on the previous page.

	up → down ↓	00 1	01 2	02 2 1	03 3	04 3 2	05 3 2 1	06 4	07 4 3	08 4 3 2	09 4 3 2 1
00	0, 1, 2, 3	r	02	g	05	05	g	09	09	09	g
10	0, 1, 2	r	12	g	15	15	g		19	19	g
20	0, 1	r	22	g		25	g			29	g
30	0	r	g	g	g	g	g	g	g	g	g
40	1, 2, 3	d	42	r	45	45	g	49	49	49	g
50	1, 2	d	52	r	55	55	g		59	59	g
60	1	d	62	r		65	g			69	g
70	2, 3	d	d	d	74	r	r	78	78	g	g
80	2	d	d	d	84	r	r		88	g	g
90	3	d	d	d	d	d	d	97	r	r	r

d direct product

r reverse semidirect product

g general bilateral semidirect product

Blanks and numbers indicate that the set theoretic product is not a semigroup. If the generated subsemigroup is covered by the table, then it is referred to by its row/column ID number.

References

[1] Aigner, M., *Combinatorial Theory*, Springer 1979.

[2] Aizenstat, A., *The defining relations of the endomorphism semigroup of a finite linearly ordered set*, Sibirsk. Mat. Z. **3** (1962), 161–169 (Russian, Math. Reviews **26** (1963), #6287).

[3] Crvenković , S. and M. Kunze, *Actions of semilattices*, Semigroup Forum **34** (1986), 139–156.

[4] Eilenberg, S., *Automata, Languages, and Machines*, vol. B. Academic Press 1976.

[5] Howie, J. M., *Products of idempotents in finite full transformation semigroups*, Proc. Roy. Soc. Edinburgh **98A** (1984), 25–35.

[6] Kimura, N., *Structure of idempotent semigroups, I*, Pac. J. Math. **8** (1958), 257–275.

[7] Kunze, M., *Lineare Parallelrechner*, EIK J. Inf. Proc. and Cybernetics **20** (1984), 9–39 and 111–147.

[8] Kunze, M., *Semidirect products of transformation semigroups*, Univ. of Arkansas, Fayetteville, 1988.

[9] Kunze, M., *Zappa products*, Acta Math. Hungar. **41** (1983), 225–239.

[10] Molchanov, V. A., *Semigroups of mappings on graphs*, Semigroup Forum **27** (1983), 155–199.

[11] Rhodes, J. and B. Tilson, *The kernel of monoid morphisms*, J. Pure Appl. Algebra **62** (1989), 227–268.

[12] Rhodes, J. and P. Weil, *Decomposition techniques for finite semigroups, using categories*, J. Pure Appl. Algebra **62** (1989), 269–312.

[13] Schein, B., *Injective monars over inverse semigroups*, "Algebraic Theory of Semigroups", Szeged, 1976. Colloq. Math. Soc. Janos Bolyai **20** (1979), 519–544.

[14] Tilson, B., *Categories as algebras: an essential ingredient in the theory of monoids*, J. Pure Appl. Algebra **48** (1987), 83–198.

U. S. Peace Corps
P. O. Box 564
Victoria, Mahé
Seychelles (Indian Ocean)

Received May 20, 1990
and in final form October 26, 1990

Semigroup Forum Vol. 45 (1992) 183–190

RESEARCH ARTICLE

On the structure of (m,n)-commutative semigroups

Attila Nagy *

Communicated by J. M. Howie

As in [2], a semigroup S is called (m,n)-commutative (m and n are positive integers) if, for all $x_1, ..., x_{m+n} \in S$,

$$(x_1 ... x_m)(x_{m+1} ... x_{m+n}) = (x_{m+1} ... x_{m+n})(x_1 ... x_m).$$

In this paper we deal with the semilattice decomposition of (m,n)-commutative semigroups, describe subdirectly irreducible (m,n)-commutative semigroups and determine (m,n)-commutative Δ-semigroups.

1. Semilattice decomposition of (m,n)-commutative semigroups

First of all we note that if a semigroup is (m,n)-commutative for some m and n, then it is (m^*,n^*)-commutative for all $m^* \geq m$ and $n^* \geq n$. This fact will be used in this paper without comment.

Theorem 1.1. *Every (m,n)-commutative semigroup is a semilattice of (m,n)-commutative archimedean semigroups.*

Proof. In [4], Putcha proved that a semigroup S is decomposable as a semilattice of archimedean semigroups if and only if, for all $a, b \in S$, the assumption $b \in S^1aS^1$ implies $b^k \in S^1a^2S^1$ for some positive integer k.

Let S be an (m,n)-commutative semigroup and a, b be arbitrary elements of S with $b \in S^1aS^1$. Then $b = xay$ for some $x, y \in S^1$. As

$$b^{m+n+2} = (xay)^{m+n+2} = (xay)^{m+1}(xay)^{n+1}$$

$$= xa(yxa)^m yxa(yxa)^n y$$

$$= xa[(yxa)^m yx][a(yxa)^n]y$$

$$= xa[a(yxa)^n][(yxa)^m yx]y = xa^2(yxa)^{n+m}yxy,$$

$b^{m+n+2} \in S^1a^2S^1$. So S is a semilattice of archimedean semigroups. It is evident that the semilattice components of S are (m,n)-commutative. ∎

Theorem 1.2. *A semigroup is simple [0-simple] and (m,n)-commutative if and only if it is a commutative group [a commutative group with 0 adjoined].*

Proof. Let S be a simple [0-simple] (m,n)-commutative semigroup. By the simplicity [0-simplicity] of S, $S^{m+n} = S$. By the (m,n)-commutativity of S, S^{m+n} is commutative. So S is a commutative group [a commutative group with zero adjoined]. As the converse statement is trivial, the theorem is proved. ∎

* The author thanks Professor J. M. Howie for his comments and support.

Theorem 1.3. *A semigroup is (m,n)-commutative archimedean and has an idempotent element if and only if it is an ideal extension of a commutative group by an (m,n)-commutative nil semigroup.*

Proof. Let S be an (m,n)-commutative archimedean semigroup with an idempotent element f. As S is archimedean, $G = SfS$ is the kernel of S. If $\mid G \mid = 1$, then S is a nil semigroup. So we may assume $\mid G \mid > 1$. Then G is a simple (m,n)-commutative semigroup. By Theorem 1.2, G is a commutative group. As S is archimedean, $Q = S/G$ (the Rees factor semigroup of S determined by the ideal G) is a nil semigroup. It is clear that Q is (m,n)-commutative.

Conversely, let S be a semigroup such that S is an ideal extension of a commutative group G by an (m,n)-commutative nil semigroup Q. Then S is an archimedean semigroup with an idempotent element.

We show that S is (m,n)-commutative. Let $x_1, \ldots, x_{m+n}$ be arbitrary elements of S. If $(x_1 \ldots x_m)(x_{m+1} \ldots x_{m+n}) \notin G$, then $x_i \notin G$ for all $i = 1, \ldots, m+n$ and so

$$(x_1 \ldots x_m)(x_{m+1} \ldots x_{m+n}) = (x_{m+1} \ldots x_{m+n})(x_1 \ldots x_m),$$

because Q is (m,n)-commutative. In the case $(x_1 \ldots x_m)(x_{m+1} \ldots x_{m+n}) \in G$, we have $(x_{m+1} \ldots x_{m+n})(x_1 \ldots x_m) \in G$, because G is an ideal of S and Q is (m,n)-commutative. Denoting the identity element of G by e,

$$es = e(es) = (es)e = e(se) = se,$$

for all $s \in S$, because es and se are in G. So

$$\begin{aligned} &(x_1 \ldots x_m)(x_{m+1} \ldots x_{m+n}) \\ &= e^m (x_1 \ldots x_m)(x_{m+1} \ldots x_{m+n}) e^n \\ &= (x_{m+1} \ldots x_{m+n}) e^n e^m (x_1 \ldots x_m) \\ &= e(x_{m+1} \ldots x_{m+n})(x_1 \ldots x_m) \\ &= (x_{m+1} \ldots x_{m+n})(x_1 \ldots x_m). \end{aligned}$$

Consequently S is (m,n)-commutative. Thus the theorem is proved. ∎

Theorem 1.4. *An (m,n)-commutative archimedean semigroup without idempotents has a non-trivial group homomorphic image.*

Proof. Let S be an (m,n)-commutative archimedean semigroup without idempotents and let a be an arbitrary element of S. We show that

$$S_a = \{x \in S : (\exists t, r, s \in N)\ a^t x a^r = a^s\}$$

(N is the set of all positive integers) is a reflexive unitary subsemigroup of S and S_a is minimal among reflexive unitary subsemigroups of S containing the element a.

Let x, y be arbitrary elements of S_a. Then there are positive integers t, r, s, i, j, k such that

$$a^t x a^r = a^s$$

and

$$a^i y a^j = a^k.$$

We may assume that $t, r, i, j \geq max\{m, n\}$. Then

$$a^{s+k} = a^t x(a^{r+i})(ya^j) = a^t xya^{j+r+i}$$

and so $xy \in S_a$. Thus S_a is a subsemigroup of S.

Assume $x, xb \in S_a$ for some $x, b \in S$. Then there are positive integers t, r, s, i, j, k such that

$$a^t xa^r = a^s$$

and

$$a^i xba^j = a^k.$$

We may assume that $t, r \geq max\{m, n\}$. Then $a^s = a^{r+t}x$. Let p be a positive integer such that $p \geq max\{i, r+t\}$. Then

$$a^p x = a^{s+p-(r+t)}$$

and so

$$a^{k+p-i} = a^p xba^j = a^{s+p-(r+t)}ba^j.$$

So $b \in S_a$, that is S_a is left unitary in S. We can prove, in a similar way, that S_a is right unitary in S. So S_a is an unitary subsemigroup of S.

To prove the reflexivity of S_a, assume that $xy \in S_a$ for some $x, y \in S$. Then

$$a^t xya^r = a^s$$

for some positive integers $t, r, s \geq max\{m, n\}$. As S is (m, n)-commutative,

$$a^s = ya^r a^t x$$

from which we get

$$a^{s+m+n} = a^m ya^r a^t xa^n = a^r a^m yxa^n a^t.$$

So $yx \in S_a$; that is, S_a is reflexive in S.

Let U be a reflexive unitary subsemigroup of S such that U contains a. If x is an arbitrary element of S_a, then $a^t xa^r = a^s \in U$ for some positive integers t, r, s. As U is unitary in S and $a^t, a^r \in U$, we have $x \in U$. Thus $S_a \subseteq U$. So S_a is the minimal reflexive unitary subsemigroup of S containing the element a.

If $S_a \neq S$, for some $a \in S$, then the principal congruence on S determined by S_a is a (non-trivial) group congruence. (See [1].) If $S_a = S$ for all elements a of S, then there is a homomorphism of S onto the additive semigroup of either the integers or the non-negative integers or the positive integers. (See the proof of Theorem 1.2 of [7].) These semigroups have non-trivial group homomorphic image. Thus the theorem is proved. ∎

2. Subdirectly irreducible (m, n)-commutative semigroups

As known, a semigroup S is said to be a subdirect product of semigroups $\{S_\alpha\}_{\alpha \in A}$ if S is isomorphic to a subsemigroup T of the direct product $\Pi_{\alpha \in A} S_\alpha$ for which $\pi_\alpha(T) = S_\alpha$ for all $\alpha \in A$, where π_α is the projection homomorphism of $\Pi_{\alpha \in A} S_\alpha$ to S_α. We say that a semigroup S is subdirectly irreducible if whenever S is written as a subdirect product of a family of semigroups $\{S_\alpha\}_{\alpha \in A}$, then for at least one $\beta \in A$, the projection homomorphism π_β maps S onto S_β isomorphically.

It is well known that every semigroup is a subdirect product of subdirectly irreducible semigroups.

It can be easily verified that every (m,n)-commutative semigroup is a subdirect product of (m,n)-commutative subdirectly irreducible semigroups.

A semigroup which is not subdirectly irreducible is called subdirectly reducible.

The semigroups S and S^0 are simultaneously subdirectly irreducible or reducible ([5]).

By [5], the least non-empty ideal of a semigroup S (if it exists) is called the kernel of S. The kernel of a semigroup with zero is trivial. We call an ideal a non-trivial ideal if it contains at least two elements. The least non-trivial ideal of a semigroup S (if it exists) is called the core of S. If K is the core of a semigroup S, then either $K^2 = K$ or $K^2 = \{0\}$, where 0 denotes the zero of S. In the first case we call K globally idempotent, in the second case K is called nilpotent.

It is known that every non-trivial subdirectly irreducible semigroup has a core. (See, for example, [5].)

As in [5], a semigroup is called a homogroup if it contains a kernel which is a group. By [5], every subdirectly irreducible homogroup without zero is a group.

We note that, by [5], a commutative group is subdirectly irreducible if and only if it is a subgroup of a quasicyclic p-group, where p is a prime.

Theorem 2.1. *A semigroup is a subdirectly irreducible (m,n)-commutative semigroup with a globally idempotent core if and only if it is isomorphic to either G or G^0 or F, where G is a non-trivial subgroup of a quasicyclic p-group (p is a prime) and F is a two-element semilattice.*

Proof. Let S be a subdirectly irreducible (m,n)-commutative semigroup with a globally idempotent core K.

Consider first the case where S has no zero element. Then K is simple and (m,n)-commutative. By Theorem 1.2, K is a commutative group. So S is a homogroup without zero, which implies that S is a commutative group. Thus S is isomorphic to a non-trivial subgroup of a quasicyclic p-group, where p is a prime.

Consider now the case where S has a zero element. Then K is 0-simple. By Theorem 1.1, S is a semilattice of archimedean (m,n)-commutative semigroups. Let S_0 denote the semilattice component of S containing the zero of S. As S_0 is an ideal of S and K is a group with 0 adjoined (Theorem 1.2), $\mid S_0 \mid = 1$.

We show that $S^* = S - \{0\}$ is a subsemigroup of S. Assume, in an indirect way, that there are elements $a, b \in S^*$ such that $ab = 0$. Then

$$A = \{x \in S : \; xb = 0\}$$

is a non-trivial ideal of S (using the fact that S_0 is a semilattice component of S). So $K \subseteq A$; that is, $Kb = \{0\}$.
It can be easily verified that

$$B = \{y \in S : \; Ky = \{0\}\}$$

is a non-trivial ideal of S. So $K \subseteq B$, from which it follows that $K^2 = \{0\}$, contradicting the assumption that K is globally idempotent. Consequently S^* is a subsemigroup of S.

If $\mid S^* \mid = 1$, then S is a two-element semilattice. If $\mid S^* \mid > 1$, then S^* has no zero element. Thus S^* is a subdirectly irreducible (m,n)-commutative semigroup with globally idempotent core and so it is isomorphic to a non-trivial subgroup G of a quasicyclic p-group, where p is a prime. Consequently S is isomorphic to G^0.

As G, G^0 and F are subdirectly irreducible commutative (and so (m,n)-commutative) semigroups with globally idempotent core, the theorem is proved. ■

Before formulating the next theorem, consider some definitions which will be needed later.

If S is a semigroup with zero 0, then let A_S denote the annihilator of S; that is,

$$A_S = \{a \in S : (\forall s \in S)\ as = sa = 0\}.$$

It is evident that A_S is an ideal of S.

As in [5], an element s of a semigroup S is called a disjunctive element if the congruence $C_{\{s\}}$ on S defined by

$$C_{\{s\}} = \{(a,b) \in S \times S : (\forall x, y \in S^1)\ xay = s \Leftrightarrow xby = s\}$$

equals id_S.

Using the concepts of [5], one may easily prove the following theorem.

Theorem 2.2. *An (m,n)-commutative semigroup S with $\mid A_S \mid > 1$ is subdirectly irreducible if and only if it has a non-zero disjunctive element.* ■

Example. Let S be a semigroup defined by the following Cayley table:

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

It can be easily verified that S is an (m,n)-commutative semigroup for all positive integers m and n with $mn \neq 1$, but S is not commutative (not $(1,1)$-commutative). We note that S is a nil semigroup.

The element a is the zero of S and $A_S = \{a,b\}$. The element b is disjunctive. So S is subdirectly irreducible.

Theorem 2.3. *If S is a subdirectly irreducible (m,n)-commutative semigroup such that $\mid A_S \mid = 1$ and the core of S is nilpotent, then S is commutative.*

Proof. Let S be a subdirectly irreducible (m,n)-commutative semigroup such that $\mid A_S \mid = 1$ and the core of S is nilpotent.

Let k_1 be an arbitrary non-zero element of K. As k_1 is not in A_S, $Sk_1 \cup k_1S \cup Sk_1S$ is a non-trivial ideal of S and so $K = Sk_1 \cup k_1S \cup Sk_1S$. Thus $k_1 = ek_1$ or $k_1 = k_1f$ or $k_1 = gk_1h$ for some $e,f,g,h \in S$. If $k_1 = k_1f$, $f \in S$, then $k_1 = k_1f^mf^n = f^nk_1f^m = f^mf^nk_1$. If $k_1 = gk_1h$, $g,h \in S$, then $k_1 = g^{mn}k_1h^{mn} = h^{mn}g^{mn}k_1$. Thus we may consider only the first case.

Assume $k_1 = ek_1$ for some $e \in S$. We note that $k_1 = e^tk_1$ for all positive integers t. Let

$$Z = \{a \in S : (\exists t \in N)\ e^ta = a\}.$$

It is evident that Z is a non-trivial right ideal of S. We show that Z is a two-sided ideal. Let $a \in Z$, $s \in S$ be arbitrary elements. Then $e^ta = a$ for some positive integer t. Then $e^{lt}a = a$ for all positive integers l. Choose the positive integers i and j such that $it, jt \geq m$. Then

$$sa = se^{it+jt}a = (se^{it})e^{jt}a = e^{jt}se^{it}a = e^{jt}sa.$$

Thus $sa \in Z$ and so Z is a (non-trivial) two-sided ideal of S.

As K is the core of S, $K \subseteq Z$. Consequently, for all $k \in K$, there is a positive integer j such that $e^j k = k$.

Define a relation α on S as follows:

$$\alpha = \{(a,b) \in S \times S : \ (\exists j \in N)\ e^j a = e^j b\}.$$

It can be easily verified that α is a right congruence. We show that α is also left compatible. Let x, a, b be arbitrary elements of S with $(a,b) \in \alpha$. Then $e^j a = e^j b$ for some positive integer j. Then $e^r a = e^r b$ for all positive integers $r \geq j$. Let r be a positive integer with $r \geq max\{j, m\}$. Then

$$e^{r+n}xa = e^r(e^n x)a = (e^n x)e^r a = e^n x e^{r-j} e^j a$$

$$= e^n x e^{r-j} e^j b = (e^n x)e^r b = e^r e^n x b = e^{r+n} x b,$$

which means that $(xa, xb) \in \alpha$. Thus α is a congruence on S.

Let k_1 and k_2 be arbitrary elements of K such that $(k_1, k_2) \in \alpha$. Then, for some positive integer j, $e^j k_1 = e^j k_2$ and so $e^t k_1 = e^t k_2$ for all positive integers $t \geq j$. As it was proved above, there are positive integers i_1, i_2 such that $e^{i_1} k_1 = k_1$ and $e^{i_2} k_2 = k_2$. Let t be a positive integer such that $t \geq j$ and $t = i_1 i_2 h$, where h is a positive integer. Then $k_1 = e^t k_1 = e^t k_2 = k_2$ and so the restriction of α to K is the equality relation on K. As S is subdirectly irreducible, α is the equality relation on S.

Let s be an arbitrary element of S. As S is (m,n)-commutative,

$$e^n e^m s = e^m e^n s = e^n s e^m;$$

that is, $(e^m s, s e^m) \in \alpha$, from which it follows that $e^m s = s e^m$. Let $e^m = f$. Then $sf = fs$, for all $s \in S$, and $fk_1 = k_1$.

Let

$$Z^* = \{a \in S : \ fa = a\}.$$

Then Z^* is a non-trivial right ideal of S. As $fsa = sfa = sa$, for all $s \in S$ and $a \in Z^*$, we get that Z^* is a (non-trivial) two-sided ideal of S. Then $K \subseteq Z^*$ and so $fk = k$ for all $k \in K$.

Let

$$\alpha^* = \{(a,b) \in S \times S : \ (\exists i, j \in N)\ f^i a = f^j b\}.$$

It is evident that α^* is a right congruence on S. We show that α^* is also left compatible. Let x, a, b be arbitrary elements of S with $(a,b) \in \alpha^*$. Then $f^i a = f^j b$ for some positive integers i and j. As $f^i xa = xf^i a = xf^j b = f^j xb$, we get $(xa, xb) \in \alpha^*$. So α^* is a congruence on S. As $fk = k$ for all $k \in K$, we have that the restriction of α^* to K is the equality relation on K. As S is subdirectly irreducible, α^* is the equality relation on S.

As $ff^2 = f^2 f$ (that is $(f, f^2) \in \alpha^*$), we get $f^2 = f$. As $fs = f^m f^n s = f^n s f^m = fsf$ (that is $(s, sf) \in \alpha^*$), for all $s \in S$, we get $s = sf$ and so $s = sf = fs$ for all $s \in S$. So f is an identity element of S. Then, for all $a, b \in S$, $ab = af^m f^n b = f^n b a f^m = ba$. Consequently S is a commutative semigroup. Thus the theorem is proved. ■

3. (m,n)-commutative Δ-semigroups

As in [6], a semigroup S is called a Δ-semigroup if the lattice of all congruences on S is a chain with respect to inclusion. By Theorem 9 of [6], if a Δ-semigroup S contains a proper two-sided ideal I then neither S nor I is homomorphic onto a non-trivial group. By Lemma 2 of [6], every homomorphic image of a Δ-semigroup is a Δ-semigroup. By Lemma 3 of [6], a semilattice is a Δ-semigroup if and only if it is of order not greater than 2. Consequently, a Δ-semigroup S is either semilattice indecomposable or a semilattice of two semilattice indecomposable semigroups (Proposition 4 of [6]).

Theorem 3.1. *A semigroup is an (m,n)-commutative semilattice indecomposable Δ-semigroup if and only if it is isomorphic to either*

(i) *a subgroup of a quasicyclic p-group, where p is a prime; or*

(ii) *a non-trivial (m,n)-commutative nil semigroup whose principal ideals form a chain with respect to inclusion.*

Proof. Let S be an (m,n)-commutative semilattice indecomposable Δ-semigroup. Then, by Theorem 1.1, S is archimedean. If S is simple then, by Theorem 1.2, S is a commutative group and so it is isomorphic to a subgroup of a quasicyclic p-group, where p is a prime. If S has a proper two-sided ideal then, by Theorem 9 of [6] and Theorem 1.4 of this paper, S has an idempotent element. By Theorem 1.3, it follows that S is an ideal extension of a commutative group G by an (m,n)-commutative nil semigroup. By Theorem 9 of [6], either $G = S$ or $\mid G \mid = 1$. In the first case S is isomorphic to a non-trivial subgroup of a quasicyclic p-group, where p is a prime. In the second case S is a non-trivial (m,n)-commutative nil semigroup whose principal ideals form a chain with respect to inclusion. (See also Lemma 1.4 of [7].) Thus the first part of the theorem is proved.

The converse statement follows from Lemma 1.4 of [7]. ∎

Theorem 3.2. *A semigroup is an (m,n)-commutative semilattice decomposable Δ-semigroup if and only if it is isomorphic to either*

(i) *G^0, where G is a subgroup of a quasicyclic p-group (p is a prime); or*

(ii) *Q^1, where Q is a non-trivial (m,n)-commutative nil semigroup whose principal ideals form a chain with respect to inclusion.*

Proof. Let us suppose that S is an (m,n)-commutative semilattice decomposable Δ-semigroup. Then, by Proposition 4 of [6] and Theorem 1.1 of this paper, S is a semilattice of two archimedean (m,n)-commutative semigroups S_0 and S_1, $S_0S_1 \subseteq S_0$. By Theorem 1.4 of this paper and Theorem 9 of [6], S_0 has an idempotent element. Then, by Theorem 1.3, S_0 is an ideal extension of a group G by an (m,n)-commutative nil semigroup. By Theorem 9 of [6], $\mid G \mid = 1$ and so S_0 is an (m,n)-commutative nil semigroup.

By Lemma 2 of [6], ${S_1}^0$ is a Δ-semigroup. Thus S_1 is also a Δ-semigroup. Then, by Theorem 3.1, S_1 is either a subgroup of a quasicyclic p-group, where p is a prime, or a non-trivial (m,n)-commutative nil semigroup whose principal ideals form a chain with respect to inclusion.

Assume first that $\mid S_1 \mid = 1$. Let $S_1 = \{e\}$. Then, for all $a \in S$, $ea = e^{m+n-1}a = e^m e^{n-1}a = e^{n-1}ae^m = eae = e^m ae^{n-1} = ae^{n-1}e^m = ae$. So $S = SeS \cup eS \cup Se = eS = Se$, that is, e is a two-sided identity element of S. Consequently S is isomorphic to either a two-element semilattice (and so (i) is satisfied) or Q^1, where Q is a non-trivial (m,n)-commutative nil semigroup whose principal ideals form a chain with respect to inclusion (and so (ii) is satisfied).

Assume now that $|S_1| > 1$. If S_1 is a (non-trivial) subgroup of a quasicyclic p-group (p is a prime) then, by the proof of Theorem 3.5 of [7], $|S_0| = 1$. Thus $S = S_1{}^0$ and so (i) is satisfied. If S_1 is a (non-trivial) nil semigroup then the Rees congruence determined by the ideal consisting of the zero elements of S_0 and S_1 is not comparable with the least semilattice congruence on S. But this is a contradiction. Thus the first part of the theorem is proved.

As the converse statement can be proved easily, the theorem is proved.■

We summarize results of Theorem 3.1 and Theorem 3.2.

Theorem 3.3. *A semigroup is an (m,n)-commutative Δ-semigroup if and only if it is isomorphic to either G or G^0 or Q or Q^1, where G is a subgroup of a quasicyclic p-group (p is a prime) and Q is a non-trivial (m,n)-commutative nil semigroup whose principal ideals form a chain with respect to inclusion.* ■

References

[1] Clifford, A.H. and G.B. Preston, "*The algebraic theory of semigroups*," Amer. Math. Soc., Providence, Volume I (1961), Volume II (1967).

[2] Lajos, S., *Fibonacci characterizations and (m,n)-commutativity in semigroup theory*, Pure Math. Appl., Ser. A, **1** (1990), 59-65.

[3] Petrich, M., "*Lectures in Semigroups*," Akademie-Verlag-Berlin, 1977.

[4] Putcha, M.S., *Semilattice decomposition of semigroups*, Semigroup Forum **6** (1973), 12–34.

[5] Schein, B.M., *Homomorphisms and subdirect decompositions of semigroups*, Pacific Journal of Math. **17** (1966), 529–547.

[6] Tamura, T., *Commutative semigroups whose lattice of congruences is a chain*, Bull. Soc. Math. France, **97** (1969), 369–380.

[7] Trotter, P.G., *Exponential Δ-semigroups*, Semigroup Forum **12** (1976), 313–331.

Department of Mathematics
Transport Engineering Faculty
Technical Univ. of Budapest
1111 Budapest, Egry J. u. 20-22
HUNGARY

Received May 30, 1990
and in final form October 15, 1990

Semigroup Forum Vol. 45 (1992) 191–204

RESEARCH ARTICLE

Semigroups of $\mathcal{I}$-density continuous functions

Krzysztof Ciesielski[1], **Lee Larson**[2], **Krzysztof Ostaszewski**[2]

Communicated by M. W. Mislove

Abstract

This paper is concerned with the classes of $\mathcal{I}$-density continuous functions and deep-$\mathcal{I}$-density continuous functions as semigroups with composition as the operation. We also analyze some of their subsemigroups. It is shown that the groups of automorphisms of these semigroups and several of their subsemigroups have the inner automorphism property.

1. Preliminaries

The notation used throughout this paper is standard. In particular, R stands for the set of real numbers and $\mathbb{N} = \{1, 2, 3, \ldots\}$. The symmetric difference of sets A and B is denoted by $A \triangle B$ and the complement of a set $A \subset \mathbb{R}$ by A^c. The symbols $\mathcal{L}$ and $\mathcal{B}$ stand for the families of subsets of R which are Lebesgue measurable and have the Baire property, respectively. $\mathcal{N}$ and $\mathcal{I}$ denote the ideals of Lebesgue measure zero and of first category subsets of R. If a statement is true everywhere except for those points of a set belonging to $\mathcal{N}$, then it is true *almost everywhere (a.e.)*. If it is true everywhere except for the points of a set belonging to $\mathcal{I}$, then it is true *$\mathcal{I}$-almost everywhere ($\mathcal{I}$-a.e.)*. If $A \in \mathcal{L}$, we denote its Lebesgue measure by $\mathrm{m}(A)$.

The natural topology on R is denoted by $\mathcal{T}_{\mathcal{O}}$. The symbols $\mathrm{int}(A)$ and $\mathrm{cl}(A)$ stand for the interior and closure of $A \subset \mathbb{R}$ with respect to $\mathcal{T}_{\mathcal{O}}$. It is also easy to see that for any set $B \in \mathcal{B}$ there is a unique open set $\tilde{B}$ such that $\tilde{B} = \mathrm{int}\left(\mathrm{cl}\left(\tilde{B}\right)\right)$ and $B = \tilde{B} \triangle I$ for some $I \in \mathcal{I}$ [18]. Any open set G for which $G = \mathrm{int}(\mathrm{cl}(G))$ is called a *regular* open set. In a sense, $\tilde{B}$ is the largest open set such that B can be written as $\tilde{B} \triangle I$ for some $I \in \mathcal{I}$.

Any of the sets $\bigcup_{n \in \mathbb{N}} [a_n, b_n]$ or $\bigcup_{n \in \mathbb{N}} (a_n, b_n)$ is a *right interval set* at 0 if $0 < b_{n+1} < a_n < b_n$ for $n \in \mathbb{N}$ and $\lim_{n \to \infty} a_n = 0$. A *left interval set* at 0 is defined similarly. A set E is a right (left) interval set at $x \in \mathbb{R}$ if $E - x$ is a right (left, respectively) interval set at 0.

For a topological space $(X, \mathcal{T})$ let $\mathrm{S}(X)$ or $\mathrm{S}((X, \mathcal{T}))$ be the semigroup of all continuous selfmaps of X; i.e., continuous functions $f : X \to X$, with composition as the operation. If there is no topology defined on X, then the discrete topology is used in the definition of $\mathrm{S}(X)$ so that $\mathrm{S}(X) = X^X$.

If G is a semigroup, then $\mathrm{Aut}(G)$ denotes the group of all automorphisms of G. A subsemigroup G of $\mathrm{S}(X)$ has the *inner automorphism property* if for every automorphism $\Psi \in \mathrm{Aut}(G)$ there is an $h \in G$ such that $\Psi(g) = h \circ g \circ h^{-1}$ for every $g \in G$.

[1] This author was partially supported by West Virginia University Senate Research Grant.

[2] These authors were partially supported by University of Louisville Arts and Sciences Research grants.

For a topological space X the property that $S(X)$ has the inner automorphism property is strongly related to the following definitions and facts.

A collection of topological spaces is *S–admissible* if for each pair of spaces X and Y from the collection, any isomorphism $\Phi : S(X) \to S(Y)$ is induced by a homeomorphism $h : X \to Y$; i.e., $\Phi(f) = h \circ f \circ h^{-1}$ for all $f \in S(X)$. In other words, within an S–admissible class of spaces, X is homeomorphic to Y if, and only if, $S(X)$ is isomorphic to $S(Y)$, so that the topological structure of the space X is fully characterized by the algebraic structure of $S(X)$. It is easy to see that for a topological space X a necessary and sufficient condition for belonging to an S–admissible class is that the semigroup $S(X)$ has the inner automorphism property. This gives a motivation for studying the topological spaces X for which $S(X)$ has the inner automorphism property. For more information on the subject see Magill [14].

The basis for studying the inner automorphism property of subsemigroups G of $S(X)$ is given by the following theorem of Schreier [21].

Proposition 1.1. *Let X be a set and let G be a subsemigroup of $S(X)$ such that every constant mapping is in G. Then, for every $\Psi \in \mathrm{Aut}(G)$ there exists a unique bijection h of X such that $\Psi(g) = h \circ g \circ h^{-1}$ for every $g \in G$.*

Every semigroup considered in this paper contains all constant functions. In particular, if Ψ is an automorphism of a semigroup $G \subset S(X)$ containing all constant functions then the unique bijection h of X for which $\Psi(g) = h \circ g \circ h^{-1}$ for all $g \in G$ will be called the *generating bijection* of Ψ.

Following Magill [14, Definition 2.2, p. 198] we say that a topological space X is *generated* if it is T_1 and the collection of complements of level sets for its continuous selfmaps $\{(f^{-1}(\{x\}))^c : x \in X \text{ and } f \in S(X)\}$ forms a subbase for X. (Compare also [22].) It is known that the class of all generated spaces is S–admissible [14, Theorem 2.3, p. 198]. In particular, there is the following.

Proposition 1.2. *If a topological space X is generated then, $S(X)$ has the inner automorphism property.*

2. Topologies

This paper is concerned with the semigroups $\mathcal{C}_{\mathcal{II}} = S((\mathbb{R}, \mathcal{T}_{\mathcal{I}}))$ of *$\mathcal{I}$-density continuous functions*, $\mathcal{C}_{\mathcal{DD}} = S((\mathbb{R}, \mathcal{T}_{\mathcal{D}}))$ of *deep-$\mathcal{I}$-density continuous functions* and $\mathcal{C}_{\mathcal{NN}} = S((\mathbb{R}, \mathcal{T}_{\mathcal{N}}))$ of *density continuous functions* and some of their subsemigroups. We start with definitions of the topologies $\mathcal{T}_{\mathcal{N}}$, $\mathcal{T}_{\mathcal{I}}$ and $\mathcal{T}_{\mathcal{D}}$. For this, we first introduce the notions of a density point, $\mathcal{I}$-density point and deep-$\mathcal{I}$-density point. (See [18, 19, 23]. For discussion, compare also [6].)

Let $A \in \mathcal{L}$. A number x, not necessarily in A, is a *density point* of A if

$$\lim_{h \to 0^+} \frac{\mathrm{m}(A \cap (x-h, x+h))}{2h} = 1. \tag{1}$$

The set of all density points of $A \in \mathcal{L}$ we denote by $\Phi_{\mathcal{N}}(A)$. It is a straightforward consequence of the Lebesgue density theorem that for every $A \in \mathcal{L}$, $\Phi_{\mathcal{N}}(A) \triangle A \in \mathcal{N}$. The family of sets

$$\mathcal{T}_{\mathcal{N}} = \{A \in \mathcal{L} : A \subset \Phi_{\mathcal{N}}(A)\}$$

forms a topology on R [18, 10] called the *density topology*.

If $A \in \mathcal{B}$, we say that 0 is an *$\mathcal{I}$-density point* of A if for every increasing sequence $\{n_m\}_{m\in\mathbb{N}}$ of natural numbers there exists a subsequence $\{n_{m_p}\}_{p\in\mathbb{N}}$ such that

$$\lim_{p\to\infty} \chi_{n_{m_p}A\cap(-1,1)} = \chi_{(-1,1)} \ \mathcal{I}\text{-a.e.} \tag{2}$$

We say that a point a is an $\mathcal{I}$-density point of $A \in \mathcal{B}$ if 0 is an $\mathcal{I}$-density point of $A-a$. The set of all $\mathcal{I}$-density points of $A \in \mathcal{B}$ we denote by $\Phi_{\mathcal{I}}(A)$. Similar to the case with Lebesgue density, as noted above, $A\triangle\Phi_{\mathcal{I}}(A) \in \mathcal{I}$ for every $A \in \mathcal{B}$ [23, Theorem 3] and

$$\Phi_{\mathcal{I}}(A) = \Phi_{\mathcal{I}}(B) \quad \text{for every } A, B \in \mathcal{B} \text{ such that } \ A\triangle B \in \mathcal{I}. \tag{3}$$

The family of sets

$$\mathcal{T}_{\mathcal{I}} = \{A \in \mathcal{B} : A \subset \Phi_{\mathcal{I}}(A)\}$$

forms a topology on R [19, 23] called the *$\mathcal{I}$-density topology*.

Finally, we say that a point $a \in \mathbb{R}$ is a *deep-$\mathcal{I}$-density point* [23] of an $A \in \mathcal{B}$ if there exists a closed set $F \subset A \cup \{a\}$ such that a is an $\mathcal{I}$-density point of F. The set of all deep-$\mathcal{I}$-density points of $A \in \mathcal{B}$ is denoted $\Phi_{\mathcal{D}}(A)$. The family of sets

$$\mathcal{T}_{\mathcal{D}} = \{A \in \mathcal{B} : A \subset \Phi_{\mathcal{D}}(A)\}$$

forms a topology on R [11, 23] called the *deep-$\mathcal{I}$-density topology*.

The inclusion relations between the topologies $\mathcal{T}_{\mathcal{O}}$, $\mathcal{T}_{\mathcal{N}}$, $\mathcal{T}_{\mathcal{I}}$ and $\mathcal{T}_{\mathcal{D}}$ are given by the following theorem [8, Theorem 2.5].

Theorem 2.1. *If $\mathcal{P}(\mathbb{R})$ stands for the discrete topology on R then*

$$\begin{array}{ccccccc} \mathcal{T}_{\mathcal{O}}\cap\mathcal{T}_{\mathcal{N}} & \subset & \mathcal{T}_{\mathcal{D}}\cap\mathcal{T}_{\mathcal{N}} & \subset & \mathcal{T}_{\mathcal{I}}\cap\mathcal{T}_{\mathcal{N}} & \subset & \mathcal{T}_{\mathcal{N}} \\ \| & & \cap & & \cap & & \cap \\ \mathcal{T}_{\mathcal{O}} & \subset & \mathcal{T}_{\mathcal{D}} & \subset & \mathcal{T}_{\mathcal{I}} & \subset & \mathcal{P}(\mathbb{R}) \end{array}$$

Moreover, all the inclusions are proper.

We also need the following fact. For the $\mathcal{I}$-density case it can be found in [4, Lemma 2.4], [20, Theorem 1] or [23, Theorem 2].

Proposition 2.2. *If $E = \bigcup_{n\in\mathbb{N}}[a_n, b_n]$ is a right interval set such that*

(i) $\lim_{n\to\infty}(b_n - a_n)/a_n = 0$*; and*

(ii) $\lim_{n\to\infty} b_{n+1}/a_n = 0$,

then 0 is a density and deep-$\mathcal{I}$-density point of E^c. In particular, $E^c \in \mathcal{T}_{\mathcal{N}} \cap \mathcal{T}_{\mathcal{D}}$.

3. Density Continuous Functions

Let Δ be the class of all differentiable functions from R to R. Moreover, let $\Delta^{(\mathcal{N})}$ stand for the class of all approximately differentiable functions ([2], see also [17]) and let $\Delta^{(\mathcal{N})}_{\mathcal{N}}$ be the class of all almost everywhere approximately differentiable functions. Often used is the group of homeomorphisms $f : (\mathbb{R}, \mathcal{T}_{\mathcal{O}}) \to (\mathbb{R}, \mathcal{T}_{\mathcal{O}})$, which is written as $\mathcal{H}$.

The classes Δ, $\Delta^{(\mathcal{N})}$, $\Delta^{(\mathcal{N})}_{\mathcal{N}}$ are connected to the class $\mathcal{C}_{\mathcal{N}\mathcal{N}}$ of density continuous functions by the following theorem.

Theorem 3.1. *If $\mathcal{F}_m$ stands for the class of measurable functions, then*

$$\begin{array}{ccccccc}
\Delta & \subset & \Delta^{(\mathcal{N})} & \subset & \Delta^{(\mathcal{N})}_{\mathcal{N}} & \subset & \mathcal{F}_m \\
\cup & & \cup & & \cup & & \cup \\
\Delta \cap C_{\mathcal{N}\mathcal{N}} & \subset & \Delta^{(\mathcal{N})} \cap C_{\mathcal{N}\mathcal{N}} & \subset & \Delta^{(\mathcal{N})}_{\mathcal{N}} \cap C_{\mathcal{N}\mathcal{N}} & \subset & C_{\mathcal{N}\mathcal{N}}
\end{array}$$

and all the inclusions are proper.

Proof. The inclusions are obvious. The vertical inclusions are proper because there is a C^∞ function which is not density continuous [7, Example 1].

$C_{\mathcal{N}\mathcal{N}} \not\subset \Delta^{(\mathcal{N})}_{\mathcal{N}}$ follows immediately from [9, Theorem 4].

$\Delta^{(\mathcal{N})}_{\mathcal{N}} \cap C_{\mathcal{N}\mathcal{N}} \not\subset \Delta^{(\mathcal{N})}$ is proved by the function $h(x) = |x|$.

$\Delta^{(\mathcal{N})} \cap C_{\mathcal{N}\mathcal{N}} \not\subset \Delta$ is easily justified by the function f defined by $f(x) = 0$ for $x \in E^c$ and $f(x) = (x - a_n)^2(x - b_n)^2$ for $x \in [a_n, b_n]$ and $n \in \mathbb{N}$, where $E = \bigcup_{n\in\mathbb{N}}[a_n, b_n]$ is a right interval set from Proposition 2.2. ∎

The motivation to study the inner automorphism property of the classes from Theorem 3.1 arises from the following theorem of Magill [13].

Theorem 3.2. *The semigroup Δ of all differentiable functions has the inner automorphism property.*

Following this path Ostaszewski proved the following [17].

Theorem 3.3. *The semigroups $C_{\mathcal{N}\mathcal{N}}$, $\Delta^{(\mathcal{N})}_{\mathcal{N}} \cap C_{\mathcal{N}\mathcal{N}}$, $\Delta^{(\mathcal{N})} \cap C_{\mathcal{N}\mathcal{N}}$ and $\Delta \cap C_{\mathcal{N}\mathcal{N}}$ have the inner automorphism property.*

Notice that the above theorem cannot be deduced from Proposition 1.2, because the density topology is not generated [5]. In fact, the density topology is the only known example of completely regular not generated topological space for which the semigroup of continuous selfmaps has the inner automorphism property. (Also note the remark after Theorem 5.5.)

Theorem 3.3 is proved by showing the following facts: for every generating bijection h of $\Psi \in \mathrm{Aut}(C_{\mathcal{N}\mathcal{N}}) \cup \mathrm{Aut}(\Delta^{(\mathcal{N})}_{\mathcal{N}} \cap C_{\mathcal{N}\mathcal{N}})$ we have $h, h^{-1} \in \mathcal{H} \cap C_{\mathcal{N}\mathcal{N}} \cap \Delta^{(\mathcal{N})}$ and for every generating bijection h of $\Psi \in \mathrm{Aut}(\Delta^{(\mathcal{N})} \cap C_{\mathcal{N}\mathcal{N}}) \cup \mathrm{Aut}(\Delta \cap C_{\mathcal{N}\mathcal{N}})$ we have $h, h^{-1} \in \mathcal{H} \cap C_{\mathcal{N}\mathcal{N}} \cap \Delta$. In fact, if we identify the group $\mathrm{Aut}(H)$ with the set of generating bijections, we obtain

Corollary 3.4. *The following relations hold:*

$$\mathrm{Aut}(\Delta \cap C_{\mathcal{N}\mathcal{N}}) = \mathrm{Aut}(\Delta^{(\mathcal{N})} \cap C_{\mathcal{N}\mathcal{N}}) \subset \mathrm{Aut}(\Delta^{(\mathcal{N})}_{\mathcal{N}} \cap C_{\mathcal{N}\mathcal{N}}) = \mathrm{Aut}(C_{\mathcal{N}\mathcal{N}})$$

and the inclusion is proper.

Theorems 3.2 and 3.3 prove that the classes Δ, $\Delta \cap C_{\mathcal{N}\mathcal{N}}$, $\Delta^{(\mathcal{N})} \cap C_{\mathcal{N}\mathcal{N}}$, $\Delta^{(\mathcal{N})}_{\mathcal{N}} \cap C_{\mathcal{N}\mathcal{N}}$ and $C_{\mathcal{N}\mathcal{N}}$ have the inner automorphism property. For the remaining three classes the question about their inner automorphism property is moot as those classes do not form semigroups. For the class $\Delta^{(\mathcal{N})}$, the proof is given in Example 4.11. For the two other classes the result follows from the example below.

Example 3.5. There exist almost everywhere differentiable functions f and g such that $g \circ f$ is not measurable. In particular, the classes $\Delta^{(\mathcal{N})}_{\mathcal{N}}$ and $\mathcal{F}_m$ are not closed under composition.

Proof. Let $P \subset \mathbb{R}$ be any perfect nowhere dense set with positive measure, $S \subset P$ be nonmeasurable and let C be a Cantor set of measure 0. Let f be a homeomorphism such that $f^{-1}(C) = P$ and let $g = \chi_S \circ f^{-1}$. Then, f is differentiable almost everywhere, as a homeomorphism and g is differentiable almost everywhere, as $g = \chi_{f(S)}$ and $f(S) \subset C$ has measure 0. On the other hand, $g \circ f = \chi_S \circ f^{-1} \circ f = \chi_S$ is not measurable, because S is nonmeasurable. ■

4. $\mathcal{I}$-approximate Derivative

In this section we introduce the notion of the $\mathcal{I}$-approximate derivative as category analogue of the approximate derivative. Let us recall that a function $f: \mathbb{R} \to \mathbb{R}$ is $\mathcal{I}$-approximately continuous (deep-$\mathcal{I}$-approximately continuous) if f is continuous with respect to the ordinary topology on the range and $\mathcal{I}$-density (deep-$\mathcal{I}$-density) topology on the domain. It is well known that the classes of $\mathcal{I}$-approximately continuous functions and deep-$\mathcal{I}$-approximately continuous functions coincide [11, Theorem 2].

We start with the following definitions. A function $f: \mathbb{R} \to \mathbb{R}$ is said to be *$\mathcal{I}$-approximately differentiable at a point x* if there exists a number $D^{(\mathcal{I})} f(x)$, called the *$\mathcal{I}$-approximate derivative* of f at x, such that for every $\varepsilon > 0$, x is an $\mathcal{I}$-density point of some Baire subset of

$$\left\{ t \in \mathbb{R}: \frac{f(t) - f(x)}{t - x} \in \left(D^{(\mathcal{I})} f(x) - \varepsilon, D^{(\mathcal{I})} f(x) + \varepsilon\right) \right\}. \tag{4}$$

(Compare also [12] and [23, Definition 8].)

A function $f: \mathbb{R} \to \mathbb{R}$ is said to be *$\mathcal{I}$-approximately differentiable* if it is $\mathcal{I}$-approximately differentiable at every point. The class of all $\mathcal{I}$-approximately differentiable functions is denoted by $\Delta^{(\mathcal{I})}$.

A function $f: \mathbb{R} \to \mathbb{R}$ is said to be *$\mathcal{I}$-approximately differentiable $\mathcal{I}$-almost everywhere* if there exists a set $A \in \mathcal{I}$ such that f is $\mathcal{I}$-approximately differentiable at every point of A^c. The class of all functions $\mathcal{I}$-approximately differentiable $\mathcal{I}$-almost everywhere is denoted by $\Delta^{(\mathcal{I})}_{\mathcal{I}}$.

Similarly, the classes of functions that are deep-$\mathcal{I}$-approximately or $\mathcal{I}$-approximately continuous $\mathcal{I}$-almost everywhere may be defined.

It is easy to introduce the idea of a function $f: \mathbb{R} \to \mathbb{R}$ being deep-$\mathcal{I}$-approximately differentiable at point x by requiring that x is a deep-$\mathcal{I}$-density point of a set from (4). However it can be proved that for the $\mathcal{I}$-approximately continuous functions the notions of $\mathcal{I}$-approximate differentiability and deep-$\mathcal{I}$-approximate differentiability at a point coincide. Thus, there is no good reason to pursue this notion.

We start our investigation with the following propositions.

Proposition 4.1. *If a function $f: \mathbb{R} \to \mathbb{R}$ is $\mathcal{I}$-approximately differentiable at x then f is $\mathcal{I}$-approximately continuous at x.*

Proof. Choose $\varepsilon > 0$. We must prove that x is an $\mathcal{I}$-density point of the set

$$\{t \in \mathbb{R}: |f(t) - f(x)| \in (-\varepsilon, \varepsilon)\}.$$

But, if we choose $\delta > 0$ such that $\delta\left(D^{(\mathcal{I})} f(x) - \varepsilon, D^{(\mathcal{I})} f(x) + \varepsilon\right) \subset (-\varepsilon, \varepsilon)$, then

$$\{t \in \mathbb{R}: |f(t) - f(x)| \in (-\varepsilon, \varepsilon)\} \supset$$

$$\{t \in \mathbb{R}: |t-x| < \delta \ \& \ |f(t)-f(x)| \in \delta(D^{(\mathcal{I})}f(x)-\varepsilon, D^{(\mathcal{I})}f(x)+\varepsilon)\} \supset$$

$$\{t \in \mathbb{R}: |t-x| < \delta \ \& \ |f(t)-f(x)| \in |t-x|(D^{(\mathcal{I})}f(x)-\varepsilon, D^{(\mathcal{I})}f(x)+\varepsilon)\} =$$

$$\left\{t \in \mathbb{R}: \frac{|f(t)-f(x)|}{|t-x|} - D^{(\mathcal{I})}f(x) \in (-\varepsilon,\varepsilon)\right\} \cap (x-\delta, x+\delta)$$

and, by assumption, x is an $\mathcal{I}$-density point of some Baire subset of the last set. This finishes the proof. ∎

Proposition 4.2. *If $f:\mathbb{R}\to\mathbb{R}$ is $\mathcal{I}$-approximately differentiable $\mathcal{I}$-almost everywhere, then f is a Baire function.*

Proof. Poreda, Wagner-Bojakowska and Wilczyński proved that a function $f:\mathbb{R}\to\mathbb{R}$ is $\mathcal{I}$-approximately continuous $\mathcal{I}$-almost everywhere if, and only if, f has the Baire property [19, Theorem 7]. Use of Proposition 4.1 completes the proof. ∎

In the sequel we also need the following two examples.

Example 4.3. For every $a < b \le c < d$ there exists a differentiable $\mathcal{I}$-density continuous and density continuous function $f:\mathbb{R}\to[0,1]$ such that $f(x)=0$ for $x \in (a,d)^c$ and $f(x)=1$ for $x\in[b,c]$.

Proof. We have to define f only on (a,b) and (c,d). So, define

$$f(x) = \frac{(x-a)^2(x-2b+a)^2}{(b-a)^4}$$

for $x \in (a,b)$ and

$$f(x) = \frac{(x-d)^2(x-2c+d)^2}{(c-d)^4}$$

for $x\in(c,d)$. It is easy to see that f is differentiable and, because it is everywhere unilaterally a polynomial, it is also density and $\mathcal{I}$-density continuous [7, Corollary 3], [3]. ∎

The previous example easily implies the existence of the following example.

Example 4.4. There exists an approximately and $\mathcal{I}$-approximately differentiable, density continuous and $\mathcal{I}$-density continuous function which is not continuous. In particular,

$$\Delta^{(\mathcal{N})} \cap \Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{N}\mathcal{N}} \cap \mathcal{C}_{\mathcal{I}\mathcal{I}} \not\subset \mathcal{C}.$$

Proof. Let $E = \bigcup_{n\in\mathbf{N}}[a_n,d_n]$ be an interval set as in Proposition 2.2. For each n, let f_n be as in Example 4.3 with $[a,d]=[a_n,d_n]$. Define $f(x)=0$ on E^c and $f=f_n$ on $[a_n,d_n]$. The rest is obvious. ∎

Now we are ready to prove the next theorem.

Theorem 4.5. *If $\mathcal{F}_c$ stands for the Baire functions, then*

$$\begin{array}{ccccccc}
\Delta & \subset & \Delta^{(\mathcal{I})} & \subset & \Delta_{\mathcal{I}}^{(\mathcal{I})} & \subset & \mathcal{F}_c \\
\cup & & \cup & & \cup & & \cup \\
\Delta \cap \mathcal{C}_{\mathcal{DD}} & \subset & \Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}} & \subset & \Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}} & \subseteq & \mathcal{C}_{\mathcal{DD}} \\
\cup & & \cup & & \cup & & \cup \\
\Delta \cap \mathcal{C}_{\mathcal{II}} & \subset & \Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}} & \subset & \Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}} & \subseteq & \mathcal{C}_{\mathcal{II}}
\end{array}$$

and all the inclusions with the symbol $\subset$ are proper.

Proof. The inclusion $\mathcal{C}_{\mathcal{II}} \subset \mathcal{C}_{\mathcal{DD}}$ can be found in [8, Theorem 4.1(iv)]. The other inclusions are obvious.

The vertical inclusions are proper because $\Delta \cap \mathcal{C}_{\mathcal{DD}} \not\subset \mathcal{C}_{\mathcal{II}}$ [8, Example 6.7] and $\Delta \not\subset \mathcal{C}_{\mathcal{DD}}$ [8, Example 5.7].

The fact that the horizontal inclusions of the form $\subset$ are proper follows from $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}} \not\subset \Delta$ (Example 4.4) and $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}} \not\subset \Delta^{(\mathcal{I})}$ which is proved by $f(x) = |x|$.

This finishes the proof. ∎

It remains open whether the inclusions of the form $\subseteq$ in the previous theorem are proper. In particular, a positive answer to this question would follow from a positive answer for the following problem.

Problem 4.6. Does there exist an $\mathcal{I}$-density continuous function which is nowhere $\mathcal{I}$-approximately differentiable?

To prove the next theorem, the following two lemmas are needed. The first one is a version of the chain rule. The analogous theorem for approximate derivatives can be found in [16].

Lemma 4.7. *If $f, g\colon \mathbb{R} \to \mathbb{R}$ are such that $\Delta^{(\mathcal{I})} f(x_0)$ and $\Delta^{(\mathcal{I})} g(f(x_0))$ exist then $\Delta^{(\mathcal{I})}(g \circ f)(x_0)$ also exists and*

$$\Delta^{(\mathcal{I})}(g \circ f)(x_0) = \Delta^{(\mathcal{I})} g(f(x_0))\, \Delta^{(\mathcal{I})} f(x_0). \tag{5}$$

Proof. First assume that $\Delta^{(\mathcal{I})} g(f(x_0)) > 0$ and $\Delta^{(\mathcal{I})} f(x_0) > 0$. Put $z_0 = f(x_0)$ and let u_1, u_2, v_1, v_2 be arbitrary positive numbers such that $0 < u_1 < \Delta^{(\mathcal{I})} f(x_0) < u_2$ and $0 < v_1 < \Delta^{(\mathcal{I})} g(z_0) < v_2$. Then, there exist sets $U, V \in \mathcal{T}_{\mathcal{I}}$, $x_0 \in U$, $z_0 \in V$, such that

$$u_1 \le \frac{f(x) - f(x_0)}{x - x_0} \le u_2 \tag{6}$$

for all $x \in U$ and

$$v_1 \le \frac{g(z) - g(z_0)}{z - z_0} \le v_2 \tag{7}$$

for all $z \in V$. Since, by Proposition 4.1, f is $\mathcal{I}$-density continuous at x_0, there exists $W \in \mathcal{T}_{\mathcal{I}}$ such that $x_0 \in W \subset f^{-1}(V)$. In particular, by (7), for $x \in W$ we have $z = f(x) \in V$ and

$$v_1 \le \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \le v_2. \tag{8}$$

Multiplying (6) by (8) we obtain

$$u_1v_1 < \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0} < u_2v_2;$$

i.e.,

$$u_1v_1 < \frac{g(f(x)) - g(f(x_0))}{x - x_0} < u_2v_2,$$

for every $x \in U \cap W \in \mathcal{T}_{\mathcal{I}}$. But,

$$u_1v_1 < \Delta^{(\mathcal{I})} g(f(x_0))\, \Delta^{(\mathcal{I})} f(x_0) < u_2v_2$$

where the numbers u_1v_1 and u_2v_2 can be choosen as close to the number $\Delta^{(\mathcal{I})} g(f(x_0))$ $\Delta^{(\mathcal{I})} f(x_0)$ as we wish. This implies (5).

The remaining cases, when $\Delta^{(\mathcal{I})} g(f(x_0)) \le 0$ or $\Delta^{(\mathcal{I})} f(x_0) \le 0$, are very similar, modulo some little technical problems with the signs of the inequalities. This finishes the proof. ∎

Lemma 4.8. *If $f \in C_{\mathcal{I}\mathcal{I}} \cap \Delta_{\mathcal{I}}^{(\mathcal{I})}$ and $g \in \Delta_{\mathcal{I}}^{(\mathcal{I})}$, then $g \circ f \in \Delta_{\mathcal{I}}^{(\mathcal{I})}$.*

Proof. Let $A, B \in \mathcal{I}$ be such that f is $\mathcal{I}$-approximately differentiable on A^c and g is $\mathcal{I}$-approximately differentiable on B^c.

By Lemma 4.7 the function $g \circ f$ is $\mathcal{I}$-approximately differentiable on the set $D = A^c \cap f^{-1}(B^c)$. Moreover, for every x, if $E_x = f^{-1}(x)$ then $g \circ f$ is $\mathcal{I}$-approximately differentiable on the set $E_x \cap \tilde{E}_x \in \mathcal{T}_{\mathcal{I}}$, as $g \circ f$ is constant on this set. Thus, if $E = \bigcup_{x \in \mathbb{R}} (E_x \cap \tilde{E}_x)$, then $g \circ f$ is $\mathcal{I}$-approximately differentiable on $D \cup E$.

To finish the proof it is enough to prove that $(D \cup E)^c \in \mathcal{I}$; in other words that every x is an $\mathcal{I}$-density point of $D \cup E$.

So, let $x \in \mathbb{R}$. Then, $f(x)$ is an $\mathcal{I}$-density point of $B^c \cup \{f(x)\}$, because B is discrete in $\mathcal{T}_{\mathcal{I}}$. So, since f is $\mathcal{I}$-density continuous, x is an $\mathcal{I}$-density point of

$$\begin{aligned} f^{-1}(B^c \cup \{f(x)\}) &= f^{-1}(B^c) \cup f^{-1}(f(x)) \\ &= f^{-1}(B^c) \cup E_{f(x)} \\ &= f^{-1}(B^c) \cup (E_{f(x)} \cap \tilde{E}_{f(x)}) \cup (E_{f(x)} \setminus \tilde{E}_{f(x)}). \end{aligned}$$

But $E_{f(x)} \setminus \tilde{E}_{f(x)} \in \mathcal{I}$ and $A \in \mathcal{I}$. Hence, x is an $\mathcal{I}$-density point of $A^c \cap (f^{-1}(B^c) \cup (E_{f(x)} \cap \tilde{E}_{f(x)})) \subset D \cup E$. This finishes the proof. ∎

As a corollary we obtain

Theorem 4.9. *The classes Δ, $C_{\mathcal{I}\mathcal{I}}$, $C_{\mathcal{D}\mathcal{D}}$, $\Delta \cap C_{\mathcal{I}\mathcal{I}}$, $\Delta \cap C_{\mathcal{D}\mathcal{D}}$, $\Delta^{(\mathcal{I})} \cap C_{\mathcal{I}\mathcal{I}}$, $\Delta^{(\mathcal{I})} \cap C_{\mathcal{D}\mathcal{D}}$ and $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap C_{\mathcal{I}\mathcal{I}}$ are closed under composition. In particular, they form semigroups.*

Proof. The classes Δ, $C_{\mathcal{I}\mathcal{I}}$ and $C_{\mathcal{D}\mathcal{D}}$ are obviously closed under composition. Therefore the same is true of the classes $\Delta \cap C_{\mathcal{I}\mathcal{I}}$ and $\Delta \cap C_{\mathcal{D}\mathcal{D}}$.

For the classes $\Delta^{(\mathcal{I})} \cap C_{\mathcal{I}\mathcal{I}}$ and $\Delta^{(\mathcal{I})} \cap C_{\mathcal{D}\mathcal{D}}$ the conclusion follows immediately from Lemma 4.7, while for the class $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap C_{\mathcal{I}\mathcal{I}}$ it follows from Lemma 4.8. ∎

Problem 4.10. Are the classes $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}}$ and $\Delta_{\mathcal{I}}^{(\mathcal{I})}$ closed under composition? If so, do they have the inner automorphism property?

We finish this section by showing that the remaining classes $\Delta^{(\mathcal{I})}$ and $\mathcal{F}_c$ of Theorem 4.5 are not closed under composition.

Example 4.11. There exists a differentiable function f and an approximately and $\mathcal{I}$-approximately differentiable function g such that $g \circ f$ is neither approximately nor $\mathcal{I}$-approximately continuous. In particular, the classes $\Delta^{(\mathcal{N})}$ and $\Delta^{(\mathcal{I})}$ are not closed under composition.

Proof. Let $D = \bigcup_{n\in\mathbb{N}}(p_n, q_n) = (0,\infty) \setminus E$ where E is from Proposition 2.2. Let us define $h:\mathbb{R} \to \mathbb{R}$ by putting $h(x) = 0$ for $x \in D^c$ and

$$h(x) = u_n(x - p_n)(x - q_n)$$

for $x \in (p_n, q_n)$, where $u_n > 0$. Define f by $f(x) = \int_0^x h(y)\,dy$. It is easy to see that f is differentiable and constant on each interval contained in D^c. Decreasing the constants u_n, if necessary, we may assume that $\lim_{n\to\infty} f(p_{n+1})/f(p_n) = 0$.

Let us choose intervals (a_n, b_n) centered at $f(p_n)$ such that the assumptions from Proposition 2.2 are satisfied and let $E = \bigcup_{n\in\mathbb{N}}[a_n, b_n]$. Then, $E^c \in \mathcal{T}_{\mathcal{N}} \cap \mathcal{T}_{\mathcal{D}}$. Let us define $g(x) = 0$ for $x \in E^c$ and

$$g(x) = v_n(x - a_n)^2(x - b_n)^2$$

for $x \in (a_n, b_n)$, where the v_n are choosen in such a way that $g(f(p_n)) = 1$. Then g is differentiable at every point $\neq 0$ and g is constantly equal 0 on $E^c \in \mathcal{T}_{\mathcal{I}} \cap \mathcal{T}_{\mathcal{N}}$. Hence, $h \in \Delta^{(\mathcal{N})} \cap \Delta^{(\mathcal{I})}$. On the other hand, $g \circ f = 1$ on the set D^c while, $g \circ f(0) = 0$. Thus, $g \circ f \notin \Delta^{(\mathcal{N})} \cup \Delta^{(\mathcal{I})}$. ■

Example 4.12. The class $\mathcal{F}_c$ is not closed under composition.

Proof. Let h be an embedding of the irrationals $\mathbb{R} \setminus \mathbb{Q}$ into the Cantor set $C \subset [0,1]$. Put $f(q) = 2$ for $q \in \mathbb{Q}$ and $f(x) = h(x)$ for $x \in \mathbb{R} \setminus \mathbb{Q}$. Then, $f \in \mathcal{F}_c$. Moreover, let $S \subset \mathbb{R}\setminus\mathbb{Q}$ be a set without the Baire property and let g be the characteristic function of $f(S)$; i.e., $g = \chi_{f(S)}$. Then $g \in \mathcal{F}_c$, since $f(S) \subset C$ and $g \circ f = \chi_S \notin \mathcal{F}_c$. ■

5. Semigroups of $\mathcal{I}$-density Continuous Functions

For the reminder of this paper let $\mathcal{E}$ stand for any of the semigroups of Theorem 4.9; e.g., one of the classes $\mathcal{C}_{\mathcal{II}}$, $\mathcal{C}_{\mathcal{DD}}$, $\Delta \cap \mathcal{C}_{\mathcal{II}}$, $\Delta \cap \mathcal{C}_{\mathcal{DD}}$, $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}}$, $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}}$ or $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}}$. In particular, $\Delta \cap \mathcal{C}_{\mathcal{II}} \subset \mathcal{E} \subset \mathcal{C}_{\mathcal{DD}}$. It will be shown that these semigroups, with the possible exception of $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{II}}$, have the inner automorphism property.

We start with the following lemma.

Lemma 5.1. *If h is a generating bijection of an automorphism Ψ of $\mathcal{E}$, then h is $\mathcal{I}$-approximately continuous.*

Proof. Let $r \in \mathbb{R}$ and $\varepsilon > 0$ be arbitrary. We will show that h is $\mathcal{I}$-approximately continuous at r.

Let $f: \mathbb{R} \to \mathbb{R}$ be as in Example 4.3 for $a = h(r) - \varepsilon$, $b = c = h(r)$ and $d = h(r) + \varepsilon$. Let $g: \mathbb{R} \to \mathbb{R}$ with $g(x) = f(x) + h(r)$. It is easy to see that $g \in \Delta \cap \mathcal{C}_{\mathcal{II}} \subset \mathcal{E}$.

Ψ is an automorphism of $\mathcal{E}$, thus there is an $\alpha \in \mathcal{E} \subset \mathcal{C}_{\mathcal{DD}}$ such that $\Psi(\alpha) = g$. We have $h \circ \alpha = g \circ h$. Note that $h(\alpha(r)) = g(h(r)) = f(h(r)) + h(r) = 1 + h(r)$. Therefore, $h(\alpha(r)) \neq h(r)$. But, h is a bijection, so $\alpha(r) \neq r$.

The function α is deep-$\mathcal{I}$-density continuous. Thus, there exists a set $U \in \mathcal{T}_{\mathcal{D}}$ such that $\alpha(x) \neq r$ for all $x \in U$. Then, for all $x \in U$, $f(h(x)) + h(r) = g(h(x)) = h(\alpha(x)) \neq h(r)$, which implies $f(h(x)) \neq 0$; i.e., $h(x) \in (a, d) = (h(r) - \varepsilon, h(r) + \varepsilon)$. So $|h(x) - h(r)| < \varepsilon$ for $x \in U$. This means precisely that h is $\mathcal{I}$-approximately continuous. ∎

Corollary 5.2. *If h is a generating bijection of an automorphism Ψ of $\mathcal{E}$, then h is a homeomorphism.*

Proof. h is an $\mathcal{I}$-approximately continuous bijection of R. Since h is also a Darboux Baire one function, it must be a homeomorphism. ∎

As a next step we need the following

Lemma 5.3. *Let $h: \mathbb{R} \to \mathbb{R}$ be an increasing homeomorphism which is not $\mathcal{I}$-density continuous at 0. Then there exists a function $f \in \Delta \cap \mathcal{C}_{\mathcal{II}} \subset \mathcal{E}$ such that $h \circ f \circ h^{-1}$ is not deep-$\mathcal{I}$-density continuous.*

Proof. Without any loss of generality we may assume that $h(0) = 0$.

It is very easy to verify that any deep-$\mathcal{I}$-density continuous homeomorphism is also $\mathcal{I}$-density continuous. (Compare e.g. [23, Theorem 23].) Thus, h is also not deep-$\mathcal{I}$-density continuous at 0. Again, without any loss of generality we may further assume that h is not right deep-$\mathcal{I}$-density continuous at 0. The left-hand side argument is essentially the same.

The sets $E \cup (-E) \cup \{0\}$, where E is an open right interval set, form basis for $\mathcal{T}_{\mathcal{D}}$ at 0 [11]. Thus, there exists a right interval set $E = \bigcup_{n\in\mathbb{N}}[a_n, b_n]$ such that 0 is an $\mathcal{I}$-density point of E^c, while 0 is not a right $\mathcal{I}$-density point of the complement of

$$D = \bigcup_{n\in\mathbb{N}} [h^{-1}(a_n), h^{-1}(b_n)].$$

Using the definition of deep-$\mathcal{I}$-density point we can easily choose the sequences $\{\alpha_n\}$ and $\{\beta_n\}$ such that $h^{-1}(a_n) < \alpha_n < \beta_n < h^{-1}(b_n)$ for $n \in \mathbb{N}$ and 0 is still not a right $\mathcal{I}$-density point of the complement of

$$\bigcup_{n\in\mathbb{N}} [\alpha_n, \beta_n]$$

while, evidently, 0 is an $\mathcal{I}$-density point of the complement of

$$\bigcup_{n\in\mathbb{N}} [h(\alpha_n), h(\beta_n)] \subset \bigcup_{n\in\mathbb{N}} [a_n, b_n].$$

For each $n \in \mathbb{N}$, let $f_n: \mathbb{R} \to [0,1]$ be a function from Example 4.3 for $h^{-1}(a_n) < \alpha_n < \beta_n < h^{-1}(b_n)$. Define $f(x) = 0$ for $x \in E^c$ and $f(x) = c_n f_n(x)$ for $x \in [a_n, b_n]$ and $n \in \mathbb{N}$, where $c_n > 0$ are choosen in such a way that

$$\lim_{n\to\infty} \frac{h(c_{n+1})}{h(c_n)} = 0. \tag{9}$$

Then it is easy to see that f is differentiable and $\mathcal{I}$-density continuous.
On the other hand

$$T = (h \circ f \circ h^{-1}) \left(\bigcup_{n \in \mathbb{N}} (\alpha_n, \beta_n) \right) = \bigcup_{n \in \mathbb{N}} h(c_n).$$

Now, by (9) and Proposition 2.2, it is easy to see that there is a right interval set $S \supset T$ such that 0 is a deep-$\mathcal{I}$-density point of S^c, while 0 is not an $\mathcal{I}$-density point of $(\bigcup_{n \in \mathbb{N}} [\alpha_n, \beta_n])^c$. Hence, $h \circ f \circ h^{-1}$ is not deep-$\mathcal{I}$-density continuous. ■

Corollary 5.4. *If h is a generating bijection of an automorphism of $\mathcal{E}$, then h and h^{-1} are $\mathcal{I}$-density continuous homeomorphisms of R.*

Proof. By Corollary 5.2, h is a homeomorphism of R. So, by Lemma 5.3 it must be $\mathcal{I}$-density continuous. The statement for h^{-1} follows from the fact that h^{-1} is a generating bijection for Ψ^{-1}. ■

As an immediate corollary we obtain the following.

Theorem 5.5. *The semigroups $\mathcal{C}_{\mathcal{I}\mathcal{I}}$ and $\mathcal{C}_{\mathcal{D}\mathcal{D}}$ have the inner automorphism property. Moreover,*

$$\mathrm{Aut}(\mathcal{C}_{\mathcal{I}\mathcal{I}}) = \mathrm{Aut}(\mathcal{C}_{\mathcal{D}\mathcal{D}}).$$

Let us notice that we can also deduce that $\mathcal{C}_{\mathcal{D}\mathcal{D}}$ has the inner automorphism property by using Proposition 1.2, as the deep-$\mathcal{I}$-density topology $\mathcal{T}_{\mathcal{D}}$ is generated. This follows immediatelly from the fact that the function which demonstrates the complete regularity of the $\mathcal{T}_{\mathcal{D}}$ [11, Theorem 5] is also deep-$\mathcal{I}$-density continuous[3]. However, Proposition 1.2 cannot be used in the case of the $\mathcal{I}$-density topology, because $\mathcal{T}_{\mathcal{I}}$ is not generated. This follows from the fact that for $\mathcal{T}_{\mathcal{I}}$ to be generated, the family $\mathcal{F} = \{(f^{-1}(\{x\}))^c : x \in \mathbb{R} \text{ and } f : (\mathbb{R}, \mathcal{T}_{\mathcal{I}}) \to (\mathbb{R}, \mathcal{T}_{\mathcal{I}})\}$ would have to form a subbase of $\mathcal{T}_{\mathcal{I}}$. This is not the case since even the family $\{(f^{-1}(\{1\}))^c : f : (\mathbb{R}, \mathcal{T}_{\mathcal{I}}) \to [0,1]\} \subset \mathcal{F}$ does not form a subbase for $\mathcal{T}_{\mathcal{I}}$, because $\mathcal{T}_{\mathcal{I}}$ is not regular [19, Theorem 5].

To discuss the inner automorphism properties of the other semigroups we need the following two lemmas.

Lemma 5.6. *The notions of $\mathcal{I}$-approximate derivative, approximate derivative and ordinary derivative coincide on the class of all homeomorphisms. In other words, if $f : \mathbb{R} \to \mathbb{R}$ is a homeomorphism and $x \in \mathbb{R}$ then $f'(x) = \Delta^{(\mathcal{I})} f(x) = \Delta^{(\mathcal{N})} f(x)$ whenever any piece of this equation exists.*

Proof. See Łazarow and Wilczyński [12], also Wilczyński [23, Theorem 38]. ■

Lemma 5.7. *Let Ψ be an automorphism of $\Delta \cap \mathcal{C}_{\mathcal{I}\mathcal{I}}$, $\Delta \cap \mathcal{C}_{\mathcal{D}\mathcal{D}}$, $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{I}\mathcal{I}}$ or $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{D}\mathcal{D}}$ and let h be its generating bijection. Then h is differentiable; i.e., $h \in \Delta$.*

---

[3] If $0 \in A \in \mathcal{T}_{\mathcal{D}}$, then there exist right interval sets $E \subset V \subset A$ composed of closed and open intervals, respectively, such that 0 is a deep-$\mathcal{I}$-density point of a closed set $P = \mathrm{cl}((-E) \cup E)$ which is a subset of $U = (-V) \cup V \cup \{0\}$. Lazarow's function is defined by $f(0) = 1$ and $f(x) = (\mathrm{dist}(x, U^c)) / (\mathrm{dist}(x, U^c) + \mathrm{dist}(x, P))$ for $x \neq 0$. Because f is piecewise linear on a neighborhood of every point $x \neq 0$, it must be density continuous at x. It is deep-$\mathcal{I}$-density continuous at 0 because f is constant on $U \in \mathcal{T}_{\mathcal{D}}$.

Proof. It is already known that h is a homeomorphism. Thus h is differentiable almost everywhere [2]. Let x_0 be a point of differentiability of h and let $x \in \mathbb{R}$ be any other point. Define $f(t) = t + (x - x_0)$. Then $f \in \Delta \cap \mathcal{C}_{II}$, so it belongs to all of the semigroups under consideration. For any positive δ we have

$$\begin{aligned}\frac{h(x+\delta) - h(x)}{\delta} &= \frac{h \circ f(x_0 + \delta) - h \circ f(x_0)}{\delta} \\ &= \frac{\Psi(f) \circ h(x_0 + \delta) - \Psi(f) \circ h(x_0)}{\delta}.\end{aligned}$$

If Ψ is an automorphism of $\Delta \cap \mathcal{C}_{II}$ or $\Delta \cap \mathcal{C}_{\mathcal{DD}}$ then $\Psi(f)$ is differentiable, and since h is differentiable at x_0, the quotient in (10) converges to $\Psi(f)'(h(x_0))\, h'(x_0)$.

If Ψ is an automorphism of $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{II}$ or $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}}$ then $\Psi(f)$ is $\mathcal{I}$-approximately differentiable. So, by Lemma 4.7, $\Psi(f) \circ h$ is $\mathcal{I}$-approximately differentiable at x_0. Hence, condition (10) guarantees that $D^{(\mathcal{I})}h(x)$ exists and is equal $D^{(\mathcal{I})}(\Psi(f) \circ h)(x_0)$. But, h is a homeomorphism. Thus, by Lemma 5.6, h must be differentiable at x as well. This ends the proof. ■

As an immediate corollary we obtain

Theorem 5.8. *The semigroups $\Delta \cap \mathcal{C}_{II}$, $\Delta \cap \mathcal{C}_{\mathcal{DD}}$, $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{II}$ and $\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}}$ have the inner automorphism property. Moreover,*

$$\begin{gathered}\mathrm{Aut}(\Delta \cap \mathcal{C}_{II}) = \mathrm{Aut}(\Delta \cap \mathcal{C}_{\mathcal{DD}}) = \mathrm{Aut}(\Delta^{(\mathcal{I})} \cap \mathcal{C}_{II}) = \mathrm{Aut}(\Delta^{(\mathcal{I})} \cap \mathcal{C}_{\mathcal{DD}}) \\ \subset \mathrm{Aut}(\mathcal{C}_{II}) = \mathrm{Aut}(\mathcal{C}_{\mathcal{DD}})\end{gathered}$$

and the inclusion is proper.

We are not able to prove or disprove that the semigroup $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{II}$ has the inner automorphism property. We proved that each automorphism of $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{II}$ is generated by a homeomorphism h of the real line which must be differentiable almost everywhere. However, to be an element of $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{II}$, h would need to be differentiable outside of a set of first category. In general, a homeomorphism of the real line need not be $\mathcal{I}$-a.e. differentiable. In fact, Belna, Cargo, Evans and Humke [1] show that there exists a strictly increasing homeomorphism $h : [0,1] \to [0,1]$ and $\tau \in (0,1)$ such that

$$\underline{D}^- h(x) = \underline{D}^+ h(x) = \tau \text{ and } \overline{D}^- h(x) = \overline{D}^+ h(x) = \infty$$

for a residual set of points $x \in [0,1]$. On the other hand, there are certain restrictions on the "severity" of the nondifferentiability of h on sets of the second category. Neugebauer [15] shows that a continuous function f has $\underline{D}^- f(x) = \overline{D}^+ f(x)$ and $\overline{D}^- f(x) = \overline{D}^+ f(x)$ on a residual set.

We are not able to find a satisfactory answer to our question. Therefore, the following remains.

Problem 5.9. Let h be a homeomorphism of R such that for every $\mathcal{I}$-density continuous, $\mathcal{I}$-a.e. $\mathcal{I}$-approximately differentiable f, $h \circ f \circ h^{-1}$ is also $\mathcal{I}$-density continuous, and $\mathcal{I}$-a.e. $\mathcal{I}$-approximately differentiable. Is h differentiable on a residual set?

Clearly, an affirmative answer to this problem is equivalent to the fact that $\Delta_{\mathcal{I}}^{(\mathcal{I})} \cap \mathcal{C}_{II}$ has the inner automorphism property.

References

[1] C.L. Belna, G.T. Cargo, M.J. Evans, and P.D. Humke. Analogues of the Denjoy-Young-Saks theorem. *Trans. Amer. Math. Soc.* **271** (1982), 253–260.

[2] A. M. Bruckner. *Differentiation of Real Functions. Lecture Notes in Mathematics* **659**, Springer-Verlag, 1978.

[3] Krzysztof Ciesielski and Lee Larson. Analytic functions are $\mathcal{I}$-density continuous. (submitted).

[4] Krzysztof Ciesielski and Lee Larson. Baire classification of $\mathcal{I}$-approximately and $\mathcal{I}$-density continuous functions. (submitted).

[5] Krzysztof Ciesielski and Lee Larson. The density topology is not generated. (submitted).

[6] Krzysztof Ciesielski and Lee Larson. Refinements of the density and $\mathcal{I}$-density topologies. (submitted).

[7] Krzysztof Ciesielski and Lee Larson. The space of density continuous functions. *Acta Math. Hung.*, to appear.

[8] Krzysztof Ciesielski and Lee Larson. Various continuities with the density, $\mathcal{I}$-density and ordinary topologies on R. (submitted).

[9] Krzysztof Ciesielski, Lee Larson, and Krzysztof Ostaszewski. Differentiability and density continuity. *Real Anal. Exchange* **15** (1989-90), 239–247.

[10] Casper Goffman and Daniel Waterman. Approximately continuous transformations. *Proc. Amer. Math. Soc.* **12** (1961), 116–121.

[11] E. Łazarow. The coarsest topology for $\mathcal{I}$-approximately continuous functions. *Comment. Math. Univ. Caroli.* **27** (1986), 695–704.

[12] E. Łazarow and W. Wilczyński. $\mathcal{I}$-approximate derivatives. to appear.

[13] K. D. Magill, Jr. Automorphisms of the semigroup of all differentiable functions. *Glasgow Math. J.* **8** (1967), 63–66.

[14] K. D. Magill, Jr. A survey of semigroups of continuous selfmaps. *Semigroup Forum* **11** (1975/76), 189–282.

[15] C. J. Neugebauer. A theorem on derivatives. *Acta Sci Math (Szeged)* **23** (1962), 79–81.

[16] Krzysztof Ostaszewski. Continuity in the density topology II. *Rend. Circ. Mat. Palermo*, **(2)32** (1983), 398–414.

[17] Krzysztof Ostaszewski. Semigroups of density continuous functions. *Real Anal. Exchange* **14** (1988–89), 104–114.

[18] J. C. Oxtoby. *Measure and Category.* Springer-Verlag, 1971.

[19] W. Poreda, E. Wagner-Bojakowska, and W. Wilczyński. A category analogue of the density topology. *Fund. Math.* **75** (1985), 167–173.

[20] W. Poreda, E. Wagner-Bojakowska, and W. Wilczyński. Remarks on $\mathcal{I}$-density and $\mathcal{I}$-approximately continuous functions. *Comm. Math. Univ. Carolinae* **26** (1985), 553–563.

[21] J. Schreier. Über Abbildungen einer abstracten Menge auf ihre Teilmengen. *Fund. Math.* **28** (1937), 261–264.

[22] J. C. Warndof. Topologies uniquely determined by their continuous selfmaps. *Fund. Math.* **66** (1969/70), 25–43.

[23] W. Wilczyński. A category analogue of the density topology, approximate continuity, and the approximate derivative. *Real Anal. Exchange* **10** (1984–85), 241–265.

Department of Mathematics
West Virginia University
Morgantown, WV 26506

Department of Mathematics
University of Louisville
Louisville, KY 40292

Received October 11, 1990
and in final form May 21, 1991

Semigroup Forum Vol. 45 (1992) 205–213

RESEARCH ARTICLE

The Structure of Morita Dual Monoids

Peeter Normak*

Communicated by F. Pastijn

Morita duality for monoids is introduced in [5]. This paper gives a general characterization of Morita dual monoids and presents a finite example.

In this paper we describe all monoids having Morita duality: they are exactly atomic Rees monoids (or, equivalently, special left chain semigroups). It follows that a finite monoid has Morita duality if and ony if it is a monogenic nilpotent semigroup with identity.

1. Basic definitions and results

Let S and T be monoids. A left *S-act* ${}_SA$ is a set A on which S acts from the left, that is $1 \cdot a = a$ and $(s \cdot t) \cdot a = s \cdot (t \cdot a)$ for all $a \in A$, $s, t \in S$. Homomorphisms between acts will be written on the side opposite to the elements of the monoids. By S-Act (Act-T) we denote the category of all left S-acts (right T-acts). If C is a left S-act and a right T-act and $s(ct) = (sc)t$ for every $s \in S$, $c \in C$, $t \in T$, then C is called S-T-biact and denoted by ${}_SC_T$. By ${}_S\mathfrak{S}$ we denote the minimal full subcategory of S-Act which is closed under subacts and factor acts, and contains S. The monoids S and T are *Morita dual* monoids, if the categories ${}_S\mathfrak{S}$ and $\mathfrak{S}_T$ are dually equivalent. In this case, the monoid S (and T) is said to *have a Morita duality.* If, in addition, S is isomorphic to T, then S is called *selfdual.* We call a monoid S with zero *atomic* if there exists an element $0 \neq u \in S$ such that $ux = xu = 0$ for all elements $1 \neq x \in S$. The element u is called an *atom* in this case. Call monoid S a *Rees monoid* if every congruence on S is a Rees congruence. A monoid S is a *(left) chain monoid* if the lattice of its (left) congruences is a chain. A linearly ordered monoid S is *divisible* if for $x \leq y$, $x, y \in S$, there exists a $z \in S$ such that $yz = x = zy$. As usual, an S-act A is called *subdirectly irreducible* if every set of congruences $\{k_i \mid i \in I\}$ on A with $\bigcap_{i \in I} k_i = \Delta_A$ contains Δ_A, where Δ_A denotes the diagonal of $A \times A$. If the S-acts ${}_SX$ and ${}_SY$ both have exactly one zero, then the coproduct of ${}_SX$ and ${}_SY$ with amalgamated zero, denoted by ${}_SX \overset{\circ}{\amalg} {}_SY$, is called the *0-coproduct.* A left S-act with one generating element is called *cyclic.* If every finitely generated subsystem of ${}_SA$ is cyclic then ${}_SA$ is called *locally cyclic.*

For a biact ${}_SC_T$, an S-act ${}_SA$ is called *${}_SC_T$-reflexive* if the canonical mapping $h_A : {}_SA \to \mathrm{Hom}_T(\mathrm{Hom}_S({}_SA, {}_SC_T), {}_SC_T)$ defined by $((a)h_A)u = (a)u$ for $u \in \mathrm{Hom}_S({}_SA, {}_SC_T)$, $a \in A$, is an isomorphism.

We recall the following:

Lemma 1.1 ([5], Proposition 3.4). *An S-act C is a cogenerator in any full subcategory $\mathfrak{B}$ of S-Act, which is closed with respect to factor acts, if and only if every subdirectly irreducible S-act in $\mathfrak{B}$ is isomorphic to a subact of C.*

* The author wishes to express his gratitude to Ulrich Knauer for many helpful comments.

Lemma 1.2 ([5], Lemma 2.4). *Let S, T be monoids with zero, let the S-T-biact ${}_SC_T$ be an injective cogenerator in ${}_S\mathfrak{S}$ and in $\mathfrak{S}_T$ with $\mathrm{End}_T(C_T) \cong S$ and $\mathrm{End}_S({}_SC) \cong T$ as monoids. Let ${}_SB$ be a ${}_SC_T$-reflexive act in ${}_S\mathfrak{S}$. Then every subact and every factor act of ${}_SB$ is ${}_SC_T$-reflexive.*

Lemma 1.3 ([5], Theorem 3.6). *Monoids S and T are Morita dual if and only if S and T contain zero and there exists an S-T-biact ${}_SC_T$ which is cyclic injective in S-Act and in Act-T and a cogenerator in ${}_S\mathfrak{S}$ and $\mathfrak{S}_T$ and, in addition, $S \cong \mathrm{End}_T({}_SC_T)$ and $T \cong \mathrm{End}_S({}_SC_T)$ as monoids.*

Lemma 1.4 ([5], Lemma 4.3). *Let the nonzero S-acts ${}_SX$ and ${}_SY$ be generated by any one of their nonzero elements. Then the 0-coproduct ${}_SX \overset{\circ}{\amalg} {}_SY$ is not ${}_SC_T$-reflexive.*

A selfdual monoid S is called *reflexive* if the biact ${}_SC_S$ in Lemma 1.3 is isomorphic to ${}_SS_S$.

Lemma 1.5. *If a monoid S has a Morita duality, via an S-T-biact ${}_SC_T$, then all elements in ${}_S\mathfrak{S}$ and $\mathfrak{S}_T$ are ${}_SC_T$-reflexive.*

Proof. If S has a Morita duality via an S-T-biact ${}_SC_T$, then, by Lemma 1.3, S is ${}_SC_T$-reflexive. Now use Lemma 1.2. ∎

In the sequel we need the following lemmas. First, from Lemmas 1.2 and 1.4 we get

Lemma 1.6. *No 0-coproduct of acts is ${}_SC_T$-reflexive.*

Lemma 1.7. *If a monoid S has a Morita duality, via an S-T-biact ${}_SC_T$, then all ${}_SC_T$-reflexive S-acts are locally cyclic.*

Proof. Let ${}_SA$ be ${}_SC_T$-reflexive. Assume that ${}_SA$ is not locally cyclic. Then there exist elements a_1, $a_2 \in A$ such that $Sa_1 \not\subseteq Sa_2$ and $Sa_2 \not\subseteq Sa_1$. By Lemmas 1.2 and 1.3 the Rees factor

$$(Sa_1 \cup Sa_2)/(Sa_1 \cap Sa_2) = (Sa_1/(Sa_1 \cap Sa_2)) \overset{\circ}{\amalg} (Sa_2/(Sa_1 \cap Sa_2))$$

is ${}_SC_T$-reflexive, which contradicts Lemma 1.6. ∎

Lemma 1.8 ([6], Theorem). *Let S be a semigroup with more than two elements. Then S is a left chain semigroup if and only if one of the following conditions is satisfied:*

1) *S is commutative and the lattice of its congruences is a chain;*
2) *$S = \{e; f; u\}$ where $se = sf = s$ and $su = u$ for all $s \in S$;*
3) *$S = \{e; u; v\}$ where $se = s$, $su = u$ and $sv = v$ for all $s \in S$;*
4) *$S = \{e; a; u\}$ where $se = s$ and $sa = su = u$ for all $s \in S$;*
5) *$S = \{e; a; u; v\}$ where $se = s$, $su = u$, $sv = v$ for all $s \in S$ and $a^2 = e$, $ua = v$, $va = u$.*

Let M be the multiplicative semigroup of reals in $[0, 1)$ and let $0 \neq g \in M$ be an element. In the sequel throughout the paper, by Q and R we shall denote the Rees factors M^1/Mg and M^1/M^1g, respectively. Clearly

$$Q = (\theta \cup [g, 1], \cdot), \text{ where } q_1q_2 = \theta \text{ iff } q_1q_2 < g \text{ and}$$
$$R = (\theta \cup (g, 1], \cdot), \text{ where } r_1r_2 = \theta \text{ iff } r_1r_2 \leq g\ .$$

Obviously, Q is atomic (with atom g) and R is not.

Lemma 1.9 ([8], Theorem 1). *The lattice of all congruences of a commutative semigroup G is a chain if and only if G is isomorphic to one of the following semigroups:*

1) *a cyclic group of order p^n or the group of type p^∞, where p is an arbitrary prime and n any positive integer;*
2) *a group of the type* 1) *with adjoined zero;*
3) *a monogenic nilpotent semigroup;*
4) *a semigroup of the type* 3) *with adjoined identity;*
5) *an infinite divisible subsemigroup of Q or R.*

In the latter case G is isomorphically embeddable either in Q or in R but not in both and the isomorphism is uniquely defined.

As we see from the following lemma, the presence of an identity element in a semigroup simplifies considerably the structure of Rees semigroups:

Lemma 1.10 ([2], Satz 3.11 and 3.15). *A monoid is a Rees monoid if and only if it is isomorphic to a divisible submonoid of Q or R, defined above.*

From Lemmas 1.9 and 1.10 it follows that the lattice of all congruences of a Rees monoid is a chain, i.e. every Rees monoid is a chain monoid.

2. Preliminary lemmas

In this section several lemmas are given which will be useful in the proof of our main result.

Lemma 2.1. *A monoid with a Morita duality is a left chain monoid.*

Proof. Let S be a monoid with a Morita duality and let k_1 and k_2 be left congruences on S. Let T be a Morita dual to the monoid S and let $F : {}_S\mathfrak{S} \to \mathfrak{S}_T$ and $G : \mathfrak{S}_T \to {}_S\mathfrak{S}$ be contravariant equivalence functors. Since there exist natural epimorphisms $p_1 : S \to S/k_1$ and $p_2 : S \to S/k_2$ we get that $F(S/k_1)$ and $F(S/k_2)$ are subacts of $F(S)$. Assume that $F(S/k_1)$ and $F(S/k_2)$ are incomparable. Because $\mathfrak{S}_T$ is closed under subacts and factors we get that the Rees factor

$$(F(S/k_1) \cup F(S/k_2)) \,/\, (F(S/k_1) \cap F(S/k_2)) =$$
$$= F(S/k_1)/\,(F(S/k_1) \cap F(S/k_2)) \overset{\circ}{\amalg} (F(S/k_2))\,/\,(F(S/k_1) \cap F(S/k_2))$$

belongs to $\mathfrak{S}_T$. But this contradicts Lemmas 1.5 and 1.6. Without loss of generality we can thus assume that there exists a monomorphism $f : F(S/k_1) \to F(S/k_2)$ such that the following diagram is commutative:

$$\begin{tikzcd}
F(S) & & F(S/k_1) \arrow[ll, "F(p_1)"'] \arrow[dl, "f"] \\
& F(S/k_2) \arrow[ul, "F(p_2)"] &
\end{tikzcd}$$

Now, applying the functor G, we get an epimorphism $G(f): S/k_2 \to S/k_1$ such that the diagram

$$\begin{tikzcd} S \arrow[rr, "p_1"] \arrow[dr, "p_2"'] & & S/k_1 \\ & S/k_2 \arrow[ur, "G(f)"'] & \end{tikzcd}$$

is commutative. Hence $k_2 \subseteq k_1$ and therefore S is a left chain monoid. ∎

Since the semigroups described in Lemma 1.8 under 2), 3), 4) and 5) do not contain both identity and zero element, we have the following

Lemma 2.2. *A monoid with a Morita duality is a commutative chain monoid.*

By [4] commutative Morita equivalent monoids are isomorphic. Using [5, Corollary 3.7], we get the following

Corollary 2.3. *If a monoid S has a Morita dual monoid T, then T is unique up to isomorphism.*

Now we proceed with some lemmas.

Lemma 2.4. *A finite Rees monoid is monogenic.*

Proof. Let $S = \{0; x_1; x_2; \dots; x_n; 1\}$ be a finite Rees monoid. By Lemma 1.10 we can assume that $0 < x_1 < x_2 < \dots < x_n < 1$. Suppose that there exists an element y in S such that $x_n^{k+1} < y < x_n^k$. Then by divisibility there exists an element $z \in S$ such that $x_n^{k+1} = yz$. We have $x_n^{k+1} = yz \le yx_n < x_n^k \cdot x_n = x_n^{k+1}$, a contradiction. Hence $S = \{x_n^k \mid k = 1, \dots, n+1\} \cup \{1\}$ and therefore S is a monogenic semigroup with adjoined identity. ∎

Lemma 2.5. *Every infinite Rees monoid S is dense (in the usual sense) in Q or R.*

Proof. We have to show that for every two elements $x < y$ in Q or R there exists an element $t \in S$ such that $x < t < y$. Let $u, v \in S$ be elements such that $0 < v - u < g(y-x)/2$ (these elements exist by infiniteness of S and by Lemma 1.10). Then $v > u > v - v(y-x)/2$ and hence $1 > u/v > 1 - (y-x)/2 > x$. Denote u/v by s. For every natural number l we have $s^l - s^{l+1} = s^l(1-s) < s^l(y-x)/2 < (y-x)/2$. Because $\lim s^k = 0$ we have that there exists an N such that $x < s^N < y$. ∎

Lemma 2.6. *Every infinite Rees monoid S is injective as an S-act.*

Proof. Let S be a Rees monoid and let $S \subseteq {}_S A$ be an inclusion of S-acts. We will construct a retraction $f: {}_S A \to S$. Let $a \in {}_S A$ be an arbitrary element. If $Sa \cap S = 0$, then define $f(a) = 0$. Otherwise, if $a \in S$ then define $f(a) = a$. Now let $a \in A \backslash S$ and let $s \in S$ be an element such that $0 \neq sa \in S$. Then $sa \le s$. For, if $s < sa$, then by Lemma 2.5 there exists an element $u \in S$ such that $us < g < u(sa)$. But then $(us)a = \theta a = \theta \neq u(sa)$ which contradicts the definition of an S-act. Hence $sa \le s$ and (by divisibility of S) there exists an element $t \in S$ such that $sa = st$. Define $f(a) = t$. Next we show that f is a homomorphism. Let $w \in S$ be an arbitrary element. If $w < s$ then $w = vs$

for some $v \in S$ and we get $wf(a) = wt = vst = vsa = f(vsa) = f(wa)$. If $w > s$ then $s = vw$ for some $v \in S$ and $v(wa) = sa \in S$. Hence there exists an element $p \in S$ such that $v(wa) = vp$. Then $f(wa) = p$ and we have $vwf(a) = vwt = st = sa = vp = vf(wa)$. By cancellability we get $wf(a) = f(wa)$. Therefore f is a well-defined S-homomorphism and a retraction. By [1, Corollary 2 of Theorem 6] this implies that S is injective as a left S-act. ∎

Lemma 2.7. *Let I be an ideal of an infinite Rees monoid S and let $s = \sup I$. Then $I = (S \cap [g,s]) \cup \theta$ or $I = (S \cap [g,s)) \cup \theta$.*

Proof. First let $s = \sup I$ and $t \in S$ with $t < s$. If $s \in I$ then by divisibility there exists an element $x \in S$ such that $xs = t$. Hence $t \in I$. If $s \notin I$ then there exists $u \in I$ such that $t < u < s$. By divisibility of S there exists an $x \in S$ such that $t = xu$. Hence $t \in I$. ∎

From Lemmas 2.5 and 2.7 it follows that if an infinite Rees monoid S is not atomic, then the intersection of all nontrivial (Rees) congruences on S is the diagonal congruence. On the other hand, if a Rees monoid S is atomic (with atom g), then the intersection of all nontrivial congruences on S is $(0,g) \cup (g,0) \cup \Delta_S$. Clearly every finite Rees monoid is atomic. Hence we get the following

Corollary 2.8. *A Rees monoid is subdirectly irreducible if and only if it is atomic.*

Corollary 2.9. *For an infinite Rees monoid S, every cyclic S-act can be considered as a Rees monoid, too.*

Proof. Let S be an infinite Rees monoid and ${}_SZ$ a cyclic (left) S-act. Then ${}_SZ \cong S/I$ with I an ideal in S. By Lemma 2.7 we have $I = (S \cap [g,s]) \cup \theta$ or $I = (S \cap [g,s)) \cup \theta$, where $s = \sup I$. Then $S/I \cong (S \cap (s,1]) \cup \theta$ or $S/I \cong (S \cap [s,1]) \cup \theta$ is a divisible submonoid of R or Q (with $g = s$), respectively. Hence ${}_SZ = S/I$ is a Rees monoid, by Lemma 1.10. ∎

From Lemma 2.7 and Corollary 2.9 we get the following

Corollary 2.10. *Let S be an infinite Rees monoid and let ${}_SZ$ be a cyclic S-act. Then a subact ${}_SX \subseteq {}_SZ$ is atomic if and only if ${}_SZ$ is atomic.*

Lemma 2.11. *Every infinite atomic Rees monoid S is a cogenerator in ${}_S\mathfrak{S}$.*

Proof. By Lemma 1.1 and Corollary 2.8 it suffices to show that S contains an isomorphic copy of every atomic S-act in ${}_S\mathfrak{S}$ or, equivalently (by Corollary 2.10), of every cyclic atomic S-act ${}_SZ$. By Lemmas 1.10 and 2.7 we have that $Z = S \backslash [g,s)$ for some $s \in S$, $s > g$. Define the mapping $f : {}_SZ \to S$, setting for every element $u \in Z$

$$f(u) = \begin{cases} \theta & , \quad \text{if } u = \theta \\ (g/s)u & , \quad \text{if } u \neq \theta \, . \end{cases}$$

It is clear that f is a monomorphism.

For our further purposes we also need the next

Lemma 2.12 ([5], Example 5.6). *Every finite monogenic semigroup with adjoined identity is reflexive.*

3. The main result

The next theorem gives us a complete characterization of monoids having Morita duality.

Theorem 3.1. *The following properties of a monoid S are equivalent:*

1) *S has a Morita dual monoid;*
2) *S is reflexive;*
3) *S is left and right selfinjective and a cogenerator in ${}_S\mathfrak{S}$ and in $\mathfrak{S}_S$.*
4) *S is an atomic Rees monoid;*
5) *S is a commutative atomic chain monoid with trivial group of units;*
6) *S is isomorphic to an atomic divisible submonoid of a monoid $Q = (\theta \cup [g, 1], \cdot)$ for some $g \in (0, 1)$, where $x \cdot y = \theta$ whenever $xy < g$.*

Proof. The implication 2) $\Longrightarrow$ 1) is trivial. The equivalence of 2) and 3) is proved in [5, Corollary 5.3]. By Lemma 1.10, 4) and 6) are equivalent (remark that the monoid $R = (\theta \cup (g, 1], \cdot)$ is not atomic). For implications 5) $\Longrightarrow$ 6) and 6) $\Longrightarrow$ 5) remark that monogenic nilpotent semigroups with adjoined identity can be considered as atomic divisible submonoids of R or Q: in the finite case, the monoid

$$\{1; x; x^2; \ldots; x^{n-1}; x^n\}$$

is obviously isomorphic, for example, to the monoid

$$\{1; 1/2; (1/2)^2; \ldots; (1/2)^{n-1}; \theta\} .$$

Now the implications mentioned above follow from Lemma 1.9.

1) $\Longrightarrow$ 5). Let S have a Morita dual monoid via an S-T-biact ${}_SC_T$. By Lemma 2.2 we have that S is a commutative chain monoid. By Lemma 1.3 the monoid S is not a nontrivial group. The monoid S cannot be a nontrivial group with adjoined zero. For, if $S = G \cup \theta$, G a group, then consider the congruence $\rho \subset S \times S$ defined by

$$s_1 \,\rho\, s_2 \Leftrightarrow \begin{cases} s_1 = s_2 & \text{or} \\ s_1, s_2 \in G. \end{cases}$$

Since $|S/\rho| = 2$, S/ρ is subdirectly irreducible. By Lemmas 1.1 and 1.3 we have that $S/\rho \subseteq {}_SC$. Since every cyclic S-act is generated by all its nonzero elements, we get that $S/\rho = {}_SC$. By [5, Proposition 5.4] we have that $|S| = |{}_SC_T|$. Hence $S = \{0; 1\}$. By Lemma 1.9 it follows that S is isomorphic to a divisible submonoid of R or Q. The same holds for T. As T is a Rees monoid (Lemma 1.10) and C_T is cyclic we get that $C_T \cong T/I$ for an ideal I. Consider a congruence k on T defined by

$$ukv \Leftrightarrow \begin{cases} u = 1 = v & \text{or} \\ u \neq 1 \neq v. \end{cases}$$

The factor act T/k has 2 elements only and is therefore subdirectly irreducible. By Lemmas 1.1 and 1.3 we have $T/k \subseteq C$. Hence C_T has an atom, say $a \in C_T$. Consider the element $f \in \mathrm{End}(C_T)$ such that

$$f(c) = \begin{cases} a, & \text{if } c = 1 \\ 0, & \text{if } c \neq 1. \end{cases}$$

It is clear that f is a well-defined T-homomorphism and that f is an atom in $\mathrm{End}(C_T)$. By Lemma 1.3 we have that $\mathrm{End}_T(C_T) = S$. Hence the monoid S is atomic.

5) $\Longrightarrow$ 2). Let S be an atomic divisible submonoid of Q. By [5, Corollary 5.3] we have to show that ${}_SS$ is injective in S-Act and a cogenerator in ${}_S\mathfrak{S}$. By Lemma 1.10, S is a Rees monoid. If S is infinite, we use Lemmas 2.6 and 2.11. If S is finite, then we use Lemmas 2.4 and 2.12. ∎

As we see from the following corollary, the structure of finite Morita dual monoids is extremely simple.

Corollary 3.2. *A finite monoid has a Morita dual if and only if it is a monogenic semigroup with adjoined identity.*

Proof. Necessity follows from Theorem 3.1 and Lemma 2.4. Sufficiency follows from Theorem 3.1 and Lemma 2.12. ∎

Using Corollary 2.3 we get the following:

Corollary 3.3. *If the monoids S and T are Morita dual then they are isomorphic.*

4. Some related problems

In this section we discuss relations between Morita duality and some other algebraic properties.

1) An S-act A is called a *cover* of B if there exists an epimorphism $A \to B$ whose restriction to every subobject of A is not onto. A monoid S is said to be *left semiperfect* if every finitely generated left S-act has a projective cover. Since every projective S-act is a coproduct of cyclic S-acts [4, Corollary 3.8], it is easy to see that a monoid S is left semiperfect if and only if every cyclic left S-act has a projective cover. If S has a Morita dual then a cyclic S-act Sz is isomorphic to a Rees factor S/I for some ideal I (cf. Theorem 3.1) and S is a projective cover of S/I (via the natural epimorphism $S \to S/I$). Hence we have the following:

Proposition 4.1. *Every monoid with Morita duality is semiperfect.*

As we see from the 3-element chain (as a lower semilattice), the class of semiperfect monoids is essentially wider than the class of all monoids having Morita duality. A monoid S is said to be *perfect* if every left S-act has a projective cover. No infinite monoid with Morita duality is perfect. For, if S is an infinite Morita dual monoid, then $S\backslash\{1\}$ as a left S-act is locally cyclic but not cyclic. But for perfect monoids, every locally cyclic act is cyclic (cf. 1.1 of [3]).

2) If S has Morita duality then by Theorem 3.1 and by the definition of the category ${}_S\mathfrak{S}$, every object in ${}_S\mathfrak{S}$ is locally cyclic. On the other hand, let ${}_SA$ be a locally cyclic S-act and let $\theta \neq a \in A$ be an element. The mapping $\varphi : Sa \to S$, defined by

$$\varphi(sa) = \begin{cases} g, & \text{if } s = 1 \\ \theta, & \text{if } s \neq 1, \end{cases}$$

is a well-defined S-homomorphism. By injectivity of S (cf. Lemma 1.3), there exists a homomorphism $\overline{\varphi} : A \to S$ such that the diagram

$$\begin{tikzcd} & S \\ A \arrow[ur, "\overline{\varphi}"] & Sa \arrow[l, tail, "i"] \arrow[u] \end{tikzcd}$$

is commutative (i an inclusion map). If a is an atom in A then it can be easily shown that $\overline{\varphi}$ is a monomorphism, that is $A \cong \overline{\varphi}(A) \subseteq S$. If $\overline{\varphi}(A)$ has a maximal element m, then A is cyclic, generated by $\overline{\varphi}^{-1}(m)$. If A has neither an atom nor a maximal element then $A \cong (S\backslash\{1\})/I$, where $I = \{s \in S \mid \forall\ a \in A, sa = \theta\}$. Hence, using Lemmas 1.5 and 1.7 we get the following:

Proposition 4.2. *If S has Morita duality then the category ${}_S\mathfrak{S}$ consists exactly of all locally cyclic S-acts, or, equivalently, of all ${}_SS_S$-reflexive S-acts.*

3) Let X be a subset of a monoid S with zero. Denote by $l(X)$ and $r(X)$ the left and right annihilator of the set X, respectively. A semigroup with zero S is called *dual* if for every ideal L of S we have $l(r(L)) = L$ and for every right ideal R of S we have $r(l(R)) = R$ (cf. [9]). If I is an ideal of a monoid S having Morita duality, then by Lemma 2.7 we have $I = (S \cap [g, s]) \cup \theta$ or $I = (S \cap [g, s)] \cup \theta$ for some $s \in (0,1)$. Then $l(I) = r(I) = (S \cap [g, g/s)) \cup \theta$ or $l(I) = r(I) = (S \cap [g, g/s]) \cup \theta$, respectively. Hence $l(r(I)) = r(l(I)) = I$ for every ideal I and we get the following:

Proposition 4.3. *Every monoid having Morita duality is dual.*

The converse to Proposition 4.3 is not true: take, for example, a nontrivial multiplicative group with external zero.

4) The problem, to what extent the terms "Morita dual ring" and "Morita dual monoid" are related, remains open. A noncommutative semisimple ring provides an example of a ring having Morita duality (cf. [7]) whose set of left ideals (with respect to the multiplication of ideals or as a lower semilattice of ideals as well) does not have Morita duality (as a monoid), since its ideals are not linearly ordered. The multiplicative semigroup of this ring also does not have Morita duality (cf. Theorem 3.1).

References

[1] Berthiaume, P., *The injective envelope of S-sets*, Canad. Math. Bull. **10** (1967), 261–273.

[2] Hotzel, E., *Halbgruppen mit ausschließlich reesschen Linkskongruenzen*, Math. Z. **112** (1969), 300–320.

[3] Isbell, J. R., *Perfect monoids*, Semigroup Forum **2** (1971), 95–118.

[4] Knauer, U., *Projectivity of acts and Morita equivalence of monoids*, Semigroup Forum **3** (1972), 359–370.

[5] Knauer, U. and P. Normak, *Morita duality for monoids*, Semigroup Forum **40** (1990), 39–57.

[6] Kozhukhov, I. B., *Left chain semigroups*, Semigroup Forum **22** (1981), 1–8.

[7] Müller, B. J., *Morita duality - a survey*, Abelian groups and modules (Udine, 1984). CISM Courses and Lectures **287** (1984), 395–414.

[8] Schein, B. M., *Commutative Semigroups where Congruences Form a Chain*, Bulletin de l'Academie Polonaise de Sciences, Serie des Sciences Math., Astr. et Phys. **17** (1969), N° 9, 523–527.

[9] Schwarz, S., *On dual semigroups*, Czechosl. Math. J. **10** (85), 1960, N° 2, 201–230.

Informaatika kateeder
Tallinna Pedagoogikaülikool
EE0102 Tallinn, Estonia

Received Febuary 4, 1991
and in final form November 11, 1991

Semigroup Forum Vol. 45 (1992) 214–225

RESEARCH ARTICLE

On some presentations of completely regular semigroups

K. S. Ajan

Communicated by F. Pastijn

In this paper we consider three types of presentations of completely regular semigroups. In each of the considered cases the solution of the word problem can be reduced to the solution of the word problem for a corresponding group presentation. As a consequence, in each of these cases the one relator presentation has a solvable word problem.

1. Preliminaries

A *unary semigroup* is an algebra of type $\langle 2,1\rangle$ which for the binary operation of multiplication is a semigroup. The unary operation will be denoted by $^{-1}$. Let X be any non empty set and $U(X)$ the smallest subsemigroup of the free semigroup on the set $X \cup \{(,)^{-1}\}$ satisfying the following conditions:

(*i*) $X \subseteq U(X)$,

(*ii*) if $u \in U(X)$, then $(u)^{-1} \in U(X)$.

It is well known [3] that $U(X)$ is free on X in the variety $\mathcal{U}$ of all unary semigroups. The variety $\mathcal{CR}$ of *completely regular semigroups* consists of those unary semigroups which satisfy the identities

$$(x^{-1})^{-1} = x, \quad xx^{-1}x = x, \quad xx^{-1} = x^{-1}x. \tag{1}$$

Whenever this last identity is satisfied, we can use the notation $x^0 = xx^{-1} = x^{-1}x$; also, if S is a completely regular semigroup and $a \in S$, then we use the notation $a^0 = aa^{-1} = a^{-1}a$. The variety $\mathcal{O}$ of *orthogroups* consists of those completely regular semigroups which satisfy the additional identity

$$x^0y^0 = (x^0y^0)^0, \tag{2}$$

and the variety $\mathcal{CS}$ of *completely simple semigroups* consists of the completely regular semigroups which satisfy the identity

$$(xyx)^0 = x^0. \tag{3}$$

It is well-known that every completely regular semigroup in the disjoint union of its maximal subgroups and conversely, every unary semigroup which is a union of groups is completely regular. Thus, in particular, if S is completely regular and $a \in S$, then a^{-1} is the inverse of a within the maximal subgroup of S to which a belongs and a^0 is the identity element of this maximal subgroup. Consequently in a completely regular semigroup S we have $e = e^2$ if and only if $e = e^0$. Thus, orthogroups are precisely the completely regular semigroups in which the idempotents form a subsemigroup. Another consequence of the above remarks is that the variety $\mathcal{G}$ of groups consists of those completely regular semigroups satisfying the identity $x^0 = y^0$. It is well-known that a completely regular semigroup is completely simple if and only if it is simple as a semigroup. Thus, $\mathcal{G}$ is a subvariety of both $\mathcal{O}$ and $\mathcal{CS}$. More details and further references on completely regular semigroups may be found in [1], [2], [3] [4] and [8].

If $\mathcal{V}$ is any subvariety of $\mathcal{U}$, then we denote by $\rho_\mathcal{V}$ the least congruence on $U(X)$ such that $U(X)/\rho_\mathcal{V} \in \mathcal{V}$. Then $\rho_\mathcal{V}$ is a fully invariant congruence on $U(X)$ and $F\mathcal{V}(X) = U(X)/\rho_\mathcal{V}$ is free on X in $\mathcal{V}$. Let κ be a binary relation on $U(X)$, and for any variety $\mathcal{V} \subseteq \mathcal{U}$, let $\kappa_\mathcal{V} = \{(u\rho_\mathcal{V},\ v\rho_\mathcal{V}) \mid (u,v) \in \kappa\}$ and $\kappa_\mathcal{V}^*$ the congruence relation on $F\mathcal{V}(X)$ generated by $\kappa_\mathcal{V}$. We call $\langle X;\ \kappa_\mathcal{V}\rangle$ a $\mathcal{V}$-*presentation* and $F\mathcal{V}(X)/\kappa_\mathcal{V}^* \cong U(X)/(\kappa \cup \rho_\mathcal{V})^*$ the member of $\mathcal{V}$ determined by this presentation. We say that the $\mathcal{V}$-presentation $\langle X; \kappa_\mathcal{V}\rangle$ has a *solvable word problem* if for any $w_1, w_2 \in U(X)$ it is decidable whether $w_1(\kappa \cup \rho_\mathcal{V})^* w_2$, or equivalently, whether $(w_1\rho_\mathcal{V})\ \kappa_\mathcal{V}^*\ (w_2\rho_\mathcal{V})$. A one-relator $\mathcal{V}$-presentation is a presentation of the form $\langle X; \kappa_\mathcal{V}\rangle$ where κ is a singleton. From a theorem by Magnus it follows that one-relator $\mathcal{G}$-presentations have a solvable word problem (see e.g. [6], pp. 198–203). We shall see that the same holds true for one-relator $\mathcal{CS}$-presentations and for some particular one-relator $\mathcal{O}$-presentations. The variety of inverse semigroups is also a subvariety of $\mathcal{U}$. For results on inverse semigroup presentations, and in particular one relator inverse semigroup presentations, we refer to [7] and [9].

One can consider unary monoids instead of unary semigroups. A unary monoid is an algebra of type $\langle 2,1,0\rangle$ where the $\langle 2,1\rangle$-reduct is a unary semigroup, and where the nullary operation corresponds with a multiplicative identity element 1 satisfying $1^{-1} = 1$. Here $U(X)^1$ is free on X in the variety $\mathcal{MU}$ of unary monoids (the identity element 1 of $U(X)^1$ will be called the empty word). Again, for every subvariety $\mathcal{V}$ of $\mathcal{MU}$ one can introduce a $\rho_\mathcal{V}$ on $U(X)^1$ much in the same way as explained before. Also, in this context the notions of $\mathcal{V}$-presentation, solvable word problem, and so on, should now be clear. The variety $\mathcal{MCR}$ of *completely regular monoids* is the subvariety of $\mathcal{MU}$ determined by the identities (1) and the variety $\mathcal{MO}$ of *monoid orthogroups* is the subvariety of $\mathcal{MCR}$ determined by the additional identity (2). The variety of all groups, considered this time as unary monoids, will again be denoted by $\mathcal{G}$. Then $\mathcal{G}$ is the subvariety of $\mathcal{MCR}$ determined by the additional identity $x^0 = y^0$, or equivalently, by $x^0 = 1$. We observe that a completely regular monoid whose $\langle 2,1\rangle$-reduct is completely simple must be a group. Therefore $\mathcal{G}$ is also the subvariety of $\mathcal{MCR}$ determined by (3).

2. A presentation of monoid orthogroups.

In this section we shall be concerned with a finite relation $\kappa = \{(u_i, 1) \mid 0 \le i < n\}$ on $U(X)^1$ and we shall show that the $\mathcal{MO}$-presentation $\langle X; \kappa_{\mathcal{MO}}\rangle$ has a solvable word problem if and only if the group presentation $\langle X; \kappa_\mathcal{G}\rangle$ has a solvable word problem.

If κ is the empty relation, then the monoid orthogroup determined by the presentation $\langle X; \kappa_{\mathcal{MO}}\rangle$ is simply the free monoid orthogroup $F\mathcal{MO}(X) \cong U(X)^1/\rho_{\mathcal{MO}}$ on the set X. It is easy to see that $F\mathcal{MO}(X) = F\mathcal{O}(X)^1$ is the free orthogroup on X with an extra identity 1 adjoined. A solution of the word problem for this presentation has been obtained in [4].

For any $w \in (X \cup \{(,)^{-1}\})^+$, let $c(w)$ be the set of letters from the alphabet X which occur in w. $c(w)$ will be called the content of w. In particular, $c(1) = \emptyset$. It follows from [4] that

$$D_{w_1\rho_{\mathcal{MO}}} \leq D_{w_2\rho_{\mathcal{MO}}} \text{ in } F\mathcal{MO}(X) \Leftrightarrow c(w_2) \subseteq c(w_1). \tag{4}$$

The following result will be used several times; it is also mentioned in [4].

Lemma 1. *If S is a [monoid] orthogroup and $a, b, e \in S$ such that $e = e^2$ and $D_a = D_b \leq D_e$, then $ab = aeb$.* ■

Lemma 2. *Let $\kappa = \{(u_i, v_i) \mid 0 \leq i < n\}$ be any finite relation on $U(X)^1$. Let $w = u_0 \dots u_{n-1}\, v_0 \dots v_{n-1}$. Then for any $w_1, w_2 \in U(X)^1$,*

$$\begin{aligned}&(w_1\rho_{\mathcal{G}})\ \kappa_{\mathcal{G}}^*\ (w_2\rho_{\mathcal{G}}) \Leftrightarrow \\ &\quad (w_1w_2w)^0\, w_1\, (w_1w_2w)^0\, \rho_{\mathcal{MO}} \quad \kappa_{\mathcal{MO}}^* \quad (w_1w_2w)^0\, w_2\, (w_1w_2w)^0\, \rho_{\mathcal{MO}}.\end{aligned}$$

Proof. Let us first assume that $w_1, w_2 \in U(X)^1$ and $w_1\rho_{\mathcal{G}}\ \kappa_{\mathcal{G}}^*\ w_2\rho_{\mathcal{G}}$. Then there exists a sequence

$$w_1 = z_0, z_1, \dots, z_k = w_2$$

where for each $j < k$,

$$(z_j, z_{j+1}) \in \rho_{\mathcal{MO}}$$

or

$$\{z_j, z_{j+1}\} = \{xu_iy,\ xv_iy\} \quad \text{for some } x, y \in U(X)^1 \text{ and } i < n,$$

or

$$\{z_j, z_{j+1}\} = \{xuu^{-1}y,\ xy\} \quad \text{for some } x, y, u \in U(X)^1.$$

We can as well assume that the letters from the alphabet X which occur in $z_0, \dots, z_k$ all occur in w_1w_2w. If not, then we can subsitute all the occurrences in $z_0, \dots, z_k$ of letters which do not occur in w_1w_2w by 1 and the resulting sequence will be of the required form.

If $(z_j, z_{j+1}) \in \rho_{\mathcal{MO}}$, then

$$(w_1w_2w)^0 z_j (w_1w_2w)^0\, \rho_{\mathcal{MO}} = (w_1w_2w)^0\, z_{j+1} (w_1w_2w)^0\, \rho_{\mathcal{MO}}.$$

If $\{z_j, z_{j+1}\} = \{xu_iy, xv_iy\}$ for some $x, y \in U(X)^1$ and $i < n$, then

$$(z_j\rho_{\mathcal{MO}})\ \kappa_{\mathcal{MO}}^*\ (z_{j+1}\rho_{\mathcal{MO}})$$

and thus

$$(w_1w_2w)^0 z_j (w_1w_2w)^0 \rho_{\mathcal{MO}} \; \kappa^*_{\mathcal{MO}} \; (w_1w_2w)^0 z_{j+1} (w_1w_2w)^0 \rho_{\mathcal{MO}}.$$

Let $\{z_j, z_{j+1}\} = \{xuu^{-1}y,\ xy\}$ for some $x, y, u \in U(X)^1$. By (4) and Lemma 1 we then have $(w_1w_2w)^0 z_j (w_1w_2w)^0 \rho_{\mathcal{MO}} = (w_1w_2w)^0 z_{j+1} (w_1w_2w)^0 \rho_{\mathcal{MO}}$. Putting $(w_1w_2w)^0 z_j (w_1w_2w)^0 \rho_{\mathcal{MO}} = z'_j$ for every $0 \le j \le k$, we obtain a sequence

$$(w_1w_2w)^0 w_1 (w_1w_2w)^0 \rho_{\mathcal{MO}} = z'_0, z'_1, \ldots, z'_k = (w_1w_2w)^0 w_2 (w_1w_2w)^0 \rho_{\mathcal{MO}}$$

where for any $j < k$, $\quad z'_j = z'_{j+1}$ or $z'_j \, \kappa^*_{\mathcal{MO}} \, z'_{j+1}$. We conclude that

$$(w_1w_2w)^0 w_1 (w_1w_2w)^0 \rho_{\mathcal{MO}} \quad \kappa^*_{\mathcal{MO}} \quad (w_1w_2w)^0 w_2 (w_1w_2w)^0 \rho_{\mathcal{MO}}. \tag{5}$$

Conversely, let $w_1, w_2 \in U(X)^1$ be such that (5) holds. From $\rho_{\mathcal{MO}} \le \rho_{\mathcal{G}}$ it follows that

$$(w_1w_2w)^0 w_1 (w_1w_2w)^0 \rho_{\mathcal{G}} \quad \kappa^*_{\mathcal{G}} \quad (w_1w_2w)^0 w_2 (w_1w_2w)^0 \rho_{\mathcal{G}}.$$

From this and

$$\begin{aligned} (w_1w_2w)^0 w_1 (w_1w_2w)^0 \quad &\rho_{\mathcal{G}} \quad w_1, \\ (w_1w_2w)^0 w_2 (w_1w_2w)^0 \quad &\rho_{\mathcal{G}} \quad w_2, \end{aligned}$$

it then follows that $w_1\rho_{\mathcal{G}} \; \kappa^*_{\mathcal{G}} \; w_2\rho_{\mathcal{G}}$, as required. ∎

Corollary 3. *Let κ be a finite relation on $U(X)^1$ such that the presentation $\langle X; \kappa_{\mathcal{MO}}\rangle$ of monoid orthogroup has a solvable word problem. Then the group presentation $\langle X; \kappa_{\mathcal{G}}\rangle$ has a solvable word problem.* ∎

Let A be a subset of X and $w \in U(X)^1$. If $c(w) \cap A \ne \emptyset$, then for some $a \in A$ we can write $w = b\,a\,c$ in the free semigroup$(X \cup \{(,)^{-1}\})^+$, where $c(b) \cap A = (c(w) \cap A) - \{a\}$. We then use the notation $\bar{0}_A(w) = a$ and we denote by $0_A(w)$ the element of $U(X)^1$ which is obtained from b by deleting unmatched parentheses. Also, for some $a' \in A$ we can write $w = b'\,a'\,c'$ in the free semigroup $(X \cup \{(,)^{-1}\})^+$, where $c(c') \cap A = (c(w) \cap A) - \{a\}$. We use the notation $\bar{1}_A(w) = a'$, and $1_A(w)$ denotes the element of $U(X)^1$ which is obtained from c' by deleting unmatched parentheses. If $|c(w) \cap A| = k$, then we define $\bar{0}^\ell_A(w) = \bar{0}_A(\bar{0}^{\ell-1}_A(w))$ and $0^\ell_A(w) = 0_A(0^{\ell-1}_A(w))$ for $1 < \ell \le k$, with $\bar{0}^1_A = \bar{0}_A(w)$ and $0^1_A(w) = 0_A(w)$. $\bar{1}^\ell_A(w)$ and $1^\ell_A(w)$ are defined dually. It should be clear that if $A \subseteq B \subseteq X$ and $w \in U(X)^1$ with $c(w) \cap A \ne \emptyset$, then $\bar{0}_A(w) = \bar{0}^\ell_B(w)$ and $0_A(w) = 0^\ell_B(w)$ for some $\ell \ge 1$.

Theorem 4. *Let κ be the finite relation $\kappa = \{(u_i, 1) \mid 0 \le i < n\}$ on $U(X)^1$. Let $A = X - c(u_0 \ldots u_{n-1})$. Let $w_1, w_2 \in U(X)^1$. Then*

$$(w_1\rho_{\mathcal{MO}}) \; \kappa^*_{\mathcal{MO}} \; (w_2\rho_{\mathcal{MO}}) \Rightarrow c(w_1) \cup c(u_0 \ldots u_{n-1}) = c(w_2) \cup c(u_0 \ldots u_{n-1}). \tag{6}$$

If $c(w_1) \subseteq c(u_0 \ldots u_{n-1})$ and $c(w_2) \subseteq c(u_0 \ldots u_{n-1})$, then

$$w_1\rho_{\mathcal{MO}} \; \kappa^*_{\mathcal{MO}} \; w_2\rho_{\mathcal{MO}} \Leftrightarrow w_1\rho_{\mathcal{G}} \; \kappa^*_{\mathcal{G}} \; w_2\rho_{\mathcal{G}}. \tag{7}$$

Otherwise,

$$w_1\rho_{\mathcal{MO}}\ \kappa^*_{\mathcal{MO}}\ w_2\rho_{\mathcal{MO}} \Leftrightarrow$$

$$\begin{aligned}
&(i)\ \ c(w_1) \cup c(u_0 \dots u_{n-1}) = c(w_2) \cup c(u_0 \dots u_{n-1}),\\
&(ii)\ \ \bar{0}_A(w_1) = \bar{0}_A(w_2), \quad \bar{1}_A(w_1) = \bar{1}_A(w_2),\\
&(iii)\ 0_A(w_1)\rho_{\mathcal{MO}}\ \ \kappa^*_{\mathcal{MO}}\ \ 0_A(w_2)\rho_{\mathcal{MO}},\ 1_A(w_1)\rho_{\mathcal{MO}}\ \ \kappa^*_{\mathcal{MO}}\ \ 1_A(w_2)\rho_{\mathcal{MO}},
\end{aligned}$$

and

$$(iv)\ w_1\rho_{\mathcal{G}}\ \kappa^*_{\mathcal{G}}\ w_2\rho_{\mathcal{G}}.$$

Remark. If κ is empty, then above reduces to a description of $F\mathcal{MO}(X) \cong U(X)^1/\rho_{\mathcal{MO}}$ ($\cong (U(X)/\rho_{\mathcal{O}})^1$) which can be derived immediately from the description of $F\mathcal{O}(X)$ which is given in [4].

Proof of Theorem 4. Let $w_1, w_2 \in U(X)^1$ be such that $w_1\rho_{\mathcal{MO}}\ \kappa^*_{\mathcal{MO}}\ w_2\rho_{\mathcal{MO}}$. Then there exists a sequence

$$w_1 = z_0, z_1, \dots, z_k = w_2$$

where for each $j < k$,

$$(z_j, z_{j+1}) \in \rho_{\mathcal{MO}} \tag{8}$$

or

$$\{z_j, z_{j+1}\} = \{xu_iy, xy\} \text{ for some } x, y \in U(X)^1 \text{ and } i < n. \tag{9}$$

If $(z_j, z_{j+1}) \in \rho_{\mathcal{MO}}$, then $c(z_j) = c(z_{j+1})$, and if (9) holds, then $c(z_j) \cup c(u_i) = c(z_{j+1}) \cup c(u_i)$. Therefore (6) is satisfied. We know from the fact that $\rho_{\mathcal{MO}} \subseteq \rho_{\mathcal{G}}$, that $(w_1\rho_{\mathcal{G}})\ \kappa^*_{\mathcal{G}}\ (w_2\rho_{\mathcal{G}})$.

Let us assume that $c(w_1),\ c(w_2) \not\subseteq c(u_0 \dots u_{n-1})$. If (8) holds, then we know from [4] that $\bar{0}^\ell_X(z_j) = \bar{0}^\ell_X(z_{j+1})$ and $0^\ell_X(z_j)\ \rho_{\mathcal{MO}}\ 0^\ell_X(z_{j+1})$ for every $\ell \le |c(z_j)| = |c(z_{j+1})|$. Thus for every such ℓ we have

$$\bar{0}^\ell_X(z_j) \in A \Leftrightarrow 0^\ell_X(z_{j+1}) \in A,$$

whence,

$$\bar{0}_A(z_j) = \bar{0}^\ell_X(z_j) = \bar{0}^\ell_X(z_{j+1}) = \bar{0}_A(z_{j+1})$$

for some $\ell \le |c(z_j)| = |c(z_{j+1})|$ and then

$$0_A(z_j) = 0^\ell(z_j)\ \rho_{\mathcal{MO}}\ 0^\ell_X(z_{j+1}) = 0_A(z_{j+1}).$$

If (9) holds, then either $\bar{0}_A(z_j) = \bar{0}_A(z_{j+1}) = \bar{0}_A(x)$ and then $0_A(z_j) = 0_A(z_{j+1}) = 0_A(x)$, or $\bar{0}_A(z_j) = \bar{0}_A(z_{j+1}) = \bar{0}_A(xy)$, and in this case

$$\left\{0_A(z_j), 0_A(z_{j+1})\right\} = \left\{xu_i0_{A-(c(x)\cap A)}(y), x0_{A-(c(x)\cap A)}(y)\right\}.$$

Thus, in all cases,

$$\bar{0}_A(z_j) = \bar{0}_A(z_{j+1}) \quad \text{and} \quad 0_A(z_j)\rho_{\mathcal{MO}}\ \ \kappa^*_{\mathcal{MO}}\ \ 0_A(z_{j+1})\rho_{\mathcal{MO}}.$$

Therefore,

$$\bar{0}_A(w_1) = \bar{0}_A(w_2) \quad \text{and} \quad 0_A(w_1)\rho_{\mathcal{MO}} \quad \kappa^*_{\mathcal{MO}} \quad 0_A(w_2)\rho_{\mathcal{MO}}.$$

Dually,

$$\bar{1}_A(w_1) = \bar{1}_A(w_2) \quad \text{and} \quad 1_A(w_1)\rho_{\mathcal{MO}} \quad \kappa^*_{\mathcal{MO}} \quad 1_A(w_2)\rho_{\mathcal{MO}}.$$

We now proceed to show that the converse holds. We first consider the case where $c(w_1), c(w_2) \subseteq c(u_0 \dots u_{n-1})$ and $w_1\rho_{\mathcal{G}}\ \kappa^*_{\mathcal{G}}\ w_2\rho_{\mathcal{G}}$. Then by Lemma 2,

$$(w_1w_2w)^0\, w_1\, (w_1w_2w)^0\, \rho_{\mathcal{MO}} \quad \kappa^*_{\mathcal{MO}} \quad (w_1w_2w)^0\, w_2\, (w_1w_2w)^0\, \rho_{\mathcal{MO}}, \tag{10}$$

where $w = u_0 \dots u_{n-1}$. It it easy to see that $u^0\rho_{\mathcal{MO}}\ \kappa^*_{\mathcal{MO}}\ 1\rho_{\mathcal{MO}}$ for every u with $c(u) \subseteq c(u_0 \dots u_{n-1})$. Thus,

$$(w_1w_2w)^0\, w_i(w_1w_2w)^0\, \rho_{\mathcal{MO}} \quad \kappa^*_{\mathcal{MO}} \quad w_i\rho_{\mathcal{MO}}, \qquad i = 1, 2,$$

which by (10) implies that $(w_1\rho_{\mathcal{MO}})\ \kappa^*_{\mathcal{MO}}\ (w_2\rho_{\mathcal{MO}})$.

We next consider the case where $c(w_1), c(w_2) \not\subseteq c(u_0 \dots u_{n-1})$ and that the conditions (i)–(iv) are satisfied. Then, using $w = u_0 \dots u_{n-1}$, we have

$$\begin{aligned}
w_1\rho_{\mathcal{MO}} \quad &= (0_A(w_1)\, \bar{0}_A(w_1))^0\, w_1\, (\bar{1}_A(w_1)\, 1_A(w_1)^0\, \rho_{\mathcal{MO}} \\
&\qquad \text{(by [4])} \\
&\kappa^*_{\mathcal{MO}}\ (0_A(w_1))\bar{0}_A(w_1)w)^0\, w_1\, (w\bar{1}_A(w_1)\, 1_A(w_1))^0\, \rho_{\mathcal{MO}} \\
&= (0_A(w_1)\bar{0}_A(w_1)w)^0(w_1w_2w)^0\, w_1\, (w_1w_2w)^0(w\bar{1}_A(w_1)1_A(w_1))^0\, \rho_{\mathcal{MO}} \\
&\qquad \text{(by } (i) \text{ and Lemma 1)} \\
&\kappa^*_{\mathcal{MO}}\ (0_A(w_1)\bar{0}_A(w_1)w)^0(w_1w_2w)^0\, w_2\, (w_1w_2w)^0(w\bar{1}_A(w_1)1_A(w_1))^0\, \rho_{\mathcal{MO}} \\
&\qquad \text{(by } (iv) \text{ and Lemma 2)} \\
&\kappa^*_{\mathcal{MO}}\ (0_A(w_2)\bar{0}_A(w_2)w)^0(w_1w_2w)^0\, w_2\, (w_1w_2w)^0(w\bar{1}_A(w_2)1_A(w_2))^0\, \rho_{\mathcal{MO}} \\
&\qquad \text{(by } (ii) \text{ and } (iii)\text{)} \\
&= (0_A(w_2)\bar{0}_A(w_2)w)^0\, w_2\, (w\bar{1}_A(w_2)\, 1_A(w_2))^0\, \rho_{\mathcal{MO}} \\
&\qquad \text{(by } (i) \text{ and Lemma 1)} \\
&\kappa^*_{\mathcal{MO}}\ w_1\rho_{\mathcal{MO}},
\end{aligned}$$

as required. ■

By the foregoing we thus obtain the following converse to Corollary 3 in this special situation.

Corollary 5. *Let κ be the finite relation $\kappa = \{(u_i, 1) \mid 0 \leq i < n\}$ on $U(X)^1$. Then the $\mathcal{MO}$-presentation $\langle X;\ \kappa_{\mathcal{MO}}\rangle$ has a solvable word problem if and only if the group presentation $\langle X;\ \kappa_{\mathcal{G}}\rangle$ has a solvable word problem.* ■

In particular, by Magnus's theorem ([6, pp. 198–203]),

Corollary 6. *Let $\kappa = \{(u, 1)\}$ for some $u \in U(X)^1$. Then the $\mathcal{MO}$-presentation $\langle X;\ \kappa_{\mathcal{MO}}\rangle$ has a solvable word problem.* ■

3. A presentation of orthogroups.

In this section we consider a relation κ of the form

$$\kappa = \{(u_i, u_i(u_i)^{-1}) \mid 0 \leq i < n\}$$

on $U(X)$ and we shall show that the orthogroup presentation $\langle X;\ \kappa_{\mathcal{O}}\rangle$ has a solvable word problem if and only if some group presentation has a solvable word problem.

The notation introduced in the preceding section will be used again. An obvious modification of the proof of Lemma 2 yields

Lemma 7. *Let $\kappa = \{(u_i, v_i) \mid 0 \leq i < n\}$ be any finite relation on $U(X)$. Let $w = u_0 \ldots u_{n-1}\, v_0 \ldots v_{n-1}$. Then for any $w_1, w_2 \in U(X)$,*

$$w_1\rho_{\mathcal{G}}\ \kappa_{\mathcal{G}}^*\ w_2\rho_{\mathcal{G}} \Leftrightarrow (w_1w_2w)^0\, w_1\, (w_1w_2w)^0\rho_{\mathcal{O}}\quad \kappa_{\mathcal{O}}^*\quad (w_1w_2w)^0\, w_2\, (w_1w_2w)^0\rho_{\mathcal{O}}. \quad \blacksquare$$

Consequently,

Corollary 8. *Let κ be a finite relation on $U(X)$ such that the orthogroup presentation $\langle X;\ \kappa_{\mathcal{O}}\rangle$ has a solvable word problem. Then the group presentation $\langle X;\ \kappa_{\mathcal{G}}\rangle$ has a solvable word problem.* $\blacksquare$

Again, the following theorem reduces to the description of $F\mathcal{O}(X) \cong U(X)/\rho_{\mathcal{O}}$ of [4] if $\kappa = \emptyset$, that is, if $n = 0$.

Theorem 9. *Let κ be the finite relation $\kappa = \{(u_i, u_i(u_i)^{-1}) \mid 0 \leq i < n\}$. For $A \subseteq X$, put $I_A = \{i \mid c(u_i) \subseteq A\}$ and $\kappa_A = \{u_i, u_i(u_i)^{-1}) \mid i \in I_A\}$. Then for $w_1, w_2 \in U(X)$, we have*

$$\begin{array}{lll} w_1\rho_{\mathcal{O}} & \kappa_{\mathcal{O}}^* & w_2\rho_{\mathcal{O}} \Leftrightarrow \\ & (i) & c(w_1) = c(w_2), \\ & (ii) & 0_X(w_1)\rho_{\mathcal{O}}\ \kappa_{\mathcal{O}}^*\ 0_X(w_2)\rho_{\mathcal{O}},\ 1_X(w_1)\rho_{\mathcal{O}}\ \kappa_{\mathcal{O}}^*\ 1_X(w_2)\rho_{\mathcal{O}}, \\ \textit{and} & & \\ & (iii) & w_1\rho_{\mathcal{G}}\ (\kappa_{c(w_1)})_{\mathcal{G}}^*\ w_2\rho_{\mathcal{G}}. \end{array}$$

Proof. Let $w_1\rho_{\mathcal{O}}\ \ \kappa_{\mathcal{O}}^*\ \ w_2\rho_{\mathcal{O}}$. Then there exists a sequence

$$w_1 = z_0, z_1, \ldots, z_k = w_2,$$

where for each $j < k$,

$$(z_j, z_{j+1}) \in \rho_{\mathcal{O}} \tag{11}$$

or

$$\{z_j, z_{j+1}\} = \{xu_i(u_i)^{-1}y, xu_iy\} \quad \text{for some } x, y \in U(X)^1 \text{ and } i < n. \tag{12}$$

If (11) holds, then by [4],

$$c(z_j) = c(z_{j+1}),\ 0_X(z_j)\rho_{\mathcal{O}} = 0_X(z_{j+1})\rho_{\mathcal{O}},\ 1_X(z_j)\rho_{\mathcal{O}} = 1_X(z_{j+1})\rho_{\mathcal{O}},\ z_j\rho_{\mathcal{G}} = z_{j+1}\rho_{\mathcal{G}}.$$

If (12) holds, then obviously $c(z_j) = c(z_{j+1})$. In view of the above we thus have that $c(w_1) = c(z_\ell) = c(w_2)$ for all $0 \leq \ell \leq k$. By (12) we have that $c(u_i) \subseteq c(w_1)$, thus

$i \in I_{c(w_1)}$ and $(u_i,\ u_i(u_i)^{-1}) \in \kappa_{c(w_1)}$. Therefore $(z_j\rho_{\mathcal{G}})(\kappa_{c(w_1)})^*_{\mathcal{G}}(z_{j+1}\rho_{\mathcal{G}})$. If $0_X(z_j)$ is an initial segment of xu_i, then $0_X(z_j) = 0_X(z_{j+1}) = x0_{X-c(x)}(u_i)$. Otherwise

$$\{0_X(z_j), 0_X(z_{j+1})\} = \{xu_i(u_i)^{-1}\, 0_{X-c(xu_i)}(y), xu_i0_{X-c(xu_i)}(y)\}$$

and so $0_X(z_j)\rho_{\mathcal{O}}\ \kappa^*_{\mathcal{O}}\ 0_X(z_{j+1})\rho_{\mathcal{O}}$. Dually, $1_X(z_j)\rho_{\mathcal{O}}\ \ \kappa^*_{\mathcal{O}}\ \ 1_X(z_{j+1})\rho_{\mathcal{O}}$. Therefore the conditions $(i), (ii), (iii)$ hold.

Let us now assume that, conversely, the conditions $(i), (ii)$ and (iii) hold. From (iii) it follows that there exists a sequence

$$w_1 = z_0, z_1, \ldots, z_k = w_2,$$

where for each $j < k$,

$$(z_j, z_{j+1}) \in \rho_{\mathcal{O}}, \tag{13}$$

or

$$\{z_j, z_{j+1}\} = \{xu_i(u_i)^{-1}y, xu_iy\} \text{ for some } x, y \in U(X)^1 \text{ and } i \in I_{c(w_1)}, \tag{14}$$

or

$$\{z_j, z_{j+1}\} = \{xa(a)^{-1}y, xy\} \text{ for some } x, y \in U(X)^1 \text{ and } a \in U(X). \tag{15}$$

One can always assume that the letters which occur in the $a \in U(X)$ which are considered in transitions of the form (15) are all contained in $c(w_1) = c(w_2)$. If not, one can substitute all the occurrences in the z_ℓ, $0 < \ell < k$, of the letters which do not occur in $c(w_1) = c(w_2)$ by some letter of $c(w_1)$, and the resulting sequence would be of the required form. Thus we can assume that $c(z_\ell) \subseteq c(w_1) = c(w_2)$ for all $0 \leq \ell < k$.

If (13) or (15) holds, then by Lemma 1,

$$\begin{aligned}&(0_X(w_1)\,\bar{0}_X(w_1))^0\, z_j\, (\bar{1}_X(w_1)\ 1_X(w_1))^0\ \rho_{\mathcal{O}}\\ &\quad = (0_X(w_1)\,\bar{0}_X(w_1))^0\, z_{j+1}\, (\bar{1}_X(w_1)\ 1_X(w_1))^0\ \rho_{\mathcal{O}}.\end{aligned}$$

and if (14) holds, then

$$\begin{aligned}&(0_X(w_1)\,\bar{0}_X(w_1))^0\, z_j\, (\bar{1}_X(w_1)\ 1_X(w_1))^0\ \rho_{\mathcal{O}}\\ &\kappa^*_{\mathcal{O}}\ (0_X(w_1)\,\bar{0}_X(w_1))^0\, z_{j+1}\, (\bar{1}_X(w_1)\ 1_X(w_1))^0\ \rho_{\mathcal{O}}.\end{aligned}$$

Therefore,

$$\begin{aligned} w_1\rho_{\mathcal{O}} &= (0_X(w_1)\,\bar{0}_X(w_1))^0\ w_1\ (\bar{1}_X(w_1)\ 1_X(w))^0\,\rho_{\mathcal{O}}\\ &\qquad\qquad \text{(by [4])}\\ &\kappa^*_{\mathcal{O}}\ (0_X(w_1)\bar{0}_X(w_1))^0\, w_2\, (\bar{1}_X(w_1)\ 1_X(w_1))^0\,\rho_{\mathcal{O}}\\ &\kappa^*_{\mathcal{O}}\ (0_X(w_2)\bar{0}_X(w_2))^0\, w_2\, (\bar{1}_X(w_2)\ 1_X(w_2))^0\,\rho_{\mathcal{O}}\\ &\qquad\qquad \text{(by } (ii))\\ &= w_2\rho_{\mathcal{O}} \qquad\quad \text{(by [4])}. \end{aligned}$$

■

Corollary 10. *Let κ be the finite relation $\kappa = \{(u_i, u_i(u_i)^{-1}) \mid 0 \leq i < n\}$. Then the orthogroup presentation $\langle X;\ \kappa_{\mathcal{O}}\rangle$ has a solvable word problem if and only if the group presentations $\langle X; (\kappa_A)_{\mathcal{G}}\rangle$, $A \subseteq X$, have a solvable word problem.* ∎

Remark. Since κ is a finite relation, the number of group presentations which must be considered in Corollary 10 is finite.

Corollary 11. *Let $\kappa = \{(u, u(u)^{-1})\}$ for some $u \in U(X)$. Then the orthogroup presentation $\langle X;\ \kappa_{\mathcal{O}}\rangle$ has a solvable word problem.*

Proof. The proof follows immediately from Magnus's theorem ([6], pp. 198–203) and Corollary 10.

Example. If X is a finite set and $\kappa = \{(x, x(x)^{-1}) \mid x \in X\}$, then the orthogroup presentation $\langle X;\ \kappa_{\mathcal{O}}\rangle$ has a solvable word problem. Of course $F\mathcal{O}(X)/\kappa^*_{\mathcal{O}}$ is the free band on X.

Example. It is easy to find an example of a relation $\kappa = \{(u_i, u_i(u_i)^{-1}) \mid 0 \leq i < n\}$ on $U(X)$ such that the converse of Corollary 8 is not satisfied. Indeed, let $X = \{x_0, \ldots, x_\ell\}$ be an $\ell + 1$-element set, and let $\kappa = \{(u_i(u_i)^{-1}, u_i) \mid 0 \leq i < n\}$ be a relation on $U(X - \{x_0\})$ such that the group presentation $\langle X;\ \kappa_{\mathcal{G}}\rangle$ does not have a solvable word problem. Let $\bar{\kappa} = \kappa \cup \{(u_{pq}(u_{pq})^{-1}, u_{pq}) \mid 1 \leq p, q \leq \ell\}$, where $u_{pq} = x_0 x_p (x_0 x_q)^{-1}$. Then the group presentation $\langle X;\ \bar{\kappa}_{\mathcal{G}}\rangle$ has a solvable word problem and $F\mathcal{G}(X)/\bar{\kappa}^*_{\mathcal{G}}$ is the free product of two cyclic groups. The orthogroup presentation $\langle X;\ \bar{\kappa}_{\mathcal{O}}\rangle$ however does not have a solvable word problem.

4. A presentation of completely simple semigroups

We consider a finite relation κ of the form $\kappa = \{u_i, v_i) \mid 0 \leq i < n\}$ on $U(X)$. We now describe a model for the free completely simple semigroup on the set X as given in [1].

We let $X = \{x_0, x_1, x_2, \ldots\}$ and $Y = \{p_{x,y} \mid x, y \in X - \{x_0\}\}$ a set which is disjoint from X and in one-to-one correspondence with $(X - \{x_0\})^2$. Let F be a free group on $X \cup Y$ and let $Q = (q_{x,y})$ be a $X \times X$-matrix with $q_{x,y} = p_{x,y}$ if $x \neq x_0 \neq y$ and $q_{x,y} = 1_F$, the identity of F, otherwise. We next consider the Rees matrix semigroup $\mathcal{M}(F; X, X; Q)$ and the embedding $\iota : X \to \mathcal{M}(F; X, X; Q)$, $x \longrightarrow (x, x, x)$. Then there exists a unique homomorphism $\varphi : U(X) \to \mathcal{M}(F; X, X; Q)$ which extends ι. It follows from [1] that $\varphi\varphi^{-1} = \rho_{\mathcal{CS}}$, thus, $F\mathcal{CS}(X) = U(X)/\rho_{\mathcal{CS}} \cong \mathcal{M}(F; X, X; Q)$.

We shall henceforth identify $F\mathcal{CS}(X)$ with $\mathcal{M}(F; X, X; Q)$, so that $\kappa_{\mathcal{CS}} = \{(u_i\varphi, v_i\varphi) \mid 0 \leq i < n\}$. For each $0 \leq i < n$,

$$(x_0 u_i x_0)(x_0 v_i x_0)^{-1}\varphi = (x_0, w_i, x_0) \tag{16}$$

for some $w_i \in F$. We observe that it is routine to calculate the $u_i\varphi, v_i\varphi$ and w_i. For each $u \in U(X)$, let $h(u)$ — the head of u — be the first letter of X to appear in u

and $t(u)$ — the tail of u — the last element of X to appear in u. The equivalence relation on X generated by $\{(h(u_i), h(v_i)) \mid 0 \leq i < n\}$ will be denoted by κ_h and the equivalence relation generated by $\{(t(u_i), t(v_i)) \mid 0 \leq i < n\}$ will be denoted by κ_t. On F we consider the relation

$$\begin{aligned} \gamma &= \{(w_i, 1_F) \mid 0 \leq i < n\} \\ &\quad \cup\{(q_{x,y_1}, q_{x,y_2}) \mid x \in X,\ (y_1, y_2) \in \kappa_h\} \\ &\quad \cup\{(q_{x_1,y}, q_{x_2,y}) \mid y \in X,\ (x_1, x_2) \in \kappa_t\}. \end{aligned} \tag{17}$$

We see that if X is finite, then γ is a finite relation.

Theorem 12. *For (x,z,y), $(x',z',y') \in \mathcal{FCS}(X)$, we have*

$$(x,z,y)\ \kappa_{\mathcal{CS}}^*\ (x',z',y') \Leftrightarrow x\ \kappa_h\ x', \quad y\ \kappa_t\ y', \quad z\ \gamma^*\ z'.$$

Proof. Let τ be the relation on $\mathcal{FCS}(X)$ which is given by

$$(x,z,y)\ \tau\ (x',z',y') \Leftrightarrow x\ \kappa_h\ x', \quad y\ \kappa_t\ y', \quad z\ \gamma^*\ z'.$$

From the descriptions of the congruence relations on completely simple semigroups given in [2] or [5], it is easy to see that τ is a congruence relation. We must show that $\kappa_{\mathcal{CS}}^* = \tau$.

For some $0 \leq i < n$, we consider $(u_i\varphi, v_i\varphi) \in \kappa_{\mathcal{CS}}$. Using [1], we know that

$$u_0\varphi = (h(u_i),\ s_i, t(u_i)),\ v_i\varphi = (h(v_i),\ t_i, t(u_i)) \tag{18}$$

for some $s_i, t_i \in F$. Clearly $h(u_i)\ \kappa_h\ h(v_i)$ and $t(u_i)\ \kappa_t\ t(v_i)$, whereas

$$\begin{aligned} (x_0, w_i, x_0) &= (x_0 u_i x_0)(x_0\, v_i\, x_0)^{-1}\varphi \\ &= (x_0, x_0, x_0)(h(u_i), s_i, t(u_i))(x_0, x_0, x_0) \\ &\qquad \cdot\ ((x_0, x_0, x_0)(h(v_i), t_i, t(v_i))(x_0, x_0, x_0))^{-1} \\ &= (x_0, x_0 q_{x_0,h(u_i)}\ s_i\ q_{t(u_i),x_0}\ x_0, x_0) \\ &\qquad \cdot\ (x_0, x_0 q_{x_0,h(v_i)}\ t_i q_{t(v_i),x_0}\ x_0, x_0)^{-1} \\ &= (x_0, x_0 s_i x_0, x_0)(x_0, x_0 t_i x_0, x_0)^{-1} \\ &= (x_0, x_0 s_i x_0, x_0)(x_0, (x_0 t_i x_0)^{-1}, x_0) \\ &= (x_0, x_0 s_i t_i^{-1}\, x_0^{-1}, x_0) \end{aligned}$$

entails that $1_F\, \gamma^*\, w_i = x_o s_i t_i^{-1}\, x_0^{-1}$, and thus $s_i\ \gamma^*\ t_i$. Therefore $u_i\varphi\ \tau\ v_i\varphi$, and we can conclude that $\kappa_{\mathcal{CS}}^* \subseteq \tau$.

We now set out to prove the converse. From (16) it follows that

$$\begin{aligned} (x_0, w_i, x_0) &= (x_0 u_i x_0)(x_0 v_i x_0)^{-1}\,\varphi \\ &\quad \kappa_{\mathcal{CS}}^*\ (x_0 u_i x_0)(x_0 u_i x_0)^{-1}\,\varphi = (x_0), 1_F, x_0). \end{aligned} \tag{19}$$

From $u_i\varphi\ \kappa_{\mathcal{CS}}\ v_i\varphi$ it follows that $(u_i\, x\, \varphi)^0\ \kappa_{\mathcal{CS}}^*\ (v_i\, x\, \varphi)^0$, for every $x \in X$, that is,

$$(h(u_i), q_{x,h(u_i)}^{-1}, x)\ \kappa_{\mathcal{CS}}^*\ (h(v_i), q_{x,h(v_i)}^{-1}, x). \tag{20}$$

Hence,

$$(x_0,\ q^{-1}_{x,h(u_i)},\ x_0) = (x_0, 1_F, x_0)(h(u_i), q^{-1}_{x,h(u_i)}, x)(x_0, 1_F, x_0)$$
$$\kappa^*_{CS}\ (x_0, 1_F, x_0)(h(v_i), q^{-1}_{x,h(v_i)}, x)(x_0, 1_F, x_0) = (x_0, q^{-1}_{x,h(v_i)}, x_0).$$

Consequently,

$$(x_0, q_{x,y_1}, x_0)\ \kappa^*_{CS}\ (x_0, q_{x,y_2}, x_0) \tag{21}$$

for every $x \in X$, $(y_1, y_2) \in \kappa_h$. Dually,

$$(x_0, q_{x_1,y}, x_0) \quad \kappa^*_{CS} \quad (x_0, q_{x_2,y}, x_0) \tag{22}$$

for every $y \in X$, $(x_1, x_2) \in \kappa_t$. From (19), (21) and (22) we conclude that

$$z\ \gamma^*\ z' \Rightarrow (x_0, z, x_0)\ \kappa^*_{CS}\ (x_0, z', x_0).$$

Thus, if $z\ \gamma^*\ z'$, then

$$(x_0, q_{x_0,h(u_i)}z, x_0)\ \kappa^*_{CS}\ (x_0, q_{x_0,h(v_i)}z', x_0)$$

and using (20), we find for every $y \in X$,

$$\begin{aligned}(h(u_i), z, y) &= (h(u_i), q^{-1}_{x_0,h(u_i)}, x_0)(x_0, q_{x_0,h(u_i)}\, z, x_0)(x_0, 1_F, y)\\ &\ \kappa^*_{CS}\ (h(v_i), q^{-1}_{x_0,h(v_i)}, x)(x_0, q_{x_0,h(v_i)}z', x_0)(x_0, 1_F, y)\\ &= (h(v_i), z', y).\end{aligned}$$

Using induction we thus obtain

$$z\ \gamma^*\ z',\ x\ \kappa_h x' \Rightarrow (x, z, y)\ \kappa^*_{CS}\ (x', z', y)$$

for every $y \in X$. From this we have by duality that $\tau \subseteq \kappa^*_{CS}$. ■

Corollary 13. *Let $\kappa = \{(u_i, v_i) \mid 0 \le i < n\}$ be a finite relation on $U(X)$. Let γ be the relation given by (16) and (17). Then the completely simple semigroup presentation $\langle X;\ \kappa_{CS}\rangle$ has a solvable word problem if and only if the group presentation $\langle X \cup Y;\ \gamma\rangle$ has a solvable word problem.* ■

Corollary 14. *Let $\kappa = \{(u, v)\}$ for some $u, v \in U(X)$. Then the completely simple semigroup presentation $\langle X;\ \kappa_{CS}\rangle$ has a solvable word problem.* ■

Acknowledgement. The author is grateful to Prof. F. Pastijn for some valuable comments.

References

[1] Clifford, A. H., *The free completely regular semigroup on a set*, J. Algebra **59** (1979), 434–451.

[2] Clifford, A. H., and G. B. Preston, *The Algebraic Theory of Semigroups*, Vol. I, Math. Surveys **7**, Amer. Math. Soc., Providence, 1961.

[3] Gerhard, J. A., *Free completely regular semigroups I*, Representation, J. Algebra **82** (1983), 135–142.

[4] Gerhard, J. A., and M. Petrich, *The word problem for orthogroups*, Canad. J. Math. **33** (1981), 893–900.

[5] Howie, J. M., *An Introduction to Semigroup Theory*, Academic Press, New York, 1976.

[6] Lyndon, R. C., and P. E. Schupp, *Combinatorial Group Theory*, Springer-Verlag, New York, 1977.

[7] Meakin, J., *Automata and word problems*, in: J. E. Pin, ed., Formal Properties of Finite Automata and Applications, Lecture Notes in Computer Science **386**, Springer-Verlag, New York, 1989, 89–103.

[8] Pastijn, F., *The lattice of completely regular semigroup varieties*, J. Austral. Math. Soc. (Series A) **49** (1990), 24–42.

[9] Stephen, J. B., *Presentations of inverse monoids*, J. Pure Appl. Algebra **63** (1990), 81–112.

Marquette University,
Department of Mathematics,
Statistics & Computer Science
Milwaukee, WI 53233

Received February 26, 1991.
and in final form May 31, 1991.

Semigroup Forum Vol. 45 (1992) 226–240

RESEARCH ARTICLE

Reduction of a mean ergodic Markov semigroup on $C(X)$ into its irreducible components

Shizuo Miyajima

Communicated by R. Nagel

Introduction

Classical Perron-Frobenius theory for positive matrices (see e.g., [3], [10]) has been generalized for positive operators on ordered Banach spaces, especially Banach lattices. Among these results, a reduction theory for a single strongly mean ergodic Markov operator on $C(X)$ was given by Sawashima–Niiro [9]. The purpose of this paper is to give the semigroup version of the results in [9]. Throughout this paper, X denotes a compact Hausdorff space and $C(X)$ the complex Banach lattice of all continuous functions on X. A C_0-semigroup $\mathcal{T} = (T(t))_{t\geq 0}$ of bounded linear operators on $C(X)$ is called a Markov semigroup if each $T(t)$ is positive $(T(t)f \geq 0$ if $f \geq 0)$ and satisfies $T(t)\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ denotes the constant function with value 1 (*cf.* Davies [1], Chap.7). A typical example of a Markov semigroup is the one induced by a continuous semiflow $\Psi = \{\psi_t\}_{t\geq 0}$ on X (see [7], B-II, §3). Recall that $\Psi = \{\psi_t\}_{t\geq 0}$ is a continuous semiflow on X if $(t,x) \mapsto \psi_t(x)$ is a continuous mapping from $\mathbb{R}_+ \times X$ into X, and $\psi_{t+s} = \psi_t \circ \psi_s, \psi_0 = id$ hold. Then by setting $T(t)f := f \circ \psi_t$, $\mathcal{T} := (T(t))_{t\geq 0}$ becomes a Markov semigroup on $C(X)$.

Now we explain our results. Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a Markov semigroup on $C(X)$ (not necessarily induced by a semiflow) and let $(1/t)\int_0^t T(s)ds$ converge strongly as $t \to \infty$. Then there exists a mutually disjoint family $\{S_\lambda\}_{\lambda\in\Lambda}$ of closed subsets of X parametrized by a compact set Λ, for which $\mathcal{T}$ naturally induces an irreducible Markov semigroup $\mathcal{T}_\lambda$ on $C(S_\lambda)$ for each $\lambda \in \Lambda$ (§1, Theorem 1.3). In case $\mathcal{T}$ is induced by a continuous semiflow $\Psi = \{\psi_t\}_{t\geq 0}$, $S_\lambda(\lambda \in \Lambda)$ is a minimal Ψ-invariant closed set, i.e., $\psi_t(S_\lambda) \subset S_\lambda$ for any $t \geq 0$ and the orbit $\{\psi_t(x) \mid t \geq 0\}$ is dense in S_λ for every $x \in S_\lambda$. Thus our reduction may be considered as a generalization of classical decomposition into ergodic components. Returning to the general case and if we further assume that $(1/t)\int_0^t T(s)\,ds$ converges in operator norm as $t \to \infty$, then the boundary spectrum $\sigma(A) \cap i\mathbb{R}$ ([7], B-III, §2) of the generator A of $\mathcal{T}$ is determined by the spectra of the generators of $\mathcal{T}_\lambda$'s (§2, Theorem 2.1).

As [9] was based on Perron-Frobenius theory for irreducible positive operators on Banach lattices (Niiro–Sawashima[8], Schaefer [10,Chap.5]), this paper depends on the semigroup version of Perron-Frobenius theory established by G. Greiner[4], and the method in [4] to use positive pseudo-resolvents works effectively for our purpose.

In §1, the necessary definitions are given and the reduction procedure is described. An example of such reduction of a mean ergodic Markov semigroup on the 2-dimensional sphere is also presented. The result about the boundary spectrum is proved in §2, where fundamental facts about positive pseudo-resolvents are recalled beforehand. For standard notions concerning Banach lattices and positive semigroups, we refer the reader to Schaefer [10] or Nagel(ed.)[7].

1. Reduction into irreducible components

A Markov semigroup $\mathcal{T} = (T(t))_{t\geq 0}$ on $C(X)$ is called strongly mean ergodic if it satisfies the following condition :

$$\frac{1}{t}\int_0^t T(s)ds \text{ converges strongly to an operator } P \text{ as } t \to \infty .$$

The operator P will be called the limit projection of $\mathcal{T}$. $\mathcal{T}$ is called irreducible if there exists no non-trivial closed $\mathcal{T}$-invariant (order) ideal (a closed order ideal is a closed subspace of the form $\{f \in C(X) \mid f = 0 \text{ on } Y\}$ for some closed set $Y \subset X$). In case $\mathcal{T}$ is given by a continuous semiflow $\{\psi_t\}_{t\geq 0}$ on X, $\{f \in C(X) \mid f = 0 \text{ on } \overline{O_+(x)}\}$ yields a $\mathcal{T}$-invariant closed order ideal for every $x \in X$, where $O_+(x) := \{\psi_t(x) \mid t \geq 0\}$. Thus as stated in the introduction, $\mathcal{T}$ is irreducible if and only if each orbit $O_+(x)$ is dense in X. Although the procedure to get irreducible components of $\mathcal{T}$ runs parallel to that for a single operator ([9]) and Davies[2] contains similar results, we state necessary facts, some without proofs and some with sketches of proofs, to make this paper rather self-contained.

Throughout this section, $\mathcal{T} = (T(t))_{t\geq 0}$ denotes a strongly mean ergodic Markov semigroup on $C(X)$. In the following proposition, we collect some well-known facts or combinations of them, which are easily verified.

Theorem 1.1. *The limit projection P of $\mathcal{T}$ has the following properties.*

(i) *P is a Markov projection of norm 1 : $P^2 = P$, $P \geq 0$ ($Pf \geq 0$ if $f \geq 0$) and $P\mathbf{1} = \mathbf{1}$.*

(ii) *$PT(t) = T(t)P = P$ holds for each $t \geq 0$.*

(iii) *The range $PC(X)$ of P becomes a Banach lattice with respect to the order and the norm induced by those of $C(X)$. The supremum $f\tilde{\vee}g$ of f, $g \in PC(X)$ in $PC(X)$ is given by $P(f \vee g)$, where $f \vee g$ denotes the supremum of f, g in $C(X)$.*

(iv) *Let Λ be the set of the extreme points of the w^*-compact convex set $\Phi := \{\varphi \in C(X)^* \mid \varphi \geq 0, P^*\varphi = \varphi, \varphi(\mathbf{1}) = 1\}$. Then Λ is compact in relative w^*-topology, and $PC(X)$ is isometrically isomorphic to $C(\Lambda)$ via the mapping $f \mapsto i(f)|_\Lambda$ where $i : C(X) \to C(X)^{**}$ is the canonical embedding.*

Proof. (i) and (ii) are well-known. Only note that the positivity of P and $P\mathbf{1} = \mathbf{1}$ imply $\|P\| = 1$.

(iii): See Schaefer[10, Chap.2, Prop. 11.5].

(iv): (i) and (ii) imply that $PC(X)$ is an AM-space with unit $\mathbf{1}$ (Schaefer [10, Chap.2, Def.7.1]). Hence by the famous theorem of Kakutani and Krein (Schaefer [10, Chap.2, Theorem 7.4]) $PC(X)$ is isometrically isomorphic to $C(K)$ for some compact Hausdorff space K. That K is homeomorphic to Λ follows from the fact $(PC(X))^* \cong P^*C(X)^*$ as Banach lattices (isometrically isomorphic via the mapping $\varphi \in P^*C(X)^* \mapsto \varphi|_{PC(X)}$). ■

Remark. Note that in general $f\tilde{\vee}g \geq f \vee g$ for f, $g \in PC(X)$ by definition. And part (iv) of the preceding proposition implies that $\lambda \in \Phi$ belongs to Λ if and only if $\lambda(f\tilde{\vee}g) = \max\{\lambda(f), \lambda(g)\}$ holds for every f, $g \in PC(X)$. We say P is strictly positive if $Pf = 0, f \geq 0$ imply $f = 0$. With this definition it is easy to see that $PC(X)$ is a sublattice of $C(X)$ if P is strictly positive (see Schaefer[10, Chap.2, Theorem 7.4]).

Proposition 1.2. (Sawashima–Niiro [9, Prop.5]) *Let Λ be the set of extreme points of Φ and let $\lambda \in \Lambda$. Let S_λ be the support of λ considered as a regular Borel measure on X. Then for each $f \in C(X)$, $Pf|_{S_\lambda}$ is a constant function with value $\lambda(f)$.*

Proof. If $f \in PC(X)$ and $x \in S_\lambda$, then $P|f| - |f| \geq 0$ and $\lambda(P|f| - |f|) = P^*\lambda(|f|) - \lambda(|f|) = 0$. This implies $P|f| = |f|$ on S_λ, hence the evaluation map $\delta_x : f \mapsto f(x)$ is a lattice homomorphism on $PC(X)$. Therefore $P^*\delta_x \in \Lambda$ by the remark preceding this proposition. That $P^*\delta_x = \lambda$ follows from $PC(X) \cong C(\Lambda)$. ∎

Now we can obtain the reduction into irreducible components. Although the part (i) and (ii) of the following theorem is not new (see [9]), we state them for completeness.

Theorem 1.3. *Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a strongly mean ergodic Markov semigroup on $C(X)$ with limit projection P, let Λ be the set of the extreme points of $\Phi = \{\varphi \in C(X)^* \mid \varphi \geq 0,\ \varphi(1) = 1$ and $P^*\varphi = \varphi\}$ and let S_λ be the support of $\lambda \in \Lambda$. Then the following assertions hold.*

(i) *For all $\lambda \in \Lambda$, S_λ is a compact subset of X and the S_λ's are mutually disjoint.*

(ii) *If P is strictly positive, then $\bigcup_{\lambda\in\Lambda} S_\lambda$ is dense in X.*

(iii) *For all $\lambda \in \Lambda$, the closed order ideal $I_\lambda := \{f \in C(X) \mid f = 0$ on $S_\lambda\}$ is $\mathcal{T}$-invariant and the quotient semigroup ([7, A-I, §3.3]) $\mathcal{T}_\lambda = (T_\lambda(t))_{t\geq 0}$ of $\mathcal{T}$ on $C(X)/I_\lambda$ is an irreducible, strongly mean ergodic Markov semigroup on $C(S_\lambda)$ under the canonical identification of $C(X)/I_\lambda$ with $C(S_\lambda)$.*

Proof. (i):(Sawashima–Niiro[9, Theorem 2,(iii)]) For any λ, $\mu \in \Lambda$ with $\lambda \neq \mu$, there exists an $f \in PC(X)$ such that $f \geq 0, \lambda(f) = 1$ and $\mu(f) = 0$ since $PC(X) \cong C(\Lambda)$. Then $f = 1$ on S_λ and $f = 0$ on S_μ by Proposition 1.2, and hence $S_\lambda \cap S_\mu = \emptyset$.

(ii):(Sawashima–Niiro [9, Prop.6]) This follows from the following equivalences:

$$\begin{aligned} f = 0 \text{ on } \bigcup_{\lambda\in\Lambda} S_\lambda &\iff \lambda(|f|) = 0 \text{ for all } \lambda \in \Lambda \\ &\iff \lambda(P|f|) = 0 \text{ for all } \lambda \in \Lambda \\ &\iff P|f| = 0. \end{aligned}$$

(iii) Let $\lambda \in \Lambda$ and $f \in I_\lambda$. Then $\lambda(T(t)|f|) = \lambda(PT(t)|f|) = \lambda(|f|) = 0$. This means $T(t)|f| \in I_\lambda$, which yields $T(t)f \in I_\lambda$ since the positivity of $T(t)$ implies $|T(t)f| \leq T(t)|f|$. To prove the rest of the theorem, note that if we identify $C(X)/I_\lambda$ with $C(S_\lambda)$, $T_\lambda(t)f$ is given by $T(t)\tilde{f}|_{S_\lambda}$ for $f \in C(S_\lambda)$, where $\tilde{f} \in C(X)$ is an arbitrary extension of f to X. From this view point, the strong mean ergodicity of $\mathcal{T}_\lambda$ is obvious, since $(1/t)\int_0^t T_\lambda(s)f\,ds$ converges uniformly on S_λ to $P\tilde{f}|_{S_\lambda}$. To prove the irreducibility of $\mathcal{T}_\lambda$, assume that $f \in C(S_\lambda), f \geq 0, f \neq 0$ and $\tilde{f}$ is a non-negative continuous extension of f to X. Then for each $x \in S_\lambda, P\tilde{f}(x) = \lambda(\tilde{f}) > 0$ by Proposition 1.2. This shows that any non-zero closed $\mathcal{T}_\lambda$-invariant ideal contains the constant functions and hence is equal to $C(S_\lambda)$, and hence $\mathcal{T}_\lambda$ is irreducible. ∎

Remark. Let $\mathcal{T}$ be as in Theorem 1.3 and set $X_\lambda = \{x \in X \mid P^*\delta_x = \lambda\}$ for $\lambda \in \Lambda$. Then it is clear that X_λ is closed for each $\lambda \in \Lambda$ and mutually disjoint. Proposition 1.2 shows $S_\lambda \subset X_\lambda$, and $X = \bigcup_{\lambda\in\Lambda} X_\lambda$ holds if P is strictly positive, as shown by Sawashima–Niiro [9, Theorem 2]. Moreover, it is easily proved that $J_\lambda := \{f \in C(X) \mid f = 0 \text{ on } X_\lambda\}$ is $\mathcal{T}$-invariant. Thus we obtain another natural reduction of $\mathcal{T}$ if we use $\{X_\lambda\}_{\lambda\in\Lambda}$ instead of $\{S_\lambda\}_{\lambda\in\Lambda}$. But the quotient semigroup on $C(X)/J_\lambda$ unfortunately is not necessarily irreducible, whence we adopt S_λ. See also Davies [2].

From now on we call the family $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$ described in Theorem 1.3 the irreducible components of $\mathcal{T}$, where $\mathcal{T}_\lambda$ is considered as a Markov semigroup on $C(S_\lambda)$, not as an abstract quotient semigroup of $\mathcal{T}$, namely $T_\lambda(t)(f|_{S_\lambda}) = (T(t)f)|_{S_\lambda}$ for $f \in C(X)$. Note here that Theorem 1.3 shows that a strongly mean ergodic Markov semigroup with a strictly positive limit projection is completely determined by its irreducible components.

Next we recall the relation between the generator of $\mathcal{T}$ and those of the $\mathcal{T}_\lambda$'s. To begin with, note that the spectrum of the generator of a Markov semigroup is contained in the left half plane $\{z \mid \operatorname{Re} z \leq 0\}$ since a Markov semigroup is a contraction semigroup. The following proposition is fundamental for our purpose, and assertion (i) is [7, A–III, Prop.4.2] while (ii) is a result of [7, A–I, 3.3] and Theorem 1.3 (ii).

Proposition 1.4. ([7, A]) *Let $\mathcal{T}$ be a strongly mean ergodic Markov semigroup on $C(X)$ and let $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$ be its irreducible components. The generator of $\mathcal{T}$* [resp. $\mathcal{T}_\lambda(\lambda \in \Lambda)$] *is denoted by A* [resp. A_λ] *and $\rho_+(A)$* [resp. $\rho_+(A_\lambda)$] *denotes the connected component of the resolvent set of A* [resp. A_λ] *containing the right half plane $\{z \mid \operatorname{Re} z > 0\}$. Then the following hold.*

(i) *For each $\lambda \in \Lambda$, $\rho_+(A) \subset \rho_+(A_\lambda)$.*

(ii) *For each $\alpha \in \rho_+(A)$, $\|R(\alpha, A)\| \geq \sup_{\lambda\in\Lambda} \|R(\alpha, A_\lambda)\|$. Moreover $\|R(\alpha, A)\| = \sup_{\lambda\in\Lambda} \|R(\alpha, A_\lambda)\|$ holds if the limit projection of $\mathcal{T}$ is strictly positive.* (*Here $R(\alpha, B)$ designates the resolvent $(\alpha - B)^{-1}$.*)

An immediate consequence of Proposition 1.4 is the following

Corollary 1.5. *Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a strongly mean ergodic Markov semigroup on $C(X)$ and let $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$ be its irreducible components. Then the following inclusion holds for the spectra of the generators A of $\mathcal{T}$ and A_λ of $\mathcal{T}_\lambda (\lambda \in \Lambda)$:*

$$\sigma(A) \cap i\mathbb{R} \supset \Big(\bigcup_{\lambda\in\Lambda} \sigma(A_\lambda)\Big)^{-} \cap i\mathbb{R}.$$

In the rest of this section, we give an example of the reduction of a mean ergodic Markov semigroup on the 2-dimensional sphere S^2.

Example 1. Let $X = S^2$ be the 2-dimensional sphere and let $\{\psi_t\}_{t\geq 0}$ be a semiflow on X governed by a smooth differential equation which generates a strongly mean ergodic Markov semigroup $\mathcal{T}$ on $C(X)$. Then the famous Poincaré-Bendixson theorem tells us that S_λ in Theorem 1.3 is either a singleton or homeomorphic to the 1-dimensional sphere (=circle) and $\mathcal{T}_\lambda$ is a rotation semigroup ([7, A-I, 2.5]). Thus the irreducible components are remarkably simple in this case. Such a situation is well illustrated by the following concrete example.

We identify S^2 with the one-point compactification $\mathbb{R}^2 \cup \{\infty\}$ of $\mathbb{R}^2$, and consider the continuous flow $\Psi := \{\psi_t\}_{t\in\mathbf{R}}$ generated by the vector field $F: (x_1, x_2) \mapsto e^{-(x_1^2+x_2^2)}(x_1 - 2x_2(x_1^2 + x_2^2), 2x_1(x_1^2 + x_2^2) - x_2)$ on $\mathbb{R}^2$. Let $f(x) := (x_1^2 + x_2^2)^2 - 2x_1x_2$ for $x = (x_1, x_2) \in \mathbb{R}^2$ and let $\Gamma_c := \{\, x \mid f(x) = c \,\}$ denote the contour of f for $c \in \mathbb{R}$. Noting that $F(x)$ and $\nabla f(x)$ are mutually orthogonal for all $x \in \mathbb{R}^2$, we see that the contours Γ_c are orbits of this flow. More specifically, $\Gamma_c = \emptyset$ for $c < -1/4$ and $\Gamma_{-1/4}$ consists of two fixed points P_1, P_2 of this flow; Γ_c for $-1/4 < c < 0$ is the union of two smooth closed Jordan curves which are periodic orbits of the flow; Γ_0 is the Lemniscate, on which the flow converges to the origin; Γ_c for $c > 0$ is a smooth closed Jordan curve which is a periodic orbit of the flow. Consequently, if $x \in \mathbb{R}^2$ satisfies $f(x) \neq -1/4, 0$, then the orbit $\{\psi_t(x) \mid t \geq 0\,\}$ is a simple closed curve and $t \mapsto \psi_t(x)$ is periodic with period $p(x) > 0$. For the case $f(x) = 0$, we set $p(x) = \infty$.

One can rigorously prove the following intuitively plausible assertions about the contour Γ_c and this flow.

1° Let $-1/4 \leq c_1 < c_2$. Then the length of the contour Γ_c is uniformly bounded for $c \in [c_1, c_2]$.

2° For each $c_1 < c_2$ with $[c_1, c_2] \subset (-1/4, 0)$ or $[c_1, c_2] \subset (0, \infty)$, the period $p(x)$ for $x \in \mathbb{R}^2$ with $f(x) \in [c_1, c_2]$ is uniformly bounded. (This follows from 1° and the fact that the velocity $\left|\frac{d}{dt}\psi_t(x)\right|$ is bounded from below by a positive constant.)

3° Let V be a neighbourhood of the origin. Then for each $\delta_0 \in (0, 1/4)$ there exists a constant $s_0 > 0$ such that

$$\operatorname{meas}\{\, t \in [0, p(x)) \mid \psi_t(x) \notin V\,\} \leq s_0$$

holds for each x with $|f(x)| \leq \delta_0$, where $\operatorname{meas}\{\cdots\}$ denotes the Lebesgue measure on $\mathbb{R}$. Moreover, for each $\varepsilon > 0$ there exists a $\delta_1 \in (0, 1/4)$ for which $p(x) > 1/\varepsilon$ holds for every x with $|f(x)| \leq \delta_1$. (Remember that we have defined $p(x) = \infty$ for x with $f(x) = 0$.)

By using these facts, we can show the strong mean ergodicity of the Markov semigroup $\mathcal{T} := \{T(t)\}_{t\geq 0}$ generated by this flow ($T(t)u := u \circ \psi_t$). Here we sketch a proof of the mean ergodicity of $\mathcal{T}$.

Since $T(t)$ is uniformly bounded, it is sufficient to show that the mean $\frac{1}{t}\int_0^t T(s)u\,ds$ converges for each u in a dense subset of $C(\mathbb{R}^2 \cup \{\infty\})$. Considering the fact $T(t)\mathbf{1} = \mathbf{1}$, we have to prove only the convergence of $\frac{1}{t}\int_0^t T(s)u\,ds$ for $u \in C(\mathbb{R}^2)$ which has compact support K and is constant in a neighbourhood V of the origin.

We take such a u and we may assume $\|u\| \leq 1$. For an arbitrary $\varepsilon > 0$, we show that there exists a $t_0 > 0$ such that t, $t' > t_0$ imply

$$\left\| \frac{1}{t}\int_0^t T(s)u\,ds - \frac{1}{t'}\int_0^{t'} T(s)u\,ds \right\| < \varepsilon.$$

By the fact 3°, there exist constants $s_0 > 0$ and $\delta_0 \in (0, 1/4)$ independent of ε, and a $\delta_1 \in (0, \delta_0)$ which satisfy the following conditions:

(i) $\operatorname{meas}\{ t \in [0, p(x)) \mid \psi_t(x) \notin V \} \leq s_0$ for each x with $|f(x)| \leq \delta_0$;

(ii) $p(x) > 8s_0/\varepsilon$ for each x with $|f(x)| \leq \delta_1$.

Now let x satisfy $|f(x)| \leq \delta_1 (< \delta_0)$ and let $t > 8s_0/\varepsilon$. Then for every $0 < \tau \leq p(x)$

$$\left| \frac{1}{\tau}\int_0^\tau T(s)u(x)\,ds - u(0) \right| \leq \frac{2}{\tau}\operatorname{meas}\{ s \mid \psi_s(x) \notin V,\ 0 \leq s < p(x) \} \leq \frac{2s_0}{\tau}.$$

(Remember that $u \equiv u(0)$ on V and $\|u\| \leq 1$.) So, by writing $t = np(x) + \tau$ for $n \in Z_+$, $0 \leq \tau < p(x)$, we have

$$\begin{aligned}
&\left| \frac{1}{t}\int_0^t T(s)u(x)\,ds - u(0) \right| \\
&= \left| \frac{np(x)}{t}\left(\frac{1}{p(x)}\int_0^{p(x)} T(s)u(x)\,ds - u(0) \right) + \frac{\tau}{t}\left(\frac{1}{\tau}\int_0^\tau T(s)u(x)\,ds - u(0) \right) \right| \\
&\leq \frac{2s_0}{p(x)} + \frac{\tau}{t}\frac{2s_0}{\tau} < \frac{\varepsilon}{2}.
\end{aligned}$$

Therefore, t_1, $t_2 > \frac{8s_0}{\varepsilon}$ implies

$$\sup_{x:|f(x)|\leq\delta_1} \left| \frac{1}{t_1}\int_0^{t_1} T(s)u(x)\,ds - \frac{1}{t_2}\int_0^{t_2} T(s)u(x)\,ds \right| < \varepsilon. \qquad (*)$$

On the other hand, there exists a neighbourhood U_1 [resp. U_2] of P_1 [resp. P_2] such that $x \in U_i$ implies $|u(x) - u(P_i)| < \varepsilon/2$ $(i = 1, 2)$. For these neighbourhoods, there exists a $c_1 \in (-1/4, -\delta_1)$ which satisfies $\{ x \mid f(x) < c_1 \} \subset U_1 \cup U_2$. Also there exists a $c_2 > 0$ which satisfies $\{ x \mid f(x) \leq c_2 \} \supset K$(=the support of u). Then by the fact 3° there exists a positive number p_1 for which $p(x) \leq p_1$ holds for any x with $f(x) \in [c_1, -\delta_1] \cup [\delta_1, c_2]$. Take any $t > 4p_1/\varepsilon$. Then for each x with $f(x) \in [c_1, -\delta_1] \cup [\delta_1, c_2]$, t can be written as $t = np(x) + \tau$ for $n \in Z_+$ and $0 \leq \tau < p(x)$ with $\tau/t < \varepsilon/4$. Hence

$$\begin{aligned}
&\left| \frac{1}{t}\int_0^t T(s)u(x)\,ds - \frac{1}{p(x)}\int_0^{p(x)} T(s)u(x)\,ds \right| \\
&\leq \frac{\tau}{t}\left| \frac{1}{p(x)}\int_0^{p(x)} T(s)u(x)\,ds \right| + \frac{\tau}{t}\left| \frac{1}{\tau}\int_0^\tau T(s)u(x)\,ds \right| \\
&< \frac{\varepsilon}{2}.
\end{aligned}$$

On the other hand, for each x with $f(x) < c_1$ or $f(x) > c_2$,

$$\left|\frac{1}{t}\int_0^t T(s)u(x)\,ds - u(x)\right| \le \frac{1}{t}\int_0^t |T(s)u(x) - u(x)|\,ds < \frac{\varepsilon}{2}$$

holds for any $t > 0$, since the sets $\{\, x \mid f(x) < c_1 \,\}$ and $\{\, x \mid f(x) > c_2 \,\}$ are invariant under the flow. Thus we obtain

$$\sup_{x:|f(x)|\ge\delta_1}\left|\frac{1}{t_1}\int_0^{t_1} T(s)u(x)\,ds - \frac{1}{t_2}\int_0^{t_2} T(s)u(x)\,ds\right| < \varepsilon \qquad (**)$$

for $t_1, t_2 > \frac{4p_1}{\varepsilon}$.

Combining $(*)$ and $(**)$, we obtain

$$t_1,\, t_2 > \max\left\{\frac{8s_0}{\varepsilon}, \frac{4p_1}{\varepsilon}\right\} \Longrightarrow \left\|\frac{1}{t_1}\int_0^{t_1} T(s)u\,ds - \frac{1}{t_2}\int_0^{t_2} T(s)u\,ds\right\| < \varepsilon.$$

Thus we have shown the strong mean ergodicity of $\{T(t)\}_{t\ge 0}$.

Next we describe the irreducible components given by Theorem 1.3. At first we define families of Radon measures $\underline{\lambda}(c)$, $\overline{\lambda}(c)$ for $c \in [-1/4, 0)$ and $\lambda(c)$ for $c \in [0, \infty]$.

As remarked before, the contour $\Gamma_{-1/4}$ consists of two fixed points P_1 and P_2 of the flow Ψ. The Dirac measure $\underline{\lambda}(-1/4)$ [resp. $\overline{\lambda}(-1/4)$] concentrated on P_1 [resp. P_2] is a $\mathcal{T}$-invariant probability measure.

In case $c \in (-1/4, 0)$, the contour Γ_c is the union of two periodic orbits. The one surrounding P_1 is called O_c^1, and the one surrounding P_2 is called O_c^2. For any $x \in O_c^1$ [resp. $x \in O_c^2$], the mapping

$$u \longmapsto \frac{1}{p(x)}\int_0^{p(x)} T(t)u(x)\,dt \quad (u \in C(\mathbb{R}^2 \cup \{\infty\})) \qquad (***)$$

defines the same Radon measure with support O_c^1 [resp. O_c^2], which is denoted by $\underline{\lambda}(c)$ [resp. $\overline{\lambda}(c)$].

Let $\lambda(0)$ denote the Dirac measure concentrated on the origin. If $c \in (0, \infty)$, then the contour Γ_c is a periodic orbit of the flow Ψ, and we denote by $\lambda(c)$ the Radon measure defined by $(***)$ for any $x \in \Gamma_c$. In this case, the support of $\lambda(c)$ is Γ_c. Finally we denote the Dirac measure concentrated on ∞ by $\lambda(\infty)$.

It is clear that $\underline{\lambda}(c)$, $\overline{\lambda}(c)$ $(-1/4 \le \lambda < 0)$ and $\lambda(c)$ $(0 \le \lambda \le \infty)$ are $\mathcal{T}$-invariant probability measures, and it is not difficult to see that these measures exhaust all the extreme points of the $\mathcal{T}$-invariant probability measures on $\mathbb{R}^2 \cup \{\infty\}$. The mappings $c \mapsto \underline{\lambda}(c)$, $c \mapsto \overline{\lambda}(c)$ and $c \mapsto \lambda(c)$ are continuous with respect to the w^*-topology and $\underline{\lambda}(c)$, $\overline{\lambda}(c) \to \lambda(0)$ as $c \uparrow 0$, $\lambda(c) \to \lambda(0)$ as $c \downarrow 0$, $\lambda(c) \to \lambda(\infty)$ as $c \uparrow \infty$. Therefore we can see that the set Λ in Theorem 1.3 is homeomorphic to the planar set in Figure 1.

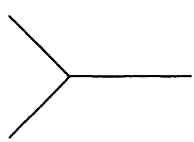
$\Lambda \cong$ 

Figure 1

The support S_λ of $\lambda \in \Lambda$ is described above, and the irreducible component $\mathcal{T}_\lambda$ is the semigroup on $C(S_\lambda)$ induced by the flow restricted to S_λ. Hence it is clear that the irreducible component $\mathcal{T}_\lambda$ is isomorphic to the rotation semigroup on the circle, except for $\lambda = \underline{\lambda}(-1/4)$, $\overline{\lambda}(-1/4)$, $\lambda(0)$ or $\lambda(\infty)$.

2. Boundary spectrum of the generator

Let $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$ be the irreducible components of a strongly mean ergodic Markov semigroup $\mathcal{T}$ on $C(X)$. It would be convenient if we can obtain the information on the properties of $\mathcal{T}$ from that of $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$, since irreducible semigroups are considered simpler than general ones and there indeed exist some precise results about irreducible positive semigroups (see e.g. [7, C-III]). But this prospect is not promising since at present we know nothing about how $\mathcal{T}_\lambda$'s are connected together to form $\mathcal{T}$. To describe the way how $\mathcal{T}_\lambda$'s are connected may be as difficult as to investigate $\mathcal{T}$ directly. Therefore it seems worthwhile to prove that the boundary spectrum of the generator of $\mathcal{T}$ is determined by the spectra of $\mathcal{T}_\lambda$'s, even under some stronger condition.

Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a Markov semigroup on $C(X)$. Then $\mathcal{T}$ is called uniformly mean ergodic if $\mathcal{T}$ satisfies the following condition:

$\frac{1}{t}\int_0^t T(s)\,ds$ converges to an operator P in operator norm as $t \to \infty$.

In this section we prove that the inclusion relation in Corollary 1.5 is replaced by the equality under this stronger condition. Namely, we have the following theorem and its immediate corollary.

Theorem 2.1. *Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a uniformly mean ergodic Markov semigroup on $C(X)$, let $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$ be the irreducible components of $\mathcal{T}$, and let A* [resp. A_λ] *be the generator of $\mathcal{T}$* [resp. $\mathcal{T}_\lambda$]. *Then*

$$\sigma(A) \cap i\mathbb{R} = \left(\bigcup_{\lambda\in\Lambda} \sigma(A_\lambda)\right)^{-} \cap i\mathbb{R}.$$

Corollary 2.2. *Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a uniformly mean ergodic C_0-group of Markov operators on $C(X)$ and let A and A_λ be defined as in Theorem 2.1. Then*

$$\sigma(A) = \left(\bigcup_{\lambda\in\Lambda} \sigma(A_\lambda)\right)^{-}.$$

To prove this theorem, we recall the basic facts about positive pseudo-resolvents (*cf.* Greiner[4], [7,C-III,§2]). Let E be a Banach lattice, D an open set of $\mathbb{C}$ containing the right half plane $\{z \mid \mathrm{Re} z > 0\}$. Then a mapping $R : D \to \mathcal{L}(E)$ is called a positive pseudo-resolvent on E with domain D if

$$R(\alpha_1) - R(\alpha_2) = (\alpha_2 - \alpha_1)R(\alpha_1)R(\alpha_2)$$

holds for all $\alpha_1, \alpha_2 \in D$ and

$$R(\alpha) \geq 0 \quad \text{for all} \quad \alpha > 0.$$

It is known that the above condition implies $|R(\alpha)u| \leq R(\mathrm{Re}\,\alpha)|u|$ for all $u \in E$ and α with $\mathrm{Re}\,\alpha > 0$ ([7,C-III,Prop.2.7]). For later use we quote the following from Greiner[4, Korollar 1.7].

Lemma 2.3. (Greiner[4]) *Let E be a complex Banach lattice and let $u \in E$ be an element for which $|u|$ is a quasi-interior positive element of E* (Schaefer[10, Chap.2, §6]). *Assume that $D = \{z \mid \operatorname{Re} z > 0\}$, $\beta \in \mathbb{R}$ and $R : D \to \mathcal{L}(E)$ is a positive pseudo-resolvent on E. Moreover, let*

$$\alpha R(\alpha + i\beta)u = u \quad \textit{for some (equivalently for all)} \quad \alpha \in D,$$

$$\alpha R(\alpha)|u| = |u| \quad \textit{for some (equivalently for all)} \quad \alpha \in D$$

hold. Then

$$R(\alpha) = M_u{}^{-1} R(\alpha + i\beta) M_u \tag{1}$$

holds for each $\alpha \in D$, where M_u is an invertible isometry defined by u.

Remark. For the definition of M_u, see Greiner [4, p.407] or Nagel (ed.) [7, C-I, §8]. In case $E = C(X)$, M_u is defined by

$$(M_u f)(x) = f(x)u(x)/|u(x)|$$

for $f \in C(X)$ and $x \in X$.

Corollary 2.4. *Besides the condition in Lemma 2.3, assume that the domain of R can be extended to a set containing $\{\alpha \mid 0 < |\alpha| < r\}$ for some $r > 0$ and this extension has a pole of order ℓ at $\alpha = 0$. Then for each $k \in \mathbb{Z}, R$ can be extended uniquely to a positive pseudo-resolvent with domain containing $\{\alpha \mid 0 < |\alpha - ik\beta| < r\}$ and having a pole of order ℓ at $ik\beta$.*

Proof. Iteration of (1) yields

$$R(\alpha) = M_u{}^{-k} R(\alpha + ik\beta) M_u{}^{k} \quad (\operatorname{Re}\alpha > 0)$$

for all $k \in \mathbb{Z}$. Using this equality we can extend the domain of definition of R so as to satisfy the condition of the corollary. The uniqueness is well-known (*cf.* [4,p.405]). ∎

The following is crucial to our theorem.

Lemma 2.5. *Let $\mathcal{T} = (T(t))_{t \geq 0}$ be a bounded semigroup on a Banach space G with generator A. Then $\mathcal{T}$ is uniformly mean ergodic if and only if the resolvent of A has a pole of order at most 1 at 0.*

Proof. Suppose that $\mathcal{T}$ is uniformly mean ergodic with limit projection P. Then P commutes with every $T(t)(t \geq 0)$, hence the subspaces $G_1 := PG$ and $G_2 := (I - P)G$ are both invariant under $\mathcal{T}$. Let $\mathcal{T}_1 = (T_1(t))_{t \geq 0}$[resp.$\mathcal{T}_2 = (T_2(t))_{t \geq 0}$] be the restriction of $\mathcal{T}$ to G_1 [resp. G_2]. Clearly $T_1(t)$ is the identity on G_1 for every $t \geq 0$, and $\mathcal{T}_2$ is uniformly mean ergodic with limit projection 0. Let A_2 be the generator of $\mathcal{T}_2$. Then by the result of Lin [6], $\alpha R(\alpha, A_2)$ converges to 0 as $\alpha \downarrow 0$, and the range of A_2 is closed. Since $\alpha R(\alpha, A_2)u = (\alpha - A_2 + A_2)R(\alpha, A_2)u = u + A_2 R(\alpha, A_2)u$ for each $u \in G_2$ and $\alpha > 0$, $\alpha R(\alpha, A_2) \to 0$ implies that u belongs to the closure of the range of A_2. Therefore the range of A_2 is equal to G_2. On the other hand it is easy to see that A_2 is one-to-one. Thus we obtain $0 \in \rho(A_2)$. This implies that $R(\alpha, A)$ has a pole of order at most 1 since $\alpha^{-1} I \oplus R(\alpha, A_2)$ on $G_1 \oplus G_2$ gives the resolvent of A for $\alpha \neq 0$ with sufficiently small modulus. The converse implication follows immediately from the result of Lin[6]. ∎

We prepare one more fact before we set about the proof of Theorem 2.1.

Lemma 2.6. *Let $\mathcal{T} = (T(t))_{t\geq 0}$ be a uniformly mean ergodic Markov semigroup on $C(X)$ with generator A and let $P = \lim_{t\to\infty}(1/t)\int_0^t T(s)\,ds$. Let S be the closed subset of X for which $I_0 := \{f \mid P|f| = 0\} = \{f \mid f = 0 \text{ on } S\}$. Then I_0 is $\mathcal{T}$-invariant and the quotient semigroup $\mathcal{T}_1 = (T_1(t))_{t\geq 0}$ of $\mathcal{T}$ on $C(X)/I_0$ is a uniformly mean ergodic Markov semigroup on $C(S)$ under the identification of $C(X)/I_0$ with $C(S)$. Moreover, the following assertions hold.*

(i) *There exists an $\eta > 0$ such that $\sigma(A) \cap \{z \mid \operatorname{Re} z > -\eta\} = \sigma(A_1) \cap \{z \mid \operatorname{Re} z > -\eta\}$, where A_1 is the generator of $\mathcal{T}_1$.*

(ii) *The limit projection $P_1 = \lim_{t\to\infty}(1/t)\int_0^t T_1(s)\,ds$ is strictly positive.*

(iii) *$R(\alpha, A_1)$ has a pole of order 1 at 0.*

(iv) *The irreducible components of $\mathcal{T}$ and $\mathcal{T}_1$ are identical and all the components are uniformly mean ergodic.*

Proof. The statement preceding (i) is clear.

proof of (i): By assumption there exists an $\eta > 0$ such that $\{z \mid 0 < |z| < \eta\} \subset \rho(A)$($=$ the resolvent set of A). Let $\mathcal{T}_2 = (T_2(t))_{t\geq 0}$ be the semigroup $\mathcal{T}$ restricted to I_0, and A_2 be its generator. Then $A_2 = A|_{I_0}$ ([7, A-I, 3.2]) and $\mathcal{T}_2$ is a uniformly mean ergodic contraction semigroup of positive operators on the Banach lattice I_0, and hence the resolvent of A_2 has a pole of order at most 1 at 0 by Lemma 2.5. But $\lim_{\alpha\downarrow 0}\alpha R(\alpha, A_2) = P|_{I_0} = 0$ implies $0 \in \rho(A_2)$. Hence $\sigma(A_2) \subset \{z \mid \operatorname{Re} z \leq -\eta\}$ holds by Greiner-Voigt-Wolff[5](or [7, C-III, 1.1]). (i) follows readily from this.

proof of (ii): Let $0 \leq f \in C(S)$ and $\tilde{f} \in C(X)$ be its non-negative extension. Then $P_1 f = P\tilde{f}|_S$ shows that $P_1 f = 0$ implies $\tilde{f} \in I_0$ (*cf.* the proof of Theorem 1.3, (ii)) and hence $f = 0$. This proves (ii).

proof of (iii): This is the consequence of Lemma 2.5.

proof of (iv): Let $\{\mathcal{T}_\lambda\}_{\lambda\in\Lambda}$ be the irreducible components of $\mathcal{T}$ and let $\{S_\lambda\}_{\lambda\in\Lambda}$ be as in Theorem 1.3. Then $S = \overline{\cup_{\lambda\in\Lambda} S_\lambda}$ holds as shown in the proof of Theorem 1.3, (ii). Therefore $\Lambda \subset C(S)^*$ and hence $\Phi = \{\varphi \in C(X)^* \mid \varphi \geq 0, \varphi(\mathbf{1}) = 1, P^*\varphi = \varphi\} \subset C(S)^*$ under the canonical embedding of $C(S)^*$ into $C(X)^*$, whence the irreducible components are identical. That $\mathcal{T}_\lambda$ $(\lambda \in \Lambda)$ is uniformly mean ergodic is clear from its definition. ■

Now we can begin with the

Proof of Theorem 2.1. Because of Corollary 1.5, it suffices to show that the assumption of the existence of an $\alpha_0 \in \sigma(A) \cap i\mathbb{R}$ belonging to $\operatorname{Int}(\cap_{\lambda\in\Lambda}\rho(A_\lambda))$ leads to a contradiction. Hereafter we suppose that α_0 is such an element. The proof is divided into 4 steps.

Step 1.(Reduction to the case where P is strictly positive) Let $\mathcal{T}_1$ be the semigroup defined in Lemma 2.6. Lemma 2.6 shows that if we can prove Theorem 2.1 for $\mathcal{T}_1$, then the conclusion of Theorem 2.1 is valid. Since $\mathcal{T}_1$ has a strictly positive limit projection, we may and do assume hereafter that the limit projection P is strictly positive. We note here that the resolvents of A and A_λ $(\lambda \in \Lambda)$ have poles of order 1 at 0 by Lemma 2.5 and Proposition 1.4.

Step 2. (Selection of a sequence in Λ) By the assumption that $\alpha_0 \in \mathrm{Int}(\bigcap_{\lambda\in\Lambda}\rho(A_\lambda))$ and Lemma 2.5, there exists an $r_0 > 0$ such that $\{\alpha \mid |\alpha - \alpha_0| \leq r_0\} \subset \rho(A_\lambda)$ for all $\lambda \in \Lambda$ and $\{\alpha \mid 0 < |\alpha| < r_0\} \subset \rho_+(A)$. Let r be an arbitrary number such that $0 < r < r_0$. If $\sup\{\|R(\alpha_0, A_\lambda)\| \mid \lambda \in \Lambda\} < \infty$, then we see that $\sup\{\|R(\alpha, A_\lambda)\| \mid \lambda \in \Lambda\}$ is uniformly bounded for α in a neighbourhood of α_0. (Use an estimate of $\|R(\alpha, A_\lambda)\|$ by $\|R(\alpha_0, A_\lambda)\|$ through the Neumann series expansion.) Therefore the boundedness of $\{\|R(\alpha_0, A_\lambda)\| \mid \lambda \in \Lambda\}$ implies the uniform boundedness of $\|R(\alpha, A)\|$ for α sufficiently close to α_0 and $\mathrm{Re}\,\alpha > 0$ (Proposition 1.4), which contradicts $\alpha_0 \in \sigma(A)$ (the Neumann series expansion near α_0 leads to this contradiction). Hence we can choose a sequence $\{\lambda_n\}_n$ in Λ for which

$$\|R(\alpha_0, A_{\lambda_n})\| > n. \tag{2}$$

Since $R(\alpha, A_{\lambda_n})$ is holomorphic on $\{\alpha \mid |\alpha - \alpha_0| \leq r\}$, $\|R(\alpha, A_{\lambda_n})\|$ is subharmonic there, hence

$$\sup\{\,\|R(\alpha, A_{\lambda_n})\| \mid |\alpha - \alpha_0| = r,\ n \in \mathbb{N}\} = \infty \tag{3}$$

follows from (2). As noted at the beginning of this step, the boundedness of $\{\|R(\alpha, A_{\lambda_n})\| \mid n \in \mathbb{N}\}$ implies the uniform boundedness of $\{\|R(\beta, A_{\lambda_n})\| \mid n \in \mathbb{N}\}$ for β in a neighbourhood of α. Hence there exists an α_1 with $|\alpha_1 - \alpha_0| = r$ and $\sup\{\|R(\alpha_1, A_{\lambda_n})\| \mid n \in \mathbb{N}\} = \infty$ by (3). Note that Proposition 1.4 implies $\mathrm{Re}\,\alpha_1 \leq 0$. By choosing a suitable subsequence if necessary, we may assume that

$$\|R(\alpha_1, A_{\lambda_n})\| > n \quad \text{and} \quad \|R(\alpha_0, A_{\lambda_n})\| > n \tag{4}$$

hold simultaneously for each $n \in \mathbb{N}$.

Step 3. In this step we prove that if r $(0 < r < r_0)$ is sufficiently small, then $\alpha_1 \in \mathbb{C}$ and the sequence $\{\lambda_n\}_n$ in Λ chosen in the 2nd step for this r automatically satisfies the following conditions A) and B), where P_λ denotes the limit projection of $\mathcal{T}_\lambda$.

A) There exists a sequence of functions $f_n \in C(S_{\lambda_n})$ with the following properties. (As in Theorem 1.3, S_{λ_n} denotes the support of the measure λ_n.)

$$\begin{cases} f_n \in \mathcal{D}(A_{\lambda_n}),\ \|f_n\| = 1,\ \|(\alpha_0 - A_{\lambda_n})f_n\| < 1/n\,, \\ \liminf_{n\to\infty} \|P_{\lambda_n}|f_n|\| > 0. \end{cases} \tag{5}$$

$$\text{If } \mathrm{Re}\,\alpha > 0, \text{ then } \quad \alpha R(\alpha + \alpha_0, A_{\lambda_n})f_n - f_n \to 0 \quad \text{as} \quad n \to \infty. \tag{6}$$

B) There exists a sequence of functions $g_n \in C(S_{\lambda_n})$ with the following properties:

$$\begin{cases} g_n \in \mathcal{D}(A_{\lambda_n}),\ \|g_n\| = 1,\ \|(\alpha_1 - A_{\lambda_n})g_n\| < 1/n\,, \\ \liminf_{n\to\infty} \|P_{\lambda_n}|g_n|\| > 0. \end{cases} \tag{7}$$

$$\text{If } \mathrm{Re}\,\alpha > r, \text{ then } \quad \alpha R(\alpha + \alpha_1, A_{\lambda_n})g_n - g_n \to 0 \quad \text{as} \quad n \to \infty. \tag{8}$$

To prove this, let us fix a constant ε with $0 < \varepsilon < 1/4$. Then there exists a $t_0 > 0$ for which $\|(1/t_0)\int_0^{t_0} T(s)\,ds - P\| < \varepsilon$. The definitions of $\mathcal{T}_\lambda$ and P_λ $(\lambda \in \Lambda)$ imply

$$\left\|(1/t_0)\int_0^{t_0} T_\lambda(s)\,ds - P_\lambda\right\| < \varepsilon \tag{9}$$

for all $\lambda \in \Lambda$. For this ε, take an r $(0 < r < r_0)$ for which

$$-r \leq \mathrm{Re}\alpha < 0 \Longrightarrow (\exp(t_0 \mathrm{Re}\,\alpha) - 1)/(t_0 \mathrm{Re}\,\alpha) > 1 - \varepsilon \tag{10}$$

holds (this is possible since $(e^{t_0 h} - 1)/t_0 h \to 1$ as $h \to 0-$). Such an r meets the conditions. To see this, first note that (10) implies

$$-r \leq \mathrm{Re}\alpha \leq 0 \Longrightarrow \int_0^{t_0} e^{(\mathrm{Re}\,\alpha)s}\, ds > t_0(1-\varepsilon), \tag{11}$$

admitting the case '$\mathrm{Re}\,\alpha = 0$'. Next let $\{\lambda_n\}_n$ and α_1 with $|\alpha_0 - \alpha_1| = r$ be as constructed in the 2nd step. Then as noted in the 2nd step, $-r \leq \mathrm{Re}\,\alpha_1 \leq 0$. Since $\|R(\alpha_1, A_{\lambda_n})\| > n$, there exists an $h_n \in C(S_{\lambda_n})$ with $\|h_n\| = 1$ and $\|R(\alpha_1, A_{\lambda_n})h_n\| > n$. Setting $g_n := R(\alpha_1, A_{\lambda_n})h_n / \|R(\alpha_1, A_{\lambda_n})h_n\|$, we obtain a $g_n \in C(S_{\lambda_n})$ which satisfies the first three conditions in (7). To see that g_n satisfy the last condition in (7), first note that the equalities

$$\begin{aligned}(T_{\lambda_n}(s) - e^{\alpha_1 s})g_n &= \int_0^s (d/du)(T_{\lambda_n}(u)e^{\alpha_1(s-u)}g_n)\, du \\ &= \int_0^s T_{\lambda_n}(u)e^{\alpha_1(s-u)}(A_{\lambda_n} - \alpha_1)g_n\, du\end{aligned}$$

imply $\|(T_{\lambda_n}(s) - e^{\alpha_1 s})g_n\| \leq (1/n)\int_0^s |e^{\alpha_1(s-u)}|du \leq s/n$ (remember $\mathrm{Re}\alpha_1 \leq 0$), and hence

$$\begin{aligned}\Big\| |T_{\lambda_n}(s)g_n| - e^{\mathrm{Re}\alpha_1 s}|g_n| \Big\| &= \Big\| |T_{\lambda_n}(s)g_n| - |e^{\alpha_1 s}g_n| \Big\| \\ &\leq \|T_{\lambda_n}(s)g_n - e^{\alpha_1 s}g_n\| \leq \frac{s}{n}.\end{aligned}$$

From this we obtain

$$\left\| \int_0^{t_0} |T_{\lambda_n}(s)g_n|\, ds - \int_0^{t_0} e^{\mathrm{Re}\alpha_1 s}|g_n|\, ds \right\| \leq (1/n)\int_0^{t_0} s\, ds \leq t_0{}^2/(2n).$$

Hence we obtain

$$\begin{aligned}\left\| \int_0^{t_0} |T_{\lambda_n}(s)g_n|\, ds \right\| &\geq \left\| \int_0^{t_0} e^{\mathrm{Re}\alpha_1 s}|g_n|\, ds \right\| - t_0{}^2/(2n) \\ &\geq t_0(1 - \varepsilon - (t_0/2n))\end{aligned} \tag{12}$$

since the first term on the right-hand side of (12) is greater than $t_0(1-\varepsilon)$ by (11).

On the other hand (9) and the positivity of $\mathcal{T}_{\lambda_n}$ yield

$$\begin{aligned}\|P_{\lambda_n}|g_n|\| &\geq (1/t_0)\left\| \int_0^{t_0} T_{\lambda_n}(s)|g_n|\, ds \right\| - \varepsilon \\ &\geq (1/t_0)\left\| \int_0^{t_0} |T_{\lambda_n}(s)g_n|\, ds \right\| - \varepsilon,\end{aligned}$$

and hence

$$\|P_{\lambda_n}|g_n|\| \geq 1 - 2\varepsilon - (t_0/2n).$$

Thus we have shown (7). To prove (8), note that $R(\alpha + \alpha_1, A_{\lambda_n})$ is uniformly bounded in n for a fixed α with $\mathrm{Re}\,\alpha > r$ since $\alpha + \alpha_1 \in \rho_+(A)$ for such an α. Then it is easy to see that the equalities

$$\begin{aligned}\alpha R(\alpha + \alpha_1, A_{\lambda_n})g_n &= \alpha R(\alpha + \alpha_1, A_{\lambda_n})R(\alpha_1, A_{\lambda_n})(\alpha_1 - A_{\lambda_n})g_n \\ &= R(\alpha_1, A_{\lambda_n})(\alpha_1 - A_{\lambda_n})g_n - \\ &\quad R(\alpha + \alpha_1, A_{\lambda_n})(\alpha_1 - A_{\lambda_n})g_n \\ &= g_n - R(\alpha + \alpha_1, A_{\lambda_n})(\alpha_1 - A_{\lambda_n})g_n\end{aligned}$$

and (7) imply (8). The similar argument applies to (5) and (6).

Step 4. In this final step we use the ultraproduct method and Greiner's theory of positive pseudo-resolvent. Let $r > 0$ be chosen so as to satisfy the condition (10) in the previous step. Then we obtain an $\alpha_1 \in \mathbb{C}$ with $|\alpha_1 - \alpha_0| = r$, a sequence $\{\lambda_n\}_n$ in Λ and sequences of functions $\{f_n\}_n$ and $\{g_n\}_n$ which satisfy (5) to (8). Now we construct an ultraproduct of Banach lattices $\{C(S_{\lambda_n})\}_n$ (*cf.* [7,A-I,3.6], [10, Chap.5, §1], [9, Prop.10]). Let $\mathcal{F}$ be a free ultrafilter on $\mathbb{N}$. Denote by $\boldsymbol{m}$ the Banach lattice $\left\{\{h_n\}_n \mid h_n \in C(S_{\lambda_n}), \sup_n \|h_n\| < \infty\right\}$, where the norm of $\{h_n\}_n$ is defined by $\sup_n \|h_n\|$, and the order is defined componentwise. Then $p(\{h_n\}_n) := \mathcal{F}\text{-}\lim \|P_{\lambda_n}|h_n|\|$ defines a continuous lattice semi-norm on $\boldsymbol{m}$. The quotient Banach lattice $E := \boldsymbol{m}/p^{-1}(0)$ is what we need. This definition is valid since $p^{-1}(0)$ is a closed order ideal of $\boldsymbol{m}$. Let π denote the quotient mapping $\boldsymbol{m} \to E$. Set $\Delta_0 := \{\alpha \mid \alpha \in \operatorname{Int} \bigcap_{\lambda\in\Lambda} \rho(A_{\lambda_n}) \text{ and } \sup_n \|R(\alpha, A_{\lambda_n})\| < \infty\}$. Then Δ_0 contains the right half plane $H_+ := \{\alpha \mid \operatorname{Re}\alpha > 0\}$ by Lemma 2.5, hence we can define Δ as the connected component of Δ_0 containing H_+. Then for each $\alpha \in \Delta$, a bounded linear operator $R(\alpha)$ on $\boldsymbol{m}$ is defined by $R(\alpha)(\{h_n\}_n) := \{R(\alpha, A_{\lambda_n})h_n\}_n$. It is easy to see that $\{R(\alpha)\}_{\alpha\in\Delta}$ is a positive pseudo-resolvent on $\boldsymbol{m}$ with domain Δ. Moreover, $p^{-1}(0)$ is invariant under $R(\alpha)$ for each $\alpha \in \Delta$. To see this, let $\alpha \in \Delta$ with $\operatorname{Re}\alpha > 0$ and $\{h_n\}_n \in p^{-1}(0)$. Then

$$P_{\lambda_n}|R(\alpha, A_{\lambda_n})h_n| \le P_{\lambda_n}R(\operatorname{Re}\alpha, A_{\lambda_n})|h_n| = P_{\lambda_n}|h_n|/\operatorname{Re}\alpha$$

holds for every $n \in \mathbb{N}$, hence $R(\alpha)\{h_n\}_n \in p^{-1}(0)$. Because of the holomorphy of $R(\alpha)$, $p^{-1}(0)$ is invariant for all $R(\alpha)$ with $\alpha \in \Delta$. Hence a positive pseudo-resolvent $\{R_1(\alpha)\}_{\alpha\in\Delta}$ on E is naturally induced by the following formula:

$$R_1(\alpha)\pi(\{h_n\}_n) := \pi(R(\alpha)\{h_n\}_n).$$

Now the sequences $\{f_n\}_n$ and $\{g_n\}_n$ belong to $\boldsymbol{m}$ and $\pi(\{f_n\}_n)$, $\pi(\{g_n\}_n) \neq 0$ by (5), (7). Moreover $\pi(\{P_{\lambda_n}|f_n|\}_n) = \pi(\{|f_n|\}_n)$ holds. In fact, let δ be an arbitrary positive number. Then there exists a $t_1 > 0$ for which

$$\left\| P_{\lambda_n} - (1/t_1)\int_0^{t_1} T_{\lambda_n}(s)\, ds \right\| < \delta \tag{13}$$

holds for every $n \in \mathbb{N}$. We obtain $\|(T_{\lambda_n}(s) - e^{\alpha_0 s})f_n\| \le s/n$ as in step 3. By using the inequality $T_{\lambda_n}(s)|f_n| \ge |f_n| - |(T_{\lambda_n}(s) - e^{\alpha_0 s})f_n|$, we have

$$(1/t_1)\int_0^{t_1} T_{\lambda_n}(s)|f_n|\, ds \ge |f_n| + h_n$$

for some $h_n \in C(S_{\lambda_n})$ with $\|h_n\| \le t_1/2n$. Together with (13), this implies that for large enough n, $P_{\lambda_n}|f_n| \ge |f_n| + k_n$ for some $k_n \in C(S_{\lambda_n})$ with $\|k_n\| < 2\delta$, hence we get

$$\lim_{n\to\infty} (P_{\lambda_n}|f_n| - |f_n|)^- = 0$$

since $\delta > 0$ is arbitrary. (f^- denotes the negative part of f.) On the other hand, $|P_{\lambda_n}|f_n| - |f_n|| = P_{\lambda_n}|f_n| - |f_n| + 2(P_{\lambda_n}|f_n| - |f_n|)^-$ implies $P_{\lambda_n}|P_{\lambda_n}|f_n| - |f_n|| = 2P_{\lambda_n}(P_{\lambda_n}|f_n| - |f_n|)^-$. Therefore

$$\lim_{n\to\infty} P_{\lambda_n}\Big|P_{\lambda_n}|f_n| - |f_n|\Big| = 0$$

since $\|P_{\lambda_n}\| = 1$. Thus we have shown that $\pi(\{P_{\lambda_n}|f_n|\}_n) = \pi(\{|f_n|\}_n)$. This implies that $\pi(\{|f_n|\}_n)$ is a quasi-interior positive element of E since

$$\begin{aligned}\pi(\{|f_n|\}_n) &= \pi(\{P_{\lambda_n}|f_n|\}_n) = \pi(\{\lambda_n(|f_n|)\mathbf{1}_n\}_n) \\ &= (\mathcal{F}\text{-}\lim \lambda_n(|f_n|))\pi(\{\mathbf{1}_n\}_n)\end{aligned}$$

(*cf.* Proposition 1.2) is non-zero and $\pi(\{\mathbf{1}_n\}_n)$ is clearly the order unit ([10, p.57]) of E, where $\mathbf{1}_n$ denotes the constant function 1 on S_{λ_n}. The equality $\pi(\{P_{\lambda_n}|f_n|\}_n) = \pi(\{|f_n|\}_n)$ also implies that $\alpha R_1(\alpha)\pi(\{|f_n|\}_n) = \pi(\{|f_n|\}_n)$ for α with $\mathrm{Re}\,\alpha > 0$ since

$$\begin{aligned}\alpha R_1(\alpha)\pi(\{|f_n|\}_n) &= \alpha R_1(\alpha)\pi(\{P_{\lambda_n}|f_n|\}_n) \\ &= \pi(\{\alpha R(\alpha, A_{\lambda_n})P_{\lambda_n}|f_n|\}_n) \\ &= \pi(\{|f_n|\}_n).\end{aligned}$$

Finally, we note that $\alpha R_1(\alpha+\alpha_0)\pi(\{f_n\}_n) = \pi(\{f_n\}_n)$ $(\mathrm{Re}\,\alpha > 0)$ and $\alpha R_1(\alpha+\alpha_1)\pi(\{g_n\}_n) = \pi(\{g_n\}_n)$ $(\mathrm{Re}\,\alpha > r)$ hold. These are consequences of (6) and (8), respectively.

Now we obtain a contradiction as follows. By Lemma 2.5 and the assumption of the theorem, R_1 is a positive pseudo-resolvent having a pole of order 1 at 0. Then Corollary 2.4 and what we have proved about $\pi(\{f_n\}_n)$ imply that R_1 can be extended uniquely to a set containing the annulus $\{\alpha \mid 0 < |\alpha-\alpha_0| < r_0\}$ and the half plane $\{\alpha \mid \mathrm{Re}\,\alpha > 0\}$ as a positive pseudo-resolvent having the first order pole at α_0. (Recall that $\{\alpha \mid 0 < |\alpha| < r_0\} \subset \rho_+(A)$. See the beginning of the second step.) Since R_1 is extended as a holomorphic function on the above region, the equation $\pi(\{g_n\}_n) = (\alpha-\alpha_1)R_1(\alpha)\pi(\{g_n\}_n)$ must also hold on the same region. We may take $\alpha = \alpha_1$ in the above equation to obtain $\pi(\{g_n\}_n) = 0$, which contradicts $\pi(\{g_n\}_n) \neq 0$. ∎

Example 2. Let X be the annulus $\{\alpha \in \mathbb{C} \mid r_1 \le |\alpha| \le r_2\}$, where r_1 and r_2 satisfy $0 < r_1 < r_2$. Define a map $\psi_t : X \to X$ for every $t \in \mathbb{R}$ by setting

$$\psi_t(re^{i\theta}) := r\exp\{i(\theta + tr)\}$$

for $re^{i\theta} \in X$ $(r, \theta \in \mathbb{R})$. Then $\Psi = \{\psi_t\}_{t\in\mathbf{R}}$ is a continuous flow on X and $T(t)f := f \circ \psi_t$ $(f \in C(X))$ defines a C_0-group $(T(t))_{t\in\mathbf{R}}$ of Markov operators on $C(X)$. It is not difficult to see that $\mathcal{T} := (T(t))_{t\ge 0}$ is uniformly mean ergodic, and the limit projection P of $\mathcal{T}$ is given by

$$Pf(re^{i\theta}) = \frac{1}{2\pi}\int_0^{2\pi} f(re^{i\varphi})\, d\varphi.$$

Under the identification of $g \in C[r_1, r_2]$ with $f \in C(X)$ given by $f(re^{i\theta}) := g(r)$, $PC(X)$ is isometrically isomorphic to $C[r_1, r_2]$ as a Banach lattice. This implies that $\Lambda \cong [r_1, r_2]$ and $S_\lambda = \{\alpha \in X \mid |\alpha| = \lambda\}$ where $\lambda \in [r_1, r_2]$ is considered as a Radon measure $f \mapsto 1/2\pi \int_0^{2\pi} f(\lambda e^{i\varphi})\, d\varphi$. Hence the irreducible components $\mathcal{T}_\lambda$ $(\lambda \in \Lambda)$ is the rotation group on the circle S_λ of period $2\pi/\lambda$. Therefore Corollary 2.2 implies that the spectrum of the generator A of $\mathcal{T}$ is given by

$$\sigma(A) = \bigcup_{k\in\mathbb{Z}_+} \{i\beta \mid \beta \in \mathbb{R}, kr_1 \le |\beta| \le kr_2\}.$$

Acknoledgement. The author would like to express his sincere thanks to the referee, who pointed out a flaw in the first version (the third step of the proof of Theorem 2.1).

References

[1] Davies, E.B., "One-parameter semigroups," Academic Press, London, 1980.

[2] Davies, E.B., *The harmonic functions of mean ergodic Markov semigroups*, Math. Z. **181** (1982), 543–552.

[3] Gantmacher, F.R., "The theory of matrices," vol. 2 (English translation), Chelsea, New York, 1974.

[4] Greiner, G., *Zur Perron-Frobenius-Theorie stark stetiger Halbgruppen*, Mathematische Zeitschrift **177** (1981), 401–423.

[5] Greiner, G., J. Voigt and M. Wolff, *On the spectral bound of the generator of semigroups of positive operators*, J. Operator Theory **5** (1981), 245–256.

[6] Lin, M.,*On the uniform ergodic theorem. II*, Proc. Amer. Math. Soc. **46** (1974), 217–225.

[7] Nagel, R.(ed.), "One-parameter semigroups of positive operators," Lecture Notes in Math. no.1184, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1986.

[8] Niiro, F. and I. Sawashima, *On the spectral properties of positive irreducible operators in an arbitrary Banach lattice and problems of H. H. Schaefer*, Sci. Papers Coll. Gen. Educ. Univ. Tokyo **16** (1966), 145–183.

[9] Sawashima, I. and F. Niiro, *Reduction of a Sub-Markov operator to its irreducible components*, Natur. Sci. Rep. Ochanomizu Univ. **24** (1973), 35–59.

[10] Schaefer, H. H., "Banach lattices and positive operators," Springer-Verlag, Berlin-Heidelberg-New York, 1974.

Department of Mathematics
Faculty of Science
Science University of Tokyo
Wakamiya-cho 26, Shinjuku-ku
Tokyo 162 Japan

electronic address
SS0849@JPNSUT20.BITNETJP

Received March 28, 1991
and in final form December 12, 1991

Semigroup Forum Vol. 45 (1992) 241–248

RESEARCH ARTICLE

Moment problems on subsemigroups of $\mathbb{N}_0^k$ and $\mathbb{Z}^k$

Nobuhisa Sakakibara

Communicated by R. Nagel

Introduction

Classical moment problems have been treated by many mathematicians. In particular, there are comprehensive studies of the trigonometric moment problem and the power moment problem. A unifying way of these problems is the moment problem on abelian $*$-semigroups. The power moment problem corresponds to the one on the semigroup $\mathbb{N}_0 = (\mathbb{N}_0, +)$ of nonnegative integers with the identical involution $n^* = n$, while the trigonometric moment problem does to the one on the group $(\mathbb{Z}, +, n^* = -n)$ of integers with the group involution $n^* = -n$. In the moment problem on abelian $*$-semigroups, the notions of perfectness and semiperfectness are quite important. To characterize perfectness or semiperfectness of an abelian $*$-semigroup is a challenging, but not yet generally solved problem. In the present paper, improving the results of [8], we will completely characterize perfect or semiperfect subsemigroups of $\mathbb{N}_0^k$ and $\mathbb{Z}^k$, $k \geq 1$.

1. Preliminaries

Let S be an abelian $*$-semigroup with the identity 0. A complex-valued function ρ on S is called a *semicharacter* if it is a nonzero multiplicative function satisfying $\rho(s^*) = \overline{\rho(s)}$ for all $s \in S$. The set of all semicharacters is denoted by S^*. We equip S^* with the topology of pointwise convergence. Then S^* is a topological semigroup under pointwise multiplication with involution $\rho \mapsto \overline{\rho}$ and the identity 1. We call S^* the *dual semigroup* of S. When $S = \mathbb{N}_0$, every element of $\mathbb{N}_0^*$ is of the form $\rho_x(n) = x^n$ for $n \in \mathbb{N}_0$, $x \in \mathbb{R}$, with the convention that $\rho_0 = \mathbf{1}_{\{0\}}$, the indicator function of $\{0\}$. The mapping $x \mapsto \rho_x$ is a topological semigroup isomorphism of $(\mathbb{R}, \cdot)$ onto $\mathbb{N}_0^*$, where the composition $\cdot$ is multiplication. The function $\varphi : S \to \mathbb{C}$ is called a *moment function* if there exists a Radon measure μ on S^* such that

$$\int_{S^*} |\rho(s)|\, d\mu(\rho) < \infty \qquad \text{for} \quad s \in S,$$

$$\varphi(s) = \int_{S^*} \rho(s) d\mu(\rho) \qquad \text{for} \quad s \in S.$$

It can be easily seen that every moment function φ is *positive definite*, that is, for any $s_1, s_2, \cdots, s_n \in S$ and $c_1, c_2, \cdots, c_n \in \mathbb{C}$

$$\sum_{i,j=1}^{n} c_i \overline{c_j} \varphi(s_i + s_j{}^*) \geq 0.$$

A representing measure μ for φ is not necessarily unique if any. Also a positive definite function on S is not necessarily a moment function. An abelian $*$-semigroup S is called *semiperfect* if every positive definite function on S is a moment function. If, furthermore, a representing measure is unique, then S is called *perfect*. The following are examples of perfect semigroups: the trivial semigroup $\{0\}$, abelian groups with the group involution $s^* = -s$ (by Bochner's Theorem), the semigroup $\mathbb{Q}_+ = (\mathbb{Q}_+, +)$ of nonnegative rational numbers with the identical involution (see [2], Proposition 6.5.6) and the semigroup $\mathbb{Q} = (\mathbb{Q}, +)$ with the identical involution (see [2], Proposition 6.5.10). The semigroups $\mathbb{N}_0$ and $\mathbb{Z} = (\mathbb{Z}, +)$ with the identical involution are semiperfect but not perfect (cf.[7],[9]). On the other hand, for $k \geq 2$ the direct product semigroups $\mathbb{N}_0^k = (\mathbb{N}_0^k, +)$ and $\mathbb{Z}^k = (\mathbb{Z}^k, +)$ with the identical involution are not semiperfect (see [1], Theorem 4 and [2], Exercise 6.4.8). The problem to determine the moment functions on $\mathbb{N}_0$ (resp. $\mathbb{Z}$, $\mathbb{N}_0^k$ $(k \geq 2)$) is called the *classical Hamburger moment problem* (resp. the *strong Hamburger moment problem*, the *classical multidimensional power moment problem*). Perfect or semiperfect semigroups have some nice properties:

(1) The direct product of two perfect semigroups is perfect (see [2], Theorem 6.5.4).
(2) Any *-homomorphic image of a perfect semigroup is perfect (see [2], Theorem 6.5.5).
(3) Any *-homomorphic image of a semiperfect semigroup is semiperfect (cf. [5], Proposition 1).
(4) The direct product of a perfect semigroup and a finitely generated semiperfect semigroup is semiperfect (see [4], Proposition 1).
(5) Any *-subsemigroup with the ideal property of a perfect *-semigroup is perfect (see [8], Theorem). Here a *-subsemigroup T of a *-semigroup S is said to have the *ideal property* if

$$t + S := \{t + s \mid s \in S\} \subset T \qquad \text{for all} \quad t \in T \setminus \{0\}.$$

See [2] for further details on related results.

2. Auxiliary results

In this section, we will present some results on semiperfectness, which will be used in proving our main results, Theorem 3.1 and Theorem 4.1. Let S be an abelian *-semigroup. A *-subsemigroup X of S is called a *face* if $s, t \in S$ and $s + t \in X$ imply $s, t \in X$.

Theorem 2.1. *A face X of a semiperfect *-semigroup S is semiperfect. Furthermore, S is perfect if and only if both X and $Y := (S \setminus X) \cup \{0\}$ are perfect.*

Proof. Remark that Y becomes a *-subsemigroup of S and has the ideal property. Let φ be a positive definite function on X. Then the function $\psi : S \to \mathbb{C}$ defined by

$$\psi(s) := \begin{cases} \varphi(s) & (s \in X) \\ 0 & (s \in S \setminus X) \end{cases}$$

is positive definite on S. Since S is semiperfect, there exists a Radon measure μ on S^* such that

$$\int_{S^*} |\rho(s)| \, d\mu(\rho) < \infty \qquad \text{for} \quad s \in S,$$

$$\psi(s) = \int_{S^*} \rho(s) d\mu(\rho) \qquad \text{for} \quad s \in S.$$

Let $f : S^* \to X^*$ be the function defined by $f(\rho) := \rho|_X$, $\rho \in S^*$. Defining a Radon measure ν on X^* such that $\nu = \mu^f$, we have

$$\int_{X^*} |\rho(s)|\, d\nu(\rho) < \infty \qquad \text{for} \quad s \in X,$$

$$\varphi(s) = \int_{X^*} \rho(s) d\nu(\rho) \qquad \text{for} \quad s \in X.$$

Hence X is semiperfect. If, furthermore, S is perfect, then uniqueness of the representing measure ν is obvious, making use of the canonical map $X^* \longrightarrow S^*$ which sends any $\rho \in X^*$ to its zero extension on S. Since Y has the ideal property, then Y is perfect by assertion (5). Conversely, if X and Y are perfect, then $X \times Y$ is perfect by assertion (1). The image of the *-homomorphism $(s,t) \mapsto s+t$ from $X \times Y$ to S, which coincides with S, is perfect by assertion (2). Hence S is perfect. ∎

Remark 2.2. Let S be a subsemigroup of $\mathbb{Q}_+^2$. For every $p, q \in \mathbb{Z}$,

$$h_{p,q}(x,y) := px + qy \qquad ((x,y) \in S)$$

is a homomorphism from S into $\mathbb{Q}$. Defining a subsemigroup $X_{p,q}$ of S by

$$X_{p,q} := \{(x,y) \in S \mid h_{p,q}(x,y) = 0\}$$

for $p, q \in \mathbb{Z}$, $(p,q) \neq (0,0)$, we have that $X_{1,0}$ and $X_{0,1}$ are faces of S. Hence if S is semiperfect then $X_{1,0}$ and $X_{0,1}$ are semiperfect by Theorem 2.1. It is natural to ask whether every $X_{p,q}$ is semiperfect under the assumption that S is semiperfect. Unfortunately the answer is no. For example, $S := \mathbb{N}_0 \times (\mathbb{Q}_+ \setminus (0,1))$ is semiperfect by assertions (5) and (4), but the subsemigroup $X_{1,-2}$ of S is not semiperfect by Corollary 4.2.

The following Corollary 2.3 can be easily obtained from Theorem 2.1. This is an improvement of [8, Corollary 3].

Corollary 2.3. *Let S be an abelian *-semigroup. Define *-subsemigroups*

$$\Gamma := \{s \in S \mid s + t = 0 \text{ for some } t \in S\} \qquad \text{and} \qquad T := (S \setminus \Gamma) \cup \{0\}.$$

If S is semiperfect, then Γ is semiperfect. Furthermore, S is perfect if and only if both Γ and T are perfect.

For $s \in S$, let χ_s denote the function of S^* into $\mathbb{C}$ given by

$$\chi_s(\rho) = \rho(s) \qquad \text{for} \quad \rho \in S^*.$$

Let $V(\mathbb{C})$ be the complex vector space spanned by the functions χ_s, $s \in S$, i.e.

$$V(\mathbb{C}) := \{\sum_{j=1}^{n} a_j \chi_{s_j} \mid n \geq 1, a_j \in \mathbb{C}, s_j \in S, j = 1, \cdots, n\}.$$

We equip $V(\mathbb{C})$ with the finest locally convex topology. Define $V(\mathbb{R})$, $V(\mathbb{R})_+$ and Σ as follows:

$$V(\mathbb{R}) := \{f \in V(\mathbb{C}) \mid f \text{ is real-valued}\};$$

$$V(\mathbb{R})_+ := \{f \in V(\mathbb{R}) \mid f \text{ is nonnegative}\};$$

$$\Sigma := \{|f_1|^2 + \cdots + |f_n|^2 \mid n \geq 1, f_j \in V(\mathbb{C}), j = 1, \cdots, n\}.$$

For a semigroup S with the identical involution, it is sufficient to consider the real vector space $V(\mathbb{R})$ spanned by the real functions χ_s, $s \in S$. We cite the following theorem as a tool to exclude semiperfectness. In [2, Theorem 6.1.11], Berg et al. have proved that the condition in the theorem is necessary and sufficient for semiperfectness under the assumption that S is finitely generated. But for only necessity the assumption is not necessary.

Theorem 2.4. *Let S be an abelian *-semigroup for which S^* separates the points. If S is semiperfect, then the closure of Σ equals to $V(\mathbb{R})_+$.*

3. Moment problem on subsemigroups of $\mathbb{N}_0^k$

Assertion (5) in §1 is not true if perfectness is replaced by semiperfectness. For example, though $\mathbb{N}_0$ is semiperfect, the subsemigroup $\mathbb{N}_0 \setminus \{1\}$ with the ideal property is not semiperfect (see [8], §4). This is a little surprising, and it is natural to ask which subsemigroups of $\mathbb{N}_0$ are semiperfect. In this section we will give a characterization of semiperfect subsemigroups of $\mathbb{N}_0^k$, $k \geq 1$.

Theorem 3.1. *A semiperfect subsemigroup S of $\mathbb{N}_0^k$ with the identical involution is either trivial or singly generated. Furthermore, the only perfect subsemigroup is trivial.*

Proof. (I) Firstly let us prove Theorem 3.1 for $k = 1$. It suffices to prove that S is not semiperfect, when S is neither trivial nor singly generated. If S dose not contain any odd numbers, then there exist a subsemigroup T of $\mathbb{N}_0$ and $k \in \mathbb{N}$ such that $S = 2^k T$ and $T \cap (2\mathbb{N}+1) \neq \emptyset$. In this case, the semiperfectness of S is equivalent to that of T by an isomorphic identification. Hence we may assume that S contains an odd number $m \geq 3$. For $x \in \mathbb{R}$ the function $\rho_x : S \to \mathbb{R}$ defined by $\rho_x(n) = x^n$ is a semicharacter on S. Conversely let $\rho \in S^*$. Defining $x = \rho(m)^{\frac{1}{m}}$ if $\rho(m) \geq 0$ and $x = -(-\rho(m))^{\frac{1}{m}}$ if $\rho(m) < 0$, we can easily show $\rho = \rho_x$ because m is odd. It can be easily checked that $x \mapsto \rho_x$ is a topological semigroup isomorphism of $(\mathbb{R}, \cdot)$ onto S^*. We note that S^* separates the points of S. We have that $V(\mathbb{R})$ is the real vector space spanned by the monomials x^n, $n \in S$, and that Σ is the convex cone generated by the squares of elements of $V(\mathbb{R})$. Modifying the proof of [2, Theorem 6.3.2], we can prove that Σ is closed in $V(\mathbb{R})$. Then it suffices to prove that $V(\mathbb{R})_+ \setminus \Sigma \neq \emptyset$ by Theorem 2.4. We shall prove that there exists an even number n_0 in $S \setminus 2S$. Let n_1 be the smallest nonzero number in S. In case that n_1 is even, n_1 suffices for n_0. In case that n_1 is odd, let n_2 be the smallest nonzero even number in S such that n_2 is not divided by n_1. Such n_2 exists because S is not singly generated. Suppose $n_2 \in 2S$. Then there exists $n_3 \in S$ such that $n_2 = 2n_3$. Since $n_1 < n_3 < n_2$ and that n_3 is not divided by n_1, n_3 must be odd. Then $n_1 + n_3$ is an even number in S such that $n_1 + n_3$ is not divided by n_1 and such that $n_1 + n_3 < n_2$. This contradicts the choice of n_2. Then n_2 suffices for n_0. We can easily check that $x^{n_0} \in V(\mathbb{R})_+ \setminus \Sigma$. (II) Secondly let us prove Theorem 3.1 for $k = 2$. Let S be a subsemigroup of $\mathbb{N}_0^2$. It suffices to prove that S is singly generated when S is semiperfect and not trivial. For every $p, q \in \mathbb{N}_0$, $h_{p,q}(m_1, m_2) := pm_1 + qm_2$ is a homomorphism from S into $\mathbb{N}_0$. For every $p, q \in \mathbb{N}_0$, $h_{p,q}(S)$ is semiperfect by (3), so that $h_{p,q}(S)$ is trivial or singly generated by (I). In particular, $h_{1,0}(S)$ and $h_{0,1}(S)$ satisfy this condition. When either $h_{1,0}(S)$ or $h_{0,1}(S)$ is trivial, we can identify S with a subsemigroup of $\mathbb{N}_0$. Then S is singly generated by (I). Therefore it suffices to consider only the case when $h_{1,0}(S)$ and $h_{0,1}(S)$ are singly generated. Furthermore we can assume that $h_{1,0}(S) = h_{0,1}(S) = \mathbb{N}_0$ by an isomorphic identification. Suppose that S contains $(0, m_2)$ for some nonzero m_2. Let n be the smallest nonzero number such that $(0, n) \in S$. Then $h_{n,1}(S)$ is singly generated and $\min(h_{n,1}(S) \setminus \{0\}) = n$. Hence $h_{n,1}(S)$ is generated by n. For every $(m_1, m_2) \in S$, we have $nm_1 + m_2 \in n\mathbb{N}_0$, so that $m_2 \in n\mathbb{N}_0$. But $h_{0,1}(S) = \mathbb{N}_0$. Hence n must be 1. By the similar argument, if any, let m be the smallest nonzero number such that $(m, 0) \in S$. Then m must be 1. When $(1, 0)$ and $(0, 1)$ belong to S, S must be $\mathbb{N}_0^2$. This contradicts that S is semiperfect. So we have that $(1, 0) \notin S$ (i.e. $(n, 0) \notin S$ for any $n \in \mathbb{N}$) or $(0, 1) \notin S$ (i.e. $(0, n) \notin S$ for any $n \in \mathbb{N}$). In case that $(1, 0) \in S$ and $(0, 1) \notin S$, $h_{p,1}(S)$

is generated by p for every $p \in \mathbb{N}$. This implies that if $(m_1, m_2) \in S$ then $m_2 \in p\mathbb{N}_0$ for any $p \in \mathbb{N}$. Only $m_2 = 0$ satisfies this condition. This contradicts that $h_{0,1}(S) = \mathbb{N}_0$. Then it suffices to consider only when neither $(1,0)$ nor $(0,1)$ belongs to S, that is, there exist no points on both axes except for $(0,0)$. Since $h_{1,0}(S) = \mathbb{N}_0$, S contains $(1, m_2)$ for some nonzero m_2. Let n be the smallest nonzero number such that $(1,n) \in S$. Since $h_{n,1}(S)$ is generated by $2n$, every m_2 such that $(m_1, m_2) \in S$ belongs to $n\mathbb{N}_0$. Since $h_{0,1}(S) = \mathbb{N}_0$, n must be 1, that is, $(1,1) \in S$. This implies that if $(m_1, m_2) \in S$ then $pm_1 + qm_2 \in (p+q)\mathbb{N}_0$ for any $p, q \in \mathbb{N}$. This is possible only when $m_1 = m_2$. Hence $S = \{(n,n) \mid n \in \mathbb{N}_0\}$, so that S is singly generated. (III) In general case we shall prove Theorem 3.1 by induction. Assume that Theorem 3.1 is verified for $k = n$ $(n \geq 2)$. Let S be a subsemigroup of $\mathbb{N}_0^{n+1}$. We may assume that every image of S by

$$(m_1, m_2, \cdots, m_{n+1}) \mapsto m_i, \qquad 1 \leq i \leq n+1,$$

is $\mathbb{N}_0$. Defining a homomorphism $h : S \to \mathbb{N}_0^n$ by

$$h(m_1, m_2, \cdots, m_n, m_{n+1}) := (m_1, m_2, \cdots, m_n),$$

we have that $h(S)$ is semiperfect. Then

$$h(S) = \{(m, m, \cdots, m) \in \mathbb{N}_0^n \mid m \in \mathbb{N}_0\}.$$

That is, every element of S is of the form $(m, m, \cdots, m, m_{n+1})$ for $m \in \mathbb{N}_0$ and some $m_{n+1} \in \mathbb{N}_0$. Let $\pi : S \to \mathbb{N}_0^2$ be the homomorphism defined by

$$\pi(m, m, \cdots, m, m_{n+1}) := (m, m_{n+1}).$$

Since $\pi(S)$ is semiperfect, then

$$\pi(S) = \{(m,m) \in \mathbb{N}_0^2 \mid m \in \mathbb{N}_0\}.$$

That is, every element of S is of the form $(m, m, \cdots, m)$ for $m \in \mathbb{N}_0$. Hence S is singly generated. This completes the proof. ∎

The following is an immediate consequence of Theorem 3.1.

Corollary 3.2. *For $k \geq 1$, let S be a finitely generated subsemigroup of $\mathbb{Q}_+^k$. Then S is semiperfect if and only if S is trivial or singly generated. Furthermore, S is perfect if and only if S is trivial.*

Proof. We assume that S is semiperfect and not trivial. Let $\{r_1, r_2, \cdots, r_n\}$ be a generator set of S. Then there exists $m \in \mathbb{N}$ such that $m\{r_1, r_2, \cdots, r_n\}$ is a subset of $\mathbb{N}_0^k$, i.e. $mS \subset \mathbb{N}_0^k$. By Theorem 3.1, mS is singly generated. Hence S is singly generated. The converse is clear. ∎

Remark 3.3. In general, it is not necessary that subsemigroups of $\mathbb{Q}_+^k$ are finitely generated. For example, $(\mathbb{N}_0^2 + (1,1)) \cup \{(0,0)\}$ is not finitely generated. As to subsemigroups of $\mathbb{N}_0$, they must be finitely generated. Any discrete nontrivial subsemigroups of $\mathbb{Q}_+$ are not perfect. Bisgaard [3] pointed out this result by modifying the method of Boas [6]. For discrete nontrivial subsemigroups of $\mathbb{Q}_+^k$, $k \geq 2$, it is true by assertion (2), too.

4. Moment problem on subsemigroups of $\mathbb{Z}^k$

In this section we shall give a characterization of semiperfect subsemigroups of $\mathbb{Z}^k$, $k \geq 1$. Here we use the word "singly generated" in a way different from that in the previous section. A nontrivial subsemigroup S of $\mathbb{Z}^k$ (resp. $\mathbb{Q}^k$) is called *singly generated* when $S = \mathbf{n}\mathbb{N}_0 := \{m\mathbf{n} \mid m \in \mathbb{N}_0\}$ or $S = \mathbf{n}\mathbb{Z}$ for some $\mathbf{n} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$ (resp. $\mathbb{Q}^k \setminus \{\mathbf{0}\}$). Then a singly generated subsemigroup is isomorphic to either $\mathbb{N}_0$ or $\mathbb{Z}$, hence semiperfect. The following theorem gives a converse.

Theorem 4.1. *Every nontrivial semiperfect subsemigroup of $\mathbb{Z}^k$ with the identical involution is singly generated. It is never perfect.*

Proof. Let S be a nontrivial semiperfect subsemigroup of $\mathbb{Z}^k$. The assertion is clear for $k = 1$. In fact, in case $S \subset \mathbb{N}_0$ or $\subset -\mathbb{N}_0$, this follows from Theorem 3.1, while in case $S \cap \mathbb{N} \neq \emptyset$ and $S \cap (-\mathbb{N}) \neq \emptyset$, S is isomorphic to $\mathbb{Z}$. Let us prove the assertion for $k = 2$. The abelian subgroup $\Gamma := \{s \in S \mid s+t = \mathbf{0} \text{ for some } t \in S\}$ of $\mathbb{Z}^2$ is semiperfect by Corollary 2.3. Since every subgroup of $\mathbb{Z}^2$ is isomorphic to $\{0\}$, $\mathbb{Z}$ or $\mathbb{Z}^2$, and $\mathbb{Z}^2$ is not semiperfect, Γ is isomorphic to $\{0\}$ or $\mathbb{Z}$. Hence S is contained in some half-plane of $\mathbb{R}^2$. We can assume that there exists a nonnegative number α such that

$$S \subset \{(m_1, m_2) \in \mathbb{Z}^2 \mid m_1 + \alpha m_2 \geq 0\}. \qquad H$$

If S is contained in a sector of $\mathbb{R}^2$ with angle $< \pi$, it is isomorphic to a semiperfect subsemigroup of $\mathbb{N}_0^2$, hence is singly generated by Theorem 3.1. Thus we have to show that S must be non-semiperfect when S is not contained in any sectors with angle $< \pi$ and any lines through the origin $(0,0)$ of $\mathbb{R}^2$. Under the above assumptions S satisfies one of the following conditions:

(i) The nonnegative number α in (H) is irrational;
(ii) α is rational and $(S \setminus \{(0,0)\}) \cap \{(m_1, m_2) \in \mathbb{Z}^2 \mid m_1 + \alpha m_2 = 0\} = \emptyset$;
(iii) α is rational and $(S \setminus \{(0,0)\}) \cap \{(m_1, m_2) \in \mathbb{Z}^2 \mid m_1 + \alpha m_2 = 0\} \neq \emptyset$.

Let $h_{1,0}$ and $h_{0,1}$ be homomorphisms on S defined by

$$h_{1,0}(m_1, m_2) := m_1 \qquad \text{and} \qquad h_{0,1}(m_1, m_2) := m_2$$

for $(m_1, m_2) \in S$. In case (i), we can assume that $h_{1,0}(S) = h_{0,1}(S) = \mathbb{Z}$, $(1,0) \in S$ and $(0,1) \in S$. In fact, as in the proof (II) of Theorem 3.1, we can assume that $h_{1,0}(S) = \mathbb{Z}$. Then there exists $n \in \mathbb{Z}$ with $(1,n) \in S$. Let π_1 be the homomorphism on S defined by

$$\pi_1(m_1, m_2) := (m_1, m_2 - nm_1).$$

Then $\pi_1(1,n) = (1,0)$ and $h_{1,0}(\pi_1(S)) = \mathbb{Z}$. As above, we can assume that $h_{0,1}(\pi_1(S)) = \mathbb{Z}$. Then there exists $m \in \mathbb{Z}$ with $(m,1) \in \pi_1(S)$. Let π_2 be the homomorphism on $\pi_1(S)$ defined by

$$\pi_2(m_1, m_2) := (m_1 - mm_2, m_2).$$

We have that $(1,0), (0,1) \in \pi_2(\pi_1(S))$ and $h_{1,0}(\pi_2(\pi_1(S))) = h_{0,1}(\pi_2(\pi_1(S))) = \mathbb{Z}$. Since $\pi_2(\pi_1(S))$ is isomorphic to S, we identify $\pi_2(\pi_1(S))$ with S, and we can assume that $h_{1,0}(S) = h_{0,1}(S) = \mathbb{Z}$, $(1,0) \in S$ and $(0,1) \in S$. Under this identification, it remains that S is contained in a half-plane of $\mathbb{R}^2$ and is not contained in any sectors with angle $< \pi$ and any lines through the origin $(0,0)$ of $\mathbb{R}^2$. Here, if α satisfying the condition (H) changes to a different number, it remains irrational. We can determine the dual semigroup S^* exactly as follows:

$$S^* = \{(m_1, m_2) \mapsto x^{m_1} y^{m_2} \mid x, y \in \mathbb{R} \setminus \{0\}\} \cup \{\mathbf{1}_{\{(0,0)\}}\}.$$

In fact, for $x, y \in \mathbb{R}\setminus\{0\}$, the function $\rho_{x,y} : S \to \mathbb{R}$ defined by $\rho_{x,y}(m_1, m_2) = x^{m_1} y^{m_2}$ and $\mathbf{1}_{\{(0,0)\}}$ are semicharacters on S. Conversely let $\rho \in S^* \setminus \{\mathbf{1}_{\{(0,0)\}}\}$. Since S is not contained in any sectors of $\mathbb{R}^2$ with angle $< \pi$ and $\rho \neq \mathbf{1}_{\{(0,0)\}}$, we have that $x \neq 0$ and $y \neq 0$, where $x = \rho(1,0)$ and $y = \rho(0,1)$. Then we can easily show $\rho = \rho_{x,y}$. Note that S^* is isomorphic to $((\mathbb{R}\setminus\{0\})^2 \cup \{(0,0)\}, \cdot\,)$ and separates the points of S. In this case, $V(\mathbb{R})$ is the real vector space spanned by the monomials $x^{m_1}y^{m_2}$, $(m_1, m_2) \in S$, on $(\mathbb{R}\setminus\{0\})^2 \cup \{(0,0)\}$, and Σ is the convex cone generated by the squares of elements of $V(\mathbb{R})$. Modifying the proof of [2, Theorem 6.3.2], we can show that Σ is closed in $V(\mathbb{R})$. Since $(4,2), (2,4), (2,2)$ and $(0,0)$ belong to S, we have

$$p := \chi_{(4,2)} + \chi_{(2,4)} - \chi_{(2,2)} + \chi_{(0,0)} \in V(\mathbb{R}),$$

and for $(x,y) \in (\mathbb{R}\setminus\{0\})^2 \cup \{(0,0)\}$

$$p(x,y) = (\chi_{(4,2)} + \chi_{(2,4)} - \chi_{(2,2)} + \chi_{(0,0)})(x,y) = x^2y^2(x^2+y^2-1)+1.$$

By a similar argument to the proof of [2, Lemma 6.3.1], we have $p \in V(\mathbb{R})_+ \setminus \Sigma$, hence $V(\mathbb{R})_+ \setminus \Sigma \neq \emptyset$. Therefore S is not semiperfect by Theorem 2.4. In case (ii) or (iii), we can reduce to the case $\alpha = 0$ by an isomorphic identification, hence S is a subsemigroup of $\mathbb{N}_0 \times \mathbb{Z}$ and is not contained in any sectors with angle $< \pi$ and any lines through the origin $(0,0)$. Furthermore we can assume that $h_{1,0}(S) = \mathbb{N}_0, h_{0,1}(S) = \mathbb{Z}$ and $(1,0) \in S$. In case (ii), the dual semigroup S^* is

$$\{(m_1,m_2) \mapsto x^{m_1}y^{m_2} \mid x,y \in \mathbb{R}\setminus\{0\}\} \cup \{\mathbf{1}_{\{(0,0)\}}\}.$$

As in (i), S^* separates the points of S and Σ is closed in $V(\mathbb{R})$. Since $h_{0,1}(S) = \mathbb{Z}$, there exists $n \in \mathbb{N}$ such that $(n,1) \in S$. Hence $(4n+2,2)$, $(4n,2)$, $(4n,4)$ and $(0,0)$ belong to S. Therefore, by a similar argument to (i),

$$\chi_{(4n+2,2)} + \chi_{(4n,4)} - \chi_{(4n,2)} + \chi_{(0,0)} \in V(\mathbb{R})_+ \setminus \Sigma.$$

We have $V(\mathbb{R})_+ \setminus \Sigma \neq \emptyset$, hence S is not semiperfect. In case (iii), the face $S \cap (\{0\}\times\mathbb{Z})$ is semiperfect by Theorem 2.1, hence is singly generated by Theorem 3.1 and the assertion for k=1. The following cases are possible:

(iiia) S is a subsemigroup of $\mathbb{N}_0 \times \mathbb{Z}$ and there exists an odd number $p \geq 1$ such that $S \cap (\{0\}\times\mathbb{Z}) = \{0\} \times p\mathbb{Z}$;
(iiib) S is a subsemigroup of $\mathbb{N}_0 \times \mathbb{Z}$ and there exists an even number $p \geq 2$ such that $S \cap (\{0\}\times\mathbb{Z}) = \{0\} \times p\mathbb{Z}$;
(iiic) S is a subsemigroup of $\mathbb{N}_0 \times \mathbb{Z}$ and there exists an odd number $p \geq 1$ such that $S \cap (\{0\}\times\mathbb{Z}) = \{0\} \times p\mathbb{N}_0$;
(iiid) S is a subsemigroup of $\mathbb{N}_0 \times \mathbb{Z}$ and there exists an even number $p \geq 2$ such that $S \cap (\{0\}\times\mathbb{Z}) = \{0\} \times p\mathbb{N}_0$.

Corresponding to each case the dual semigroup S^* is

$$\begin{aligned}
&\{(m_1,m_2) \mapsto x^{m_1}y^{m_2} \,|x,y \in \mathbb{R}\setminus\{0\}\} \\
&\qquad \cup \{(m_1,m_2) \mapsto \mathbf{1}_{\{0\}}(m_1)y^{m_2} \,|y \in \mathbb{R}\setminus\{0\}\} && \text{if (iiia)},\\
&\{(m_1,m_2) \mapsto x^{m_1}y^{m_2} \mid x,y \in \mathbb{R}\setminus\{0\}\} \cup \{(m_1,m_2) \mapsto \mathbf{1}_{\{0\}}(m_1)y^{m_2} \mid y > 0\} \\
&\qquad \cup \{(m_1,m_2) \mapsto \mathbf{1}_{\{0\}}(m_1)(-1)^{\frac{m_2}{p}}y^{m_2} \mid y > 0\} && \text{if (iiib)},\\
&\{(m_1,m_2) \mapsto x^{m_1}y^{m_2} \mid x,y \in \mathbb{R}\setminus\{0\}\} \\
&\qquad \cup \{(m_1,m_2) \mapsto \mathbf{1}_{\{0\}}(m_1)y^{m_2} \mid y \in \mathbb{R}\} && \text{if (iiic)},\\
&\{(m_1,m_2) \mapsto x^{m_1}y^{m_2} \mid x,y \in \mathbb{R}\setminus\{0\}\} \cup \{(m_1,m_2) \mapsto \mathbf{1}_{\{0\}}(m_1)y^{m_2} \mid y \geq 0\} \\
&\qquad \cup \{(m_1,m_2) \mapsto \mathbf{1}_{\{0\}}(m_1)(-1)^{\frac{m_2}{p}}y^{m_2} \mid y > 0\} && \text{if (iiid)}.
\end{aligned}$$

Then, if S satisfies one of the conditions (iiia)–(iiid), we can prove that S is not semiperfect as in (ii). As in the proof (III) of Theorem 3.1, we can easily prove the assertions for $k \geq 3$ by induction. ■

The following is an immediate consequence of Theorem 4.1.

Corollary 4.2. *For $k \geq 1$, let S be a finitely generated subsemigroup of $\mathbb{Q}^k$. Then S is semiperfect if and only if S is trivial or singly generated. Furthermore, S is perfect if and only if S is trivial.* ■

References

[1] Berg, C., J. P. R. Christensen and C. U. Jensen, *A remark on the multidimensional moment problem*, Math. Ann. **243** (1979), 163–169.

[2] Berg, C., J. P. R. Christensen and P. Ressel, " Harmonic Analysis on Semigroups," Springer-Verlag, New York-Berlin-Heidelberg-Tokyo, 1984.

[3] Bisgaard, T. M., *Characterization of perfect involution groups*, Math. Scand. **65** (1989), 245–258.

[4] Bisgaard, T. M., *The two-sided complex moment problem*, Ark. Mat. **27** (1989), 23–28.

[5] Bisgaard, T. M. and P. Ressel, *Unique disintegration of arbitrary positive definite functions on *-divisible semigroups*, Math. Z. **200** (1989), 511–525.

[6] Boas, R. P., Jr., *The Stieltjes moment problem for functions of bounded variation*, Bull. Amer. Math. Soc. **45** (1939), 399–404.

[7] Jones, W. B., O. Njåstad and W. J. Thron, *Orthogonal Laurent polynomials and the strong Hamburger moment problem*, J. Math. Anal. Appl. **98** (1984), 528–554.

[8] Nakamura, Y. and N. Sakakibara, *Perfectness of certain subsemigroups of a perfect semigroup*, Math. Ann. **287** (1990), 213–220.

[9] Shohat, J. A. and J. D. Tamarkin, "The Problem of Moments," Math. Surveys 1, American Mathematical Society, Providence, R. I., 1943.

Asahikawa National College of
Technology
Syunkohdai 2-2
Asahikawa 071, Japan

Received May 26, 1991
and in final form February 20, 1992

Semigroup Forum Vol. 45 (1992) 249–260

RESEARCH ARTICLE

Presentations of Alternating Semigroups

Stephen L. Lipscomb[1]

Communicated by J. M. Howie

Abstract. One presentation of the alternating group A_n has $n-2$ generators $s_1, \ldots, s_{n-2}$ and relations $s_1^3 = s_i^2 = (s_{i-1}s_i)^3 = (s_j s_k)^2 = 1$, where $i > 1$ and $|j-k| > 1$. Against this backdrop, a presentation of the alternating semigroup $A_n^c \supset A_n$ is introduced: It has $n-1$ generators $s_1, \ldots, s_{n-2}, e$, the A_n-relations (above), and relations $e^2 = e$, $(es_1)^3 = (es_1)^4$, $(es_j)^2 = (es_j)^4$, $es_i = s_i s_1^{-1} e s_1$, where $j > 1$ and $i \geq 1$.

Path Notation and Background

Generally, the semigroup notations/concepts follow either Howie's text [2], or Petrich's text [8]. In addition however, the members of the *symmetric inverse semigroup* C_n on the n symbols $1, 2, \ldots, n$ will be called *charts* instead of (the usual) one-one partial transformations. Also, $S_n \subset C_n$ denotes the symmetric group on these same n symbols.

Path notation [3],[4],[5] is to *charts*, as cycle notation is to permutations. And path notation is to *even charts*, as cycle notation is to even permutations [5]. To make this precise, we start with a review of the path notation.

Define the 0-*path* to be the 0 or empty transformation. Let $i_1, \ldots, i_k$ be distinct elements of $N = \{1, 2, \ldots, n\}$. If $a \in C_n$ has domain $\{i_1, \ldots, i_k\}$ and

$$i_1 a = i_2,\ i_2 a = i_3,\ \ldots,\ i_{k-1} a = i_k,\ i_k a = q;$$

then a is called a *path*. If $q = i_1$, then a is a k-*circuit*, or a k-*path of length* k. In the other case, when $q \neq i_1$, we have that q is not in the domain $\mathbf{d}a$ of a, and a is called a *proper* k-*path of length* $(k+1)$. Let $\mathbf{r}a$ denote the range of the chart a. Then for charts a, b with $(\mathbf{d}a \cup \mathbf{r}a)$ and $(\mathbf{d}b \cup \mathbf{r}b)$ disjoint, define a chart c with domain $\mathbf{d}a \cup \mathbf{d}b$ by setting $xc = xa, x \in \mathbf{d}a$, and $xc = xb, x \in \mathbf{d}b$. Call c the *join of* a *and* b and write $c = ab = ba$.

If the join of a and b exists, then a *and* b *are disjoint.* The cycle decomposition of a permutation is a join of disjoint circuits.

The chart a: *moves* x whenever $x \in \mathbf{d}a$ and $xa \neq x$; and *fixes* x whenever $x \in \mathbf{d}a$ and $xa = x$. The notation for $0 \in C_n$ is $(1](2] \cdots (n]$; for a 1-circuit with domain $\{i\}$, $(1] \cdots (i-1](i)(i+1] \cdots (n]$; and for a proper 1-path a, say a maps 1 to 2, $(12](3] \cdots (n]$. More generally, represent a k-circuit a as

$$a = (i_1, \ldots, i_k)(j_1](j_2] \cdots (j_m]$$

where a moves $i_1, \ldots, i_k$ and is not defined at $j_1, \ldots, j_m$; and denote a proper k-path a as

$$a = (i_1, \ldots, i_k, q](j_1] \cdots (j_t]$$

[1] Part of this research was supported by a Mary Washington College faculty development grant.

where a moves $i_1, \ldots, i_k$ and is not defined at $q, j_1, \ldots, j_t$. Within certain contexts, the explicit appearance of length-one paths, e.g., $(j]$ and (j), is unnecessary.

Multiplication of charts in path notation is like multiplication of permutations. Consider, $a = (12345]$, $a^2 = (135](24]$, $a^3 = (14](3](25]$, $a^4 = (15](4](3](2]$, and $a^5 = (1](2](3](4](5] = 0$.

When studying charts via path notation, the next two theorems are fundamental [3].

Theorem 1 (Unique Representation of Charts). *Every chart in C_n is the join of disjoint paths. Furthermore, if $a = b_1 b_2 \cdots b_r$ is a join representation of paths with each b_i of length at least 2, then this factorization is unique except for the order in which the paths are written.* ■

Theorem 2 (Proper Path Multiplication). *If the chart $a = (123 \cdots k]$ denotes a proper path, then $a^k = 0$ and $a^m \neq 0$ for each $m < k$.* ■

Even Charts and the Alternating Semigroup

The idea of an *even chart* was introduced in [5]: Just as (i, j) denotes a permutation, i.e., the length-one cycles do not explicitly appear in the notation, $(i, j]$ will denote a rank $n-1$ chart, i.e., again the length-one circuits do not explicitly appear. With this convention, for $2 \leq r \leq n$, consider the factorization of $(12 \cdots r) \in S_n$ into transpositions

$$(123 \cdots r) = (r-1, r)(r-2, r-1) \cdots (1, 2),$$

and the analogous factorization (replace each ")" with "]")

(1)1 $$(12 \cdots r] = (r-1, r] \circ (r-2, r-1] \circ \cdots \circ (1, 2]$$

of $(12 \cdots r] \in C_n$. However, not every such (transposition) factorization induces a factorization of the corresponding rank $n-1$ path. For example,

$$(123 \cdots r) = (r, 1)(r, 2) \cdots (r, r-1),$$

but

(2)2 $$(12 \cdots r] \neq (r, 1] \circ (r, 2] \circ \cdots \circ (r, r-1].$$

In fact, the composition on the right side of (2) is a rank $(n - r + 1)$ chart moving r to 1 and fixing all $k > r$.

A *transpositional* is a chart that is either a transposition (i, j) or a *semi-transposition* $(i, j]$, i.e., a rank $(n-1)$ chart of the form $(i, j]$. A chart is *even* if it is a product of an even number of transpositionals. First, $A_1^c = C_1$. Second, for $n \geq 2$, the *alternating semigroup* $A_n^c \subset C_n$ is the subsemigroup containing the even charts. Clearly, A_2^c contains all idempotents in C_2. From Theorem 3 (below), every member of A_2^c is an idempotent.

Unlike permutations, some charts are both "even" and "odd," e.g.,

$$(12](3] = (12] \circ (13] = (12] \circ (23] \circ (32].$$

However, when n is odd, then the rank $n-1$ chart $(12\cdots n]$ is "even" and not odd. This follows from a result of Gomes and Howie [1].

To see the connection, some terminology in [1] is translated into path notation. First,

$$C_n = J_0 \cup J_1 \cup J_2 \cup \cdots \cup J_{n-2} \cup J_{n-1} \cup J_n$$

where J_k, $0 \leq k \leq n$, denotes the charts in C_n of rank k. A *nilpotent* chart is a join of proper paths. Every chart $c = \rho\eta$ is a join of a *permutation part* ρ (on some subset of N), and a *nilpotent part* η. In [1], N_1 denotes the nilpotent elements of rank $n-1$. Thus, in path notation, we picture the elements in N_1 as those charts having the form $(12\cdots n]$. Also, we define a new version of the *completion* $\mathbf{c}$ *of a chart* $c \in C_n$: mechanically we replace, in the path decomposition of c, each instance of "]" with ")". This "completion operation" maps C_n into S_n, and is the identity on S_n. For charts of rank $n-1$, this "completion" agrees with the "completion" presented in [1].

The structure of A_n^c can be deduced from the next three Theorems: (Proofs of these Theorems appear in [6]. A proof of Theorem 3 involving Gomes' and Howie's [1] "tidy product" appears in [5].)

Theorem 3 (The even charts of rank $n-1$). *If $a \in C_n$ has rank $n-1$, i.e., $a \in J_{n-1}$, then $a \in A_n^c$ if, and only if, its completion $\mathbf{a}$ is an even permutation of N.* ∎

Theorem 4 (All charts of rank $< n-1$ are even). *If $a \in C_n$, and a is of rank at most $n-2$, then $a \in A_n^c$.* ∎

Theorem 5 (There are enough even charts). *Let $n \geq 3$. If $A, B \subset \{1, 2, \ldots, n\}$ are of the same size, then there is a chart $c \in A_n^c$ whose domain is A and whose range is B.* ∎

Let J_{n-1}^e denote the rank $n-1$ charts whose completions are even, then from Theorems 3 and 4 we have

$$A_n^c = A_n \cup J_{n-1}^e \cup J_{n-2} \cup \cdots \cup J_0. \tag{3}$$

Moreover, from Theorem 5, the parts of this partition (3) of A_n^c are Green's $\mathcal{D}$-classes.

As for generating sets of A_n^c, we have the following Theorem:

Theorem 6. *Let $n \geq 3$. If Y is a generating set for the alternating semigroup A_n and $c \in A_n^c$ is any even chart of rank $n-1$, then $X = Y \cup \{c\}$ is a generating set for A_n^c.*

Proof. By [6, Theorem 5] A_n^c is generated by the totality of paths that are either length 3 cycles (i, j, k), or rank $n-1$ charts of the form $(i, j, k]$. Thus, since Y is a generating set for $A_n \subset A_n^c$, we only need to show that every (rank $n-1$) chart $(ijk]$ is a product of members of X. First, since the given chart $c \in A_n^c$ is of rank $n-1$, there is a $q \notin \mathbf{d}c$ such that

$$cc^{-1} = (1)\cdots(q-1)(q](q+1)\cdots(n)$$

is an idempotent. Second, since $n \geq 3$, choose $\gamma = (kqm) \in A_n$ and calculate the idempotent

$$e_k = \gamma c c^{-1} \gamma^{-1} = (1) \cdots (k-1)(k](k+1) \cdots (n).$$

Then, since both (ijk) and γ are products of members of Y, we see that

$$(ijk] = e_k \circ (ijk) = \gamma c c^{-1} \gamma^{-1} \circ (ijk)$$

is a product of members of X. ■

Defining Relations for $A_n^c \longleftarrow n \geq 3 \longleftarrow$

Recall [10, p 298] that the alternating group A_n has the presentation

$$A_n \equiv \langle\, s_1, \dots, s_{n-2} \mid s_1^3 = s_i^2 = (s_{i-1}s_i)^3 = (s_j s_k)^2 = 1,\ i > 1,\ |j-k| > 1\,\rangle.$$

In particular, A_n is generated by the $n-2$ permutations

$$(1,2,3), (1,2)(3,4), \dots, (1,2)(i+1, i+2), \dots, (1,2)(n-1,n).$$

Set

$$s_1 = (1,2,3) \quad \text{and, for } i > 1, \quad s_i = (1,2)(i+1, i+2).$$

Then, using cycle notation, we can calculate and verify that the defining relations among the members of $\{s_1 = (1,2), s_2 = (1,2)(3,4), \dots, s_{n-2} = (1,2)(n-1,n)\}$ do indeed hold.

For the alternating semigroup A_n^c, it is natural to ask about a corresponding presentation. By Theorem 6, for $n \geq 3$, any set of generators of A_n coupled with any rank $n-1$ even chart yield a generating set for A_n^c. Thus, for a generating set of A_n^c we may take

$$X = \{\ s_1 = (1,2,3),\ s_2 = (1,2)(3,4), \dots,\ s_{n-2} = (1,2)(n-1,n),\\ e = (1](2) \cdots (n)\ \}.$$

Then, using path notation, we can calculate and verify that the defining relations (listed in Theorem 7 below) do indeed hold.

Just as presentations of A_n involve free groups, presentations of A_n^c involve free inverse monoids. Here, the appropriate free inverse monoid M_X on a generating set X is the one constructed via the Wagner Congruence (see [8 p. 356] or [9 p. 480]).

Theorem 7 (Defining Relations for A_n^c). *Let $n \geq 3$. Let M_X be the free inverse monoid generated by $X = \{s_1, s_2, \dots, s_{n-2}, e\}$. Let $R \subset M_X \times M_X$ be induced from the relations*

$$\begin{aligned} &s_1^3 = (s_i)^2 = (s_{i-1}s_i)^3 = (s_j s_k)^2 = 1, && \textit{for}\quad i > 1,\ \ |j-k| > 1,\\ &e^2 = e,\ \ (es_1)^3 = (es_1)^4, && \textit{and}\\ (4)\quad &(es_j)^2 = (es_j)^4,\ \ es_i = s_i s_1^{-1} e s_1, && \textit{for}\quad j > 1, i \geq 1. \end{aligned}$$

Let $\rho = R^{\#}$ be the smallest congruence on M_X that contains R. Then M_X/ρ is isomorphic to A_n^c. ■

The proof of Theorem 7, given below just prior to the reference list, requires (a technical) Lemma 8 and (an inductive) Lemma 10. To state and prove these requisite lemmas, we first identify, in "word terminology," the rank $n-1$ idempotents. To do this, let

$$\begin{aligned}
e_1 &= e \\
e_2 &= s_1^{-1} e_1 s_1 \quad \text{(Also, } e_2 = s_2 e_1 s_2 \text{ since, from (4), } es_2 = s_2 e_2\text{)} \\
e_3 &= s_1^{-1} e_2 s_1 \\
e_4 &= s_2 e_3 s_2 \\
e_5 &= s_3 e_4 s_3 \\
&\ \vdots \\
e_n &= s_{n-2} e_{n-1} s_{n-2}.
\end{aligned} \tag{5}$$

These "word forms" were motivated by the realization that all rank $n-1$ idempotents in C_n have the same path structure, i.e., all are "conjugate" [5] to $(1](2)\cdots(n)$. For example, $e_2 = s_1^{-1} e_1 s_1$ was motivated by

$$(1)(2](3)\cdots(n) = (321) \circ (1](2)\cdots(n) \circ (123).$$

The Technical Lemma

Lemma 8 and its proof translate (from C_n into word terminology) various compositions of generators of A_n with rank $n-1$ idempotents. The "word forms" appearing in Lemma 8 were discovered by experimenting with path-multiplication of those charts that correspond to members of X.

Lemma 8. *Let R and M_X/ρ be as in Theorem 7. If "equality" is understood to mean "equality modulo ρ," then*

(1) *Each $e_k\rho$, $1 \le k \le n$, is an idempotent in M_X/ρ.*
(2) *In addition to the relations in R, we have relations $(es_1)^3 = (s_1 e)^3 = (s_1 e)^4$. Also, $(es_1)^3 = (s_1^{-1}e)^3 = (s_1^{-1}e)^4 = (es_1^{-1})^3 = (es_1^{-1})^4$.*
(3) *For relations among the e_k and the s_i we have*

for $k=1$, $\quad e_1 s_i = s_i e_2$, $i \ge 1$, *and* $e_1 s_1^{-1} = s_1^{-1} e_3$;

for $k=2$, $\quad e_2 s_i = s_i e_1$, $i \ge 2$, $e_2 s_1^{-1} = s_1^{-1} e_1$, *and* $e_2 s_1 = s_1 e_3$;

for $k=3$, $\quad e_3 s_i = s_i e_3$, $i \ge 3$, $e_3 s_1^{-1} = s_1^{-1} e_2$, $e_3 s_1 = s_1 e_1$, *and* $e_3 s_2 = s_2 e_4$;

for $k \ge 4$, $\quad e_{k+1} s_{k-1} = s_{k-1} e_k$ *and* $e_k s_{k-1} = s_{k-1} e_{k+1}$;
$e_k s_i = s_i e_k$ *where* $i \ne k-1, k-2$;
$e_k s_1^{-1} = s_1^{-1} e_k$.

(4) *For any word w in M_X,*

$$w = t e_{i_1} e_{i_2} \cdots e_{i_k} \text{ where } t = s_{j_1} s_{j_2} \cdots s_{j_q}.$$

(5) *For s_1 and e_1, e_2, e_3 we have the relations*

$$s_1 e_1 e_2 e_3 = e_1 e_2 e_3 s_1 = e_1 e_2 e_3 \quad \text{and} \quad s_1^{-1} e_1 e_2 e_3 = e_1 e_2 e_3 s_1^{-1} = e_1 e_2 e_3.$$

Proof. (Of (1)): Use induction on k, $1 \le k \le n$: Clearly, $e_1\rho = e\rho$ is an idempotent since $(e^2, e) \in \rho$. Then calculate $e_k e_k$ in terms of e_{k-1}, and observe that $e_k\rho$ is an idempotent if $e_{k-1}\rho$ is. (Of (2)): To simplify notation let $s = s_1$. Then, the first equality holds since

$$\begin{aligned}(es)^3 &= (ss^{-1})(es)^3 = s(s^{-1}es)eses = se_2e_1ses \\ &= se(s^{-1}es)ses = ses^{-1}e(s^{-1}es) \qquad (\text{since } s^2 = s^{-1}) \\ &= ses^{-1}(s^{-1}es)e = sesese \qquad (\text{since } s^{-2} = s).\end{aligned}$$

Using $(se)^3 = (es)^3$ and $(es)^3 = (es)^4$, the second equality will follow if

$$(es)^4 = (se)^4. \tag{6}$$

To see that (6) holds, use the fact that es commutes with $(es)^3 = (se)^3$:

$$(es)^4 = es(se)^3 = es(se)^3e = \big((se)^3es\big)e = (se)^4.$$

The equality $(es)^3 = (s^{-1}e)^3$ holds since $(es)^3\rho$ is an idempotent:

$$(es)^3(es)^3 = (es)^4(es)^2 = (es)^3(es)^2 = (es)^4(es) = (es)^3(es) = (es)^4 = (es)^3,$$

and $(s^{-1}e)^3\rho$ is the inverse of (and therefore equal to) $(es)^3\rho$. The equality $(s^{-1}e)^3 = (s^{-1}e)^4$ follows since $(es)^3 = (es)^4$ and inverses are unique. Similarly, $(s^{-1}e)^3 = (es^{-1})^3$ and $(s^{-1}e)^4 = (es^{-1})^4$ since $(es)^3 = (se)^3$ and $(es)^4 = (se)^4$. (Of (3)): (Continue the notation $s = s_1$.) For $k = 1$, the first set of equalities follows from $es_i = s_is^{-1}es$ and $e_2 = s^{-1}es$; while the other equality follows from

$$se_1s^{-1} = s^{-1}s^{-1}ess = s^{-1}e_2s = e_3$$

and multiplication (on the left) by s^{-1}. For $k = 2$, the first set (of equalities) follows by uniqueness of the inverse of $e_1s_i = s_ie_2$ (since $i \ge 2 \Rightarrow s_i^2 = 1$); while the other equalities, those involving either s^{-1} or s, follow from the definitions of e_2 and e_3, respectively. For $k = 3$, let $i \ge 3$, and consider

$$\begin{aligned}e_3s_i &= s^{-1}s^{-1}esss_i = sesss_i \qquad (\text{since } s^{-2} = s) \\ &= ses(s_is^{-1}) \qquad (\text{since } (s_1s_i)^2 = 1 \text{ for } i \ge 3) \\ &= s(es_i)s^{-1}s^{-1} = ss_ie_2s^{-1}s^{-1} \qquad (\text{from case } k = 1) \\ &= s_is^{-1}e_2s^{-1}s^{-1} = s_is^{-1}e_2s = s_ie_3.\end{aligned}$$

And for equalities involving either s^{-1} or s, use the uniqueness of the inverse of $e_2s_1 = s_1e_3$ and the uniqueness of the inverse of $e_1s_1^{-1} = s_1^{-1}e_3$. The equality $e_3s_2 \doteq s_2e_4$ follows from the definition of e_4 and $s_2^2 = 1$. Next, let $k \ge 4$. Then first,

$$e_{k+1} = s_{k-1}e_ks_{k-1} \quad \text{and} \quad s_{k-1}^2 = 1$$

yield

$$e_{k+1}s_{k-1} = s_{k-1}e_k \quad \text{and } e_ks_{k-1} = s_{k-1}e_{k+1}.$$

Second, continuing with $k \geq 4$, we show $e_k s_i = s_i e_k$, for $i \neq k-1, k-2$, by considering three subcases, $i \geq k$, and $1 < i < k-2$, and $i = 1$: Assuming $i \geq k \geq 4$, suppose $e_{k-1} s_i = s_i e_{k-1}$ for $i \geq k-1 \geq 3$. Then we have

$$\begin{aligned} e_k s_i = (s_{k-2} e_{k-1} s_{k-2}) s_i &= s_{k-2} e_{k-1} s_i s_{k-2} && (\text{since } i \geq k \Rightarrow (s_{k-2} s_i)^2 = 1) \\ &= s_{k-2} s_i e_{k-1} s_{k-2} && (\text{since } i \geq k-1 \geq 3) \\ &= s_i e_k. \end{aligned}$$

Thus, in particular, $e_k s_k = s_k e_k$ for all $k \geq 3$. Assuming $1 < i < k-2$, we have

$$\begin{aligned} e_k s_i &= (s_{k-2} \cdots s_{i+1} s_i e_{i+1} s_i s_{i+1} \cdots s_{k-2}) s_i \\ &= \cdots s_{i+1} s_i e_{i+1} (s_i s_{i+1} s_i) \cdots \\ &= \cdots s_{i+1} s_i e_{i+1} (s_{i+1} s_i s_{i+1}) \cdots && (\text{since } (s_i s_{i+1})^3 = 1) \\ &= \cdots s_{i+1} s_i (s_{i+1} e_{i+1}) s_i s_{i+1} \cdots && (\text{since } i+1 \geq 3) \\ &= \cdots (s_i s_{i+1} s_i e_{i+1} s_i s_{i+1}) \cdots \\ &= s_i e_k. \end{aligned}$$

Assuming $i = 1$ (and still using $s = s_1$), we have

$$e_k s = (s_{k-2} \cdots e_3 \cdots s_{k-2}) s.$$

And since $s_q s = s^{-1} s_q$ and $s_q s^{-1} = s s_q$ when $q > 2$, we also have, for $k \geq 4$,

$$e_k s = \begin{cases} \cdots s_2 s_1^{-1} e_2 s_1 s_2 s \cdots, & \text{when } k \text{ even} \\ \cdots s_3 s_2 s_1^{-1} e_2 s_1 s_2 s^{-1} s_3 \cdots, & \text{when } k \text{ odd.} \end{cases}$$

When $k \geq 4$ is even, we continue with

$$\begin{aligned} e_k s_1 &= \cdots s_2 s^{-1} e_2 (s s_2 s) \cdots \\ &= \cdots s_2 s^{-1} e_2 s_2 s^{-1} s_2 \cdots && (\text{since } s s_2 s = s_2 s^{-1} s_2) \\ &= \cdots s_2 s^{-1} (s_2 e_1) s^{-1} s_2 \cdots && (\text{since } e_2 s_2 = s_2 e_1) \\ &= \cdots s s_2 s e_1 s^{-1} s_2 \cdots \\ &= s (s_{k-2} \cdots s_2 s e_1 s^{-1} s_2 \cdots s_{k-2}) && (\text{since } k \text{ even}) \\ &= s (s_{k-2} \cdots s_2 s^{-2} e_1 s^2 s_2 \cdots s_{k-2}) \\ &= s_1 e_k. \end{aligned}$$

For the case where $k \geq 4$ is odd, we shall need

$$e_4 s^{-1} = s^{-1} e_4 :$$

But, since inverses are unique, this follows from the known equality $e_4 s = s e_4$. Now we are ready to consider the case where $k \geq 4$ is odd:

$$\begin{aligned} e_k s_1 &= \cdots s_3 s_2 (e_3 s_2) s^{-1} s_3 \cdots \\ &= \cdots s_3 s_2 (s_2 e_4) s^{-1} s_3 \cdots \\ &= \cdots s_3 s_2 s_2 s^{-1} e_4 s_3 \cdots && (\text{since } e_4 s^{-1} = s^{-1} e_4) \\ &= s (s_{k-2} \cdots s_3 e_4 s_3 \cdots s_{k-2}) && (\text{since } s_2^2 = 1 \text{ and } k \text{ odd}) \\ &= s_1 e_k. \end{aligned}$$

Finally, we consider

$$e_k s_1^{-1} = s_1^{-1} e_k.$$

This case follows from the uniqueness of the inverse of $e_k s_1 = s_1 e_k$.
(Of (4)): This follows from part (3) since we can "move the e_k's to the right."
(Of (5)): From the definitions of e_1, e_2, e_3 and part (2) we have

$$\begin{aligned} s_1 e_1 e_2 e_3 &= se(s^{-1}es)(s^{-1}s^{-1}ess) = s(es^{-1}es^{-1}es^{-1}) \\ &= s(eseses) = (sesese)s = (seseses e)s \\ &= (es^{-1}es^{-1}es^{-1}es^{-1})s = (es^{-1}es^{-1}es^{-1})es^{-1}s \\ &= (es^{-1}es^{-1}es^{-1})e = (es^{-1}e(ss^{-1})s^{-1}ess)e \\ &= (e(s^{-1}es)(s^{-1}s^{-1}ess))e = e_1e_2e_3e_1 = e_1e_2e_3. \end{aligned}$$

To see that s_1 commutes with $e_1e_2e_3$, use part (3), $k = 1, 2, 3$, and calculate $s_1e_1e_2e_3 = e_3s_1e_2e_3 = e_3e_1s_1e_3 = e_3e_1e_2s_1$. Thus, the first string of equalities in part (5) holds. The last string of equalities now follows by taking the inverse of each "word," i.e., ρ-class, in the first string. This finishes the proof. ■

The Inductive Lemma

We have $(M_X/\rho, \iota)$ as the freest inverse monoid generated by $X = \{s_1, \ldots, s_{n-2}, e\}$ and satisfying the relations (4) (Theorem 7). In particular, we have $\iota : X \to M_X/\rho$ where $x\iota = x\rho$, and $\phi : X \to A_n^c$ where $s_1\phi = (1,2,3)$, $s_i\phi = (1,2)(i+1, i+2)$, $i > 1$, and $e\phi = (1](2)\cdots(n)$. The relative freeness of M_X/ρ says that there is a unique epimorphism $\psi : M_X/\rho \to A_n^c$ that makes the commutative diagram:

$$\begin{tikzcd} X \arrow[r, "\phi"] \arrow[d, "\rho = \iota"'] & A_n^c \\ M_X/\rho \arrow[ur, dashed, "\psi"'] & \end{tikzcd}$$

That is, $(x\iota)\psi = (x\rho)\psi = x\phi$. To simplify notation, denote **both** ι (from X to M_X/ρ) and the natural map $\rho^\natural$ (from M_X to M_X/ρ) as ρ. Moreover, for $B = \{i_1, i_2, \ldots, i_m\} \subset \{1, 2, \ldots, n\}$ define

$$e_B = e_{i_1} e_{i_2} \cdots e_{i_m},$$

where the ρ-class containing the word on the right is well-defined since idempotents commute in M_X/ρ.

Next, define the composition of $\rho : M_X \to M_X/\rho$ followed by ψ as the "hat" map:

$$\widehat{\ } = \rho \circ \psi \quad \text{or} \quad \widehat{w} = (w\rho)\psi.$$

Thus, using t as in part (4) of Lemma 8, we have

$$\widehat{te_B} = (te_B)\rho\psi = (t)\rho\psi\,(e_B)\rho\psi = \widehat{t}\widehat{e_B}.$$

Using this notation, we can state and prove the following Lemma 9 (which is needed to prove the Inductive Lemma 10).

Lemma 9. *If t_{ijk} is a word in M_X such that $\widehat{t}_{ijk} = (i,j,k) \in A_n$ is a 3-cycle, then $t_{ijk}e_ie_je_k = e_ie_je_k$ (mod ρ).*

Proof. Certainly each of $\widehat{s}_1 = (1,2,3)$ and $\widehat{s_1^{-1}} = (1,3,2)$ is conjugate within the symmetric group S_n to $\widehat{t}_{ijk} = (i,j,k)$. If $\sigma^{-1}\widehat{s}_1\sigma = (i,j,k)$, where σ is odd, then

$$\big((2,3)\sigma\big)^{-1}\widehat{s_1^{-1}}\big((2,3)\sigma\big) = (i,j,k),$$

and $(2,3)\sigma$ is even. Thus there exists $w \in M_X$ with $w\rho \in \langle\{s_i\rho\}\rangle$ and $\widehat{w} \in A_n$ such that either

$$\widehat{w}^{-1}\widehat{s}_1\widehat{w} = \widehat{t}_{ijk} \quad \text{or} \quad \widehat{w}^{-1}\widehat{s_1^{-1}}\widehat{w} = \widehat{t}_{ijk}.$$

Second, recall from the group theory case that ψ restricted to $\langle\{s_i\rho\}\rangle$ is an isomorphism onto $A_n \subset A_n^c$. It then follows that either

$$w^{-1}s_1w = t_{ijk} \ (\text{mod } \rho) \quad \text{or} \quad w^{-1}s_1^{-1}w = t_{ijk} \ (\text{mod } \rho).$$

(We shall assume the former ρ-equivalence since the other case is similar). Third, we show that for each $m \in \{1,2,3\}$,

$$w^{-1}e_mw = e_{m\widehat{w}} \ (\text{mod } \rho). \tag{7}$$

To see that (7) is indeed true, apply Lemma 5 part (4) to "move e_m to the right," i.e.,

$$w^{-1}e_mw = ze_\ell \ (\text{mod } \rho),$$

where $\widehat{z} \in A_n$. From this we deduce that $\widehat{w}^{-1}\widehat{e}_m\widehat{w} = \widehat{z}\widehat{e}_\ell$. On the other hand, it is clear that $\widehat{w}^{-1}\widehat{e}_m\widehat{w} = \widehat{e}_{m\widehat{w}}$, and so $\widehat{z}\widehat{e}_\ell = \widehat{e}_{m\widehat{w}}$. A comparison of the images of these two maps then gives $\ell = m\widehat{w}$, and it then follows that $i\widehat{z} = i$ except possibly for $i = m\widehat{w}$. Since $\widehat{z}$ is a permutation, we must have $\widehat{z} = 1$. Since the restriction of ψ to $\langle\{s_i\rho\}\rangle$ is an isomorphism, $z = 1$ (mod ρ), and (7) now follows. From (7) and $w\rho \in \langle\{s_i\rho\}\rangle$ we have

$$w^{-1}e_1e_2e_3w = e_ie_je_k \ (\text{mod } \rho).$$

We can now calculate that

$$(w^{-1}s_1w)(w^{-1}e_1e_2e_3w) = t_{ijk}e_ie_je_k \ (\text{mod } \rho).$$

And from this and Lemma 5 part (5), we see the conclusion of this Lemma is true. ∎

Lemma 10. *Let t and t_0 be members of $A_n\widehat{(\ )}^{-1}$, i.e., as the t in part (4) of Lemma 8. Then $\widehat{te}_B = \widehat{t}_0\widehat{e}_B$ implies $(te_B, t_0e_B) \in \rho$.*

Proof. Let p_m be the statement:

$$p_m \ \equiv \ \widehat{te}_B = \widehat{t}_0\widehat{e}_B \ \text{ when } \ |B| \le m \ \Rightarrow \ (te_B, t_0e_B) \in \rho.$$

We show p_0, p_1, p_2 and (the implication) "$p_{m-1} \Rightarrow p_m$" are true. If $m = 0$, then $B = \Gamma$ and $(t, t_0) \in \rho$ since, from group theory [10, p. 298], the homomorphism ψ carries the subgroup $\langle\{s_i\rho\}\rangle$ of M_X/ρ isomorphically onto the (alternating) subgroup A_n of A_n^c. If $m = 1$, then both charts $\widehat{t}\widehat{e}_i$ and $\widehat{t}_0\widehat{e}_i$ have rank $n-1$ and are one-one. It follows that $\widehat{t} = \widehat{t}_0$, and again from the group theory, $(t, t_0) \in \rho$. Hence, $(te_i, t_0e_i) \in \rho$. If $m = 2$, then

$$\widehat{t}\widehat{e}_i\widehat{e}_j = \widehat{t}_0\widehat{e}_i\widehat{e}_j$$

where the indices $i, j \notin \mathbf{r}(\widehat{t}\widehat{e}_i\widehat{e}_j) = \mathbf{r}(\widehat{t}_0\widehat{e}_i\widehat{e}_j)$. In this case we have $B = \{i, j\}$, and we need to show $(te_B, t_0e_B) \in \rho$. Again, it suffices to show that $\widehat{t} = \widehat{t}_0$. And to see that these *even* permutations are equal, we first pick an even permutation $\alpha \in A_n$ such that

$$\widehat{t}\alpha = \widehat{t}_0.$$

Second, we show that α is the identity permutation. Indeed, for $x \in N = \{1, 2, \ldots, n\}$, if $x\widehat{t} \notin B$, then $x\widehat{t}_0 \notin B$ and

$$x\widehat{t}\alpha = x\widehat{t}_0 = x\widehat{t}.$$

Thus, α fixes each member of $N - B$. It follows that $\alpha = 1$ or $\alpha = (ij)$. But since α is even, it cannot be the transposition (ij). Thus, $\widehat{t} = \widehat{t}_0$.

Next suppose $m = |B| \geq 3$, that p_{m-1} is true, and

$$c = \widehat{t}\widehat{e}_B = \widehat{t}_0\widehat{e}_B.$$

If for every $x \in \{1, 2, \ldots, n\}$ $x\widehat{t} = x\widehat{t}_0$ then, again by the group theory case, we are finished. Otherwise, there is an $x \notin \mathbf{d}(c)$ such that

$$j = x\widehat{t} \neq x\widehat{t}_0 = i \quad \text{and} \quad i, j \in B. \tag{8}$$

Since $|B| \geq 3$, we have the representation

$$c = \widehat{t}\widehat{e}_B = \widehat{t}\widehat{e}_i\widehat{e}_j\widehat{e}_k\widehat{e}_L = \widehat{t}_0\widehat{e}_i\widehat{e}_j\widehat{e}_k\widehat{e}_L = \widehat{t}_0\widehat{e}_B,$$

where $B = L \cup \{i, j, k\}$, $L \cap \{i, j, k\} = \Gamma$, and (8) holds. Now let

$$\widehat{t}_{ijk} = (i, j, k),$$

and apply Lemma 9 to obtain

$$t_0e_B = t_0e_ie_je_ke_L = t_0t_{ijk}e_ie_je_ke_L \pmod{\rho}.$$

It follows that if

$$t_* = t_0t_{ijk},$$

then

$$t_0e_B = t_*e_B \pmod{\rho}.$$

We now observe that $\widehat{t}$ agrees with $\widehat{t}_*$ on

$$\mathbf{d}(\widehat{t}\widehat{e}_B) \cup \{x\}.$$

Thus, if $C = B - \{j\}$, then

$$\widehat{t}\widehat{e}_C = \widehat{t}_*\widehat{e}_C$$

where $|C| = m - 1$. Hence, by the inductive hypothesis, $(te_C, t_*e_C) \in \rho$. But then

$$(te_Ce_j, t_*e_Ce_j) = (te_B, t_*e_B) \in \rho.$$

Since $t_0e_B = t_*e_B \pmod{\rho}$ we are finished. ∎

Proof (Of Theorem 7). The proof of Theorem 7 is now obvious: Indeed, from Lemma 8 we can assume that every $w \in M_X$ is in a ρ-class with a word of the form te_B. Then, the assumption that ψ is not one-one would contradict Lemma 10. Hence, the epimorphism ψ is also a monomorphism. Therefore, ψ is an isomorphism. ∎

Comments

Here, the approach to finding defining relations for A_n^c runs parallel to the approach used in the C_n case [7]. Indeed, the common elements in the mathematics behind these two cases may indicate a general approach to finding presentations of certain inverse semigroups.

Consider, for example, a given permutation group $G \subset S_n$ with presentation $\langle Y \mid R \rangle$. Suppose also that

$$S = \{\, c \,|\, c \text{ is a restriction of some } g \in G\} \subset C_n.$$

Then, what can be said about adding (idempotent) generators, via relations of the form $e = e^2$, to obtain a presentation of S?

In both of the A_n^c and C_n cases, the generators in the group presentations were first identified with permutations in cycle notation. An idempotent generator $e = e^2$ was then written in path notation, and new relations were "mechanically generated" by inspecting various path-multiplications.

How did we know which relations to incorporate? Examples in the A_n^c case illustrate one answer. Consider the relations $(es_1)^3 = (es_1)^4$ and $(es_j)^2 = (es_j)^4$, $j > 1$. We calculate in path notation that es_1 must correspond to the chart with path form

$$(1](2)\cdots(n)\circ(1,2,3) = (2,3,1](4)\cdots(n).$$

Such a path form yields a cyclic semigroup of index 3 and period 1 (see [3]). That is, the period 1 yields the exponent 4 on $(es_1)^4$. In contrast, the chart that corresponds to es_j, where $j > 1$, has path form

$$(1](2)\cdots(n)\circ(1,2)(j+1,j+2) = (2,1](j+1,j+2)(k_1)\cdots(k_{n-4}).$$

And, in this case, we have a cyclic semigroup of index 2 and period 2.

The basic idea was to pick those relations that show how the idempotent generator e "interacts" with each generator $y \in Y$.

Having expanded the generating set Y to X, and the relations R to R', we then encountered the word problem, i.e., we needed to establish that the unique homomorphism ψ was one-one. To do this, a technical lemma was needed to obtain canonical forms for words, and an inductive lemma was also needed to make use of the group presentation as the initial case in an inductive argument.

I thank J. M. Howie for his comments concerning the first and third parts of the proof of Lemma 9. Both parts are now more concise and more readily understandable.

References

[1] Gomes, G. M. S., and J. M. Howie, *Nilpotents in finite symmetric inverse semigroups*, Proc. Edinburgh Math. Soc. **30** (1987), 383–395.

[2] Howie, J. M., "An introduction to semigroup theory," Academic Press, New York, 1976.

[3] Lipscomb, S. L., *Cyclic subsemigroups of symmetric inverse semigroups*, Semigroup Forum **34** (1986), 243-248.

[4] Lipscomb, S. L., *The structure of the centralizer of a permutation*, Semigroup Forum **37** (1988), 301-312.

[5] Lipscomb, S. L., *Problems and applications of finite inverse semigroups*, in "Words, Languages, & Combinatorics," World Scientific Publishing Co Pte Ltd, Singapore, New Jersey, London, Hong Kong, 1992.

[6] Lipscomb, S.L., *The alternating semigroups: generators and congruences*, Semigroup Forum **44** (1992), 96–106.

[7] Lipscomb, S.L., *Presentations of symmetric inverse semigroups*, To appear.

[8] Petrich, M., "Inverse semigroups," John Wiley & Sons, New York, 1984.

[9] Reilly, N. R., *Free inverse semigroups*, Colloquia Mathematica Societatis János Bolyai, in Algebraic theory of semigroups, Amsterdam (1976), 479–508 (Proc. Conf. Szeged University 1976).

[10] Suzuki, M., "Group theory I," Springer-Verlag, Berlin, Heidelberg, New York, 1982.

[11] Wagner, V. V., *Generalized heaps and generalized groups with a transitive compatibility relation*, Učenye Zapiski Saratov Gos. Univ., bf 70 (1961), 25–39 (in Russian).

Department of Mathematics
Mary Washington College
Fredericksburg, VA 22401 (USA)

Received June 5, 1991
and in final form September 6, 1991

Semigroup Forum Vol. 45 (1992) 261–263

SHORT NOTE

Semigroups Associated with Generalized Brownian Functionals

Takeyuki Hida and Hui-Hsiung Kuo

Communicated by D. R. Brown and J. D. Lawson

Dedicated to Professor R. J. Koch on the occasion of his sixtieth birthday

1. Three examples of semigroups

The purpose of this note is to introduce the following three semigroups:

Semigroup R. Let R denote the half-open interval $[0, \frac{1}{2})$ of the real number system $\mathbb{R}$. R is a semigroup under the following operation

$$\alpha \circ \beta = \frac{\alpha + \beta - 4\alpha\beta}{1 - 4\alpha\beta}.$$

Semigroup J. Let $J = \{(a, \alpha) : a \in \mathbb{R},\ 0 \leq \alpha < \frac{1}{2}\}$. J is a semigroup under the following operation

$$(a, \alpha) \circ (b, \beta) = \left(\frac{a + b - 2(a\beta + b\alpha)}{1 - 4\alpha\beta}, \frac{\alpha + \beta - 4\alpha\beta}{1 - 4\alpha\beta} \right)$$

Semigroup K. Let $K = \{(n, r, \alpha) : n$ is a nonnegative integer, $r \in \mathbb{R}$, $r \neq 0$, $0 \leq \alpha < \frac{1}{2}\}$. K is a semigroup under the following operation

$$(m, q, \alpha) \circ (n, r, \beta) = \left(m + n, \frac{(q\sqrt{1 - 2\beta})^{\frac{n}{m+n}} (r\sqrt{1 - 2\alpha})^{\frac{m}{m+n}}}{\sqrt{1 - 4\alpha\beta}}, \frac{\alpha + \beta - 4\alpha\beta}{1 - 4\alpha\beta} \right).$$

These three semigroups arise in a very natural way in the study of generalized Brownian functionals.

2. Generalized Brownian functionals

Let E^* be the nuclear space consisting of tempered distributions on $\mathbb{R}$. Let μ be the standard Gaussian measure on E^*. Being motivated by the Wiener-Ito decomposition of $L^2(E^*)$, we can introduce generalized multiple Wiener integrals [2]. The space $(L^2)^-$ of generalized Brownian functionals is defined to be the orthogonal direct sum of generalized multiple Wiener integrals [1,4]. For ϕ in $(L^2)^-$, its U-functional is defined by

$$U[\phi](\xi) = e^{-\|\xi\|^2/2} \int_{E^*} e^{\langle x, \xi \rangle} \phi(x)\, d\mu(x), \quad \xi \in E.$$

where $\|\cdot\|$ is the $L^2(\mathbb{R})$ norm and E is the nuclear space of rapidly decreasing smooth functions on $\mathbb{R}$.

Example 1 [5]. The Gaussian Brownian functional with parameter α,

$$\phi_\alpha =: exp[\alpha \int_T \dot{B}(u)^2\, d\mu]: , 0 \le \alpha < \frac{1}{2}.$$

It U-functional is given by

$$U[\phi_\alpha](\xi) = exp\left[\frac{\alpha}{1-2\alpha}\int_T \xi(u)^2\, du\right].$$

Example 2. The Gaussian Brownian functional with parameters a and α,

$$\phi_{(a,\alpha)} =: exp[a\dot{B}(t) + \alpha \int_T \dot{B}(u)^2\, d\mu]: , a \in \mathbb{R}, 0 \le \alpha < \frac{1}{2}.$$

Its U-functional is given by

$$U[\phi_{(a,\alpha)}](\xi) = exp\left[\frac{a}{1-2\alpha}\xi(t) + \frac{\alpha}{1-2\alpha}\int_T \xi(u)^2\, du\right].$$

Example 3 [6]. The Hermite Brownian functional with parameters n, r, and α,

$$\phi_{(n,r,\alpha)} = n!: H_n\left(r\sqrt{1-2\alpha}\dot{B}(t); \frac{r^2}{dt}\right) exp[\alpha \int_T \dot{B}(u)^2\, du]: ,$$

$n \ge 0$, $r \ne 0$, $0 \le \alpha < \frac{1}{2}$. Here $H_n(x;\sigma^2)$ is the Hermite polynomial of degree n with parameter σ^2. Its U-functional is given by

$$U[\phi_{(n,r,\alpha)}](\xi) = \left(\frac{r\xi(t)}{\sqrt{1-2\alpha}}\right)^n exp[\frac{\alpha}{1-2\alpha}\int_T \xi(u)^2\, du].$$

3. Renormalization multiplication

The renormalization multiplication $\phi \circ \psi$ in $(L^2)^-$ has been introduced in [3] by the condition $U[\phi \circ \psi] = U[\phi]U[\psi]$, i.e., the U-functional of $\phi \circ \psi$ is the product of the U-functionals of ϕ and ψ.

Example 1. Consider the Gaussian Brownian fuctionals ϕ_α, $0 \le \alpha < \frac{1}{2}$. We have $\phi_\alpha \circ \phi_\beta = \phi_{\alpha\circ\beta}$. Thus the parameter space $R = [0, \frac{1}{2})$ becomes a semigroup with the operation $\alpha \circ \beta$.

Example 2. For the Gaussian Brownian functionals $\phi_{(a,\alpha)}$, $a \in \mathbb{R}$, $0 \le \alpha < \frac{1}{2}$, we have $\phi_{(a,\alpha)} \circ \phi_{(b,\beta)} = \phi_{(a,\alpha)\circ(b,\beta)}$. Thus the parameter space $J = \mathbb{R} \times R$ is a semigroup with the operation $(a,\alpha) \circ (b,\beta)$.

Example 3. For the Hermite Brownian functionals $\phi_{(n,r,\alpha)}$, $(n,r,\alpha) \in K$, we have $\phi_{(m,q,\alpha)} \circ \phi_{(n,r,\beta)} = \phi_{(m,q,\alpha)\circ(n,r,\beta)}$. Therefore, the parameter space K is a semigroup with the operation $(m,q,\alpha) \circ (n,r,\beta)$.

References

[1] Hida, T., Analysis of Brownian fuctionals, Carleton Math. Lecture Notes, **no. 13**, Carleton Univ., Ottawa, 1975.

[2] Hida, T., *Generalized multiple Wiener integrals*, Proc. Japan Acad. **54A** (1978), 55-58..

[3] Hida, T. and Ikeda, N., *Analysis on Hilbert space with reproducing kernel arising from multiple Wiener integral*, Proc. Fifth Berkeley Symp. Math Stat. Probability, **vol 2** (1965), 117-143..

[4] Kuo, H.-H., *Brownian functionals and applications*, Acta Applicandae Mathematicae **1** (1983), 175-188..

[5] Kuo, H.-H., *A Fourier transform characterization of Gaussian Brownian functionals*, Bulletin Inst. Math., Academia Sinica **11** (1983), 407-413..

[6] Kuo, H.-H., *Fourier-Mehler transforms of generalized Brownian functionals*, Proc. Japan Academy **59A** (1983), 312-314..

Department of Mathematics
Meijo University
Nagoya 468
JAPAN

Department of Mathematics
Louisiana State University
Baton Rouge, LA 70803
USA

Received October 22, 1989

Semigroup Forum Vol. 45 (1992) 264–266

SHORT NOTE

The translational hull of a subsemigroup of a semilattice of groups

Azeeza Ali Ismaeel and György Pollák*

Communicated by L. Márki

It is known that if a semigroup S is embeddable into a group, so is the translational hull $\Omega(S)$ (see e.g. [1], Exercise V.3.10 (vi)). The main aim of this paper is to prove that if a semigroup S is embeddable into a semilattice of groups, so is the translational hull $\Omega(S)$. Thus we partially solve the problem [1], V.3.11 (if a semigroup S is embeddable into an inverse semigroup whether or not $\Omega(S)$ is embeddable into an inverse semigroup).

First we prove two simple propositions.

Proposition 1. *Let S be a semilattice of groups, and T its subsemigroup. Then T is weakly reductive.*

Proof. Let $a, b \in T$ and $ax = bx$, $xa = xb$ for each $x \in T$. Then we have $a^2 = ba = ab = b^2$. Hence $G_a = G_{a^2} = G_{b^2} = G_b$ for the maximal subgroups containing a and b, respectively, and so the subsemigroup generated by a and b is cancellative, which implies $a = b$.

Proposition 2. *Let T be a subsemigroup of a semilattice of groups S. Then T admits a faithful representation by one-to-one partial transformations of T.*

Proof. Let $a \in T \subseteq S$, $w^a : x \to xa$, dom $w^a = Sa^{-1}$. Put $\overline{w}^a = w^a\big|_{T\cap Sa^{-1}}$. So $\overline{w}^a$ is one-to-one since w^a is. In order to prove that $\overline{w}$ is a homomorphism, all we have to show is that $T \cap S(ab)^{-1} = \operatorname{dom} \overline{w}^{ab} = \operatorname{dom}(\overline{w}^a \circ \overline{w}^b) = T \cap (T \cap Sb^{-1})a^{-1}$. Now let $x \in T \cap S(ab)^{-1}$, then $x = xaa^{-1} \in Ta^{-1}$, $xa \in Sb^{-1}$ whence $x \in (T \cap Sb^{-1})a^{-1}$, i.e. $x \in \operatorname{dom}(\overline{w}^a \circ \overline{w}^b)$. Conversely, if $x \in T \cap (T \cap Sb^{-1})a^{-1}$, then $xa \in Sb^{-1}a^{-1}a$, and $xaa^{-1} = x$. Hence $x \in Sb^{-1}a^{-1}$ and $x \in \operatorname{dom} \overline{w}^{ab}$.

Now to prove $\overline{w} : T \to \mathcal{I}(T)$ is one-to-one, where $\mathcal{I}(T)$ is the symmetric inverse semigroup on T, let $a, b \in T$ such that $\overline{w}^a = \overline{w}^b$, i.e. $T \cap Sa^{-1} \cap Sb^{-1}$. Since S is a semilattice of groups, we have $a = a^2a^{-1} \in T \cap Sa^{-1} = T \cap Sb^{-1}$, hence $a = xb^{-1} = xb^{-1}bb^{-1} = abb^{-1}$ for some $x \in S$. This implies $aa^{-1} \leqq bb^{-1}$. By the same argument $bb^{-1} \leqq aa^{-1}$, so $aa^{-1} = bb^{-1}$, $a\overline{w}^a = a\overline{w}^b \to a^2 = ab$, which together imply $a = b$. This proves the proposition.

Our main result is:

Theorem. *Let S be a semilattice of groups, T its subsemigroup. Then the translational hull $\Omega(T)$ is embeddable in a semilattice of groups.*

---

* The authors are deeply indebted to Mrs. Meredith R. Mickel for having taken the trouble of typesetting the manuscript in TEX form.

Remark. It is well known that the translational hull of a semilattice of groups is a semilattice of groups itself (see [2], [3]).

Proof. First we show that $(x\rho)(x\rho)^{-1} \leqq xx^{-1}$ holds for $(\lambda, \rho) \in \Omega(T)$, $x \in T$. Indeed, $x\rho \cdot (x\rho)^{-1} = (x\rho)^2 \cdot (x\rho)^{-2} = x \cdot (\lambda(x\rho)) \cdot (x\rho)^{-2} = xx^{-1} \cdot x \cdot (\lambda(x\rho)) \cdot (x\rho)^{-2} = xx^{-1} \cdot (x\rho)^2 \cdot (x\rho)^{-2} = xx^{-1} \cdot (x\rho) \cdot (x\rho)^{-1}$, whence $x\rho \cdot (x\rho)^{-1} \leqq xx^{-1}$. Similarly, $(\lambda x) \cdot (\lambda x)^{-1} \leqq xx^{-1}$.

Now define $A_\rho = \{x \in T \mid x\rho \cdot (x\rho)^{-1} = xx^{-1}\}$. This set is not empty. Indeed, let $x \in T$. Then

$$\begin{aligned}(x\rho \cdot x)\rho \cdot ((x\rho \cdot x)\rho)^{-1} &= x\rho \cdot x\rho \cdot (x\rho)^{-1} \cdot (x\rho)^{-1} = x\rho \cdot (x\rho)^{-1} \cdot x\rho \cdot (x\rho)^{-1} \\ &= x\rho \cdot (x\rho)^{-1} = x\rho \cdot (x\rho)^{-1} \cdot xx^{-1} \\ &= x\rho \cdot xx^{-1} \cdot (x\rho)^{-1} = x\rho \cdot x(x\rho \cdot x)^{-1} .\end{aligned}$$

Thus, $x\rho \cdot x \in A_\rho$.

Since S is a semilattice of groups, $y \in A_\rho$ means that y an $y\rho$ are contained in the same subgroup G_α of S where $\alpha = yy^{-1}$. Hence $y\rho = y \cdot t_{\alpha,\rho}$ for some $t_{\alpha,\rho} \in G_\alpha$. We shall prove that $x\rho = x \cdot t_{\alpha,\rho}$ for all $x \in G_\alpha \cap A_\rho$. We have

$$(1) \qquad y \cdot \lambda x = y\rho \cdot x = y \cdot t_{\alpha,\rho} \cdot x \in G_\alpha , \text{ hence } (\lambda x)(\lambda x^{-1}) \geqq yy^{-1} = xx^{-1} ;$$

but we have also $\lambda x(\lambda x)^{-1} \leqq xx^{-1}$, so $\lambda x \in G_\alpha$. This means that we can cancel y in (1), and obtain $\lambda x = t_{\alpha,\rho} \cdot x$. Using again the cancellation law in

$$x\rho \cdot x = x \cdot \lambda x = xt_{\alpha,\rho}x ,$$

we get $x\rho = x \cdot t_{\alpha,\rho}$, i.e. $\rho|_{A_\rho \cup G_\alpha} = \rho_{t_{\alpha,\rho}}$. Obviously, $\rho_{t_{\alpha,\rho}}$ is one-to-one.

Define the maps $\overline{\rho}$ and ρ^* on $B_\rho = \bigcup_{A_\rho \cap G_\alpha \neq \emptyset} G_\alpha$ by $x\overline{\rho} = x \cdot t_{\alpha,\rho}$ and $x\rho^* = x \cdot t_{\alpha,\rho}^{-1}$ if $x \in G_\alpha$, respectively. It is clear that $\overline{\rho} \circ \rho^* = \rho^* \circ \overline{\rho} = i_{B_\rho}$. Let $\chi : (\lambda, \rho) \to \overline{\rho}$, $(\lambda, \rho) \in \Omega(T)$. Now χ is a homomorphism. It is enough to prove that, if (λ_1, ρ_1), $(\lambda_2, \rho_2) \in \Omega(T)$, then dom $\overline{\rho_1 \circ \rho_2} =$ dom $\overline{\rho}_1 \circ \overline{\rho}_2$. Let $y \in$ dom $\overline{\rho_1 \circ \rho_2} = B_{\rho_1 \circ \rho_2}$. This implies that there exists an idempotent α such that $y \in G_\alpha \cap A_{\rho_1 \circ \rho_2}$, so we have

$$yy^{-1} = (y(\rho_1 \circ \rho_2)) \cdot (y(\rho_1 \circ \rho_2))^{-1} \leqq (y\rho_1) \cdot (y\rho_1)^{-1} \leqq yy^{-1} = \alpha ,$$

hence $y \in A_{\rho_1}$ and $y\rho_1 \in A_{\rho_2}$. This means that $G_\alpha \cap A_{\rho_1} \neq \emptyset$ and $G_\alpha \cap A_{\rho_2} \neq \emptyset$, so $G_\alpha \subseteq$ dom $\overline{\rho}_1$, $G_\alpha \subseteq$ dom $\overline{\rho}_2$. Hence $G_\alpha \overline{\rho}_1 = G_\alpha$ and $G_\alpha \subseteqq$ dom $\overline{\rho}_1 \circ \overline{\rho}_2$, which implies dom $\overline{\rho_1 \circ \rho_2} \subseteq$ dom $\overline{\rho}_1 \circ \overline{\rho}_2$. It is clear that dom $\overline{\rho}_1 \circ \overline{\rho}_2 \subseteq$ dom $\overline{\rho_1 \circ \rho_2}$, so χ is a homomorphism.

Now in order to prove χ is one-to-one, let (λ_1, ρ_1), $(\lambda_2, \rho_2) \in \Omega(T)$ such that $\overline{\rho}_1 = \overline{\rho}_2$. This means that $G_\alpha \cap A_{\rho_1} \neq \emptyset$ if and only if $A_{\rho_2} \cap G_\alpha \neq \emptyset$. Let $a \in G_\alpha \cap A_{\rho_1}$ and $b \in A_{\rho_2} \cap G_\alpha$. Then

$$a\rho_1 = a \cdot t_{\alpha,\rho_1} = a\overline{\rho}_1 = a\overline{\rho}_2 = a \cdot t_{\alpha,\rho_2} .$$

Hence $t_{\alpha,\rho_1} = t_{\alpha,\rho_2} = t$ and $(a\rho_2) \cdot b = a \cdot (\lambda_2 b) = a \cdot tb$, so

$$(2) \qquad a\rho_2 = a \cdot t = a\rho_1 .$$

Now let $x \in T$,

$$x\rho_1 = x\rho_1 \cdot (x\rho_1)^{-1} \cdot x\rho_1 = (x\rho_1)^{-1} \cdot x\rho_1 \cdot x\rho_1$$
$$= (x\rho_1)^{-1} \cdot (x\rho_1 \cdot x)\rho_1 \ .$$

Moreover, we have $x\rho_1 \cdot x \in A_{\rho_1}$, so $(x\rho_1 \cdot x)\rho_1 = (x\rho_1 \cdot x)\rho_2$ by (2). Hence $x\rho_1 = (x\rho_1)^{-1} \cdot (x\rho_1 \cdot x)\rho_2 = (x\rho_1)^{-1} \cdot (x\rho_1) \cdot x\rho_2$, which implies $x\rho_1 \leqq x\rho_2$. By symmetry, $x\rho_2 \leqq x\rho_1$, which yield $\rho_1 = \rho_2$. This implies also $\lambda_1 = \lambda_2$ because $x \cdot \lambda_1 x = x\rho_1 \cdot x = x\rho_2 \cdot x = x \cdot \lambda_2 x$ and by virtue of $\lambda_i x \cdot (\lambda_i x)^{-1} \leqq xx^{-1}$ $(i = 1, 2)$ we can cancel x. Thus $(\lambda_1, \rho_1) = (\lambda_2, \rho_2)$, which proves the theorem.

References

[1] Petrich, M., *Inverse semigroups*, John Wiley & Sons, Inc., New York, 1984.

[2] Petrich, M., *L'enveloppe de translation d'un demi-treillis de groupes*, Canad. J. Math. **25** (1973), 164–177.

[3] Schein, B. M., *Completions, translational hulls and ideal extensions of inverse semigroups*, Czechoslovak Math. J. **23** (1973), 575–610.

Bolyai Intézet
Szeged
Aradi vértanuk tere 1.
H-6722

Mathematical Research Institute
Budapest
Pf. 127
H-1364

Received July 4, 1990
and in final form February 8, 1991

Semigroup Forum Vol. 45 (1992) 267–271

RESEARCH ARTICLE

Subdirectly irreducible completely symmetrical semigroups

Attila Nagy

Communicated by M.Petrich

A semigroup S is said to be completely symmetrical if $axb = bxa$ for all $a, b, x \in S$. The purpose of this paper is to determine the subdirectly irreducible completely symmetrical semigroups.

1. Completely symmetrical semigroups

In [7], Putcha characterized semigroups which are decomposable as a semilattice of archimedean semigroups. He proved that a semigroup S is a semilattice of archimedean semigroups if and only if, for every $a, b \in S$, the assumption $b \in S^1aS^1$ implies that there is a positive integer i such that $b^i \in S^1a^2S^1$. By this result, the following lemma can be proved easily.

Lemma 1.1. *Every completely symmetrical semigroup is decomposable as a semilattice of archimedean completely symmetrical semigroups.* ∎

The reader can prove easily that every completely symmetrical semigroup is an $E-2$ semigroup ($(xy)^2 = x^2y^2$ for all $x, y \in S$). Thus, using Proposition 1.3, Proposition 1.4 and Proposition 1.6 of [5], it follows that

Lemma 1.2. *A semigroup is completely symmetrical and 0-simple if and only if it is a commutative group with 0-adjoined.* ∎

It can be easily verified that every completely symmetrical semigroup is a medial semigroup ($axyb = ayxb$ for all $x, y, a, b \in S$). In [1], it has been proved that a medial semigroup is archimedean and contains an idempotent element if and only if it is an ideal extension of a rectangular Abelian group by a medial nil semigroup. Using also this result, we can prove the following.

Lemma 1.3. *A semigroup is completely symmetrical archimedean and contains an idempotent element if and only if it is an ideal extension of a commutative group by a completely symmetrical nil semigroup.*

Proof. Let S be a completely symmetrical archimedean semigroup with an idempotent element. As S is also a medial semigroup, it is an ideal extension of a rectangular Abelian group K by a (medial) nil semigroup Q. As K is simple and completely symmetrical, it is a commutative group (Lemma 1.2). It is evident that Q is also a completely symmetrical semigroup.

Conversely, assume that the semigroup S is an ideal extension of a commutative group K by a completely symmetrical nil semigroup Q. Then S is an archimedean semigroup with an idempotent element. We show that S is completely symmetrical. Let a, b, x be arbitrary elements of S. If axb, $bxa \notin K$ then $axb = bxa$, because Q is completely symmetrical. If $axb \in K$ and $bxa \notin K$, then $a, x, b \notin K$ and so $axb = 0$, $bxa \neq 0$ (0 is the zero of

Q) which contradicts the assumption that Q is completely symmetrical. We get a similar result if consider case $bxa \in K$, $axb \notin K$. If axb, $bxa \in K$, then (denoting the identity element of K by 1)

$$axb = 1(axb) = (1a)xb = ((1a)1x)1b = (1a)(1x)(1b)$$

$$= (1b)(1x)(1a) = 1(bxa) = bxa,$$

because K is commutative. Consequently S is completely symmetrical. ■

2. Subdirectly irreducible completely symmetrical semigroups

As is known, a semigroup S is said to be a *subdirect product of semigroups* $\{S_\alpha\}_{\alpha\in A}$ if S is isomorphic to a subsemigroup T of the direct product $\Pi_{\alpha\in A}S_\alpha$ for which $\pi_\alpha(T) = S_\alpha$ for all $\alpha \in A$, where π_α is the projection homomorphism of $\Pi_{\alpha\in A}S_\alpha$ to S_α. We say that a semigroup S is *subdirectly irreducible* if whenever S is written as a subdirect product of a family of semigroups $\{S_\alpha\}_{\alpha\in A}$, then for at least one $\beta \in A$, the projection homomorphism π_β maps S onto S_β isomorphically. It is well known that every semigroup is a subdirect product of subdirectly irreducible semigroups. A semigroup which is not subdirectly irreducible is called *subdirectly reducible.*

Semigroups S and S^0 are simultaneously subdirectly irreducible or reducible ([8]). The least ideal of a semigroup S (if it exists) is called the *kernel* of S. The kernel of a semigroup with zero is trivial. We call an ideal a non-trivial ideal if it contains at least two elements. The least non-trivial ideal of a semigroup S (if it exists) is called the *core* of S. If K is the core of a semigroup S, then either $K^2 = K$ or $K^2 = \{0\}$, where 0 denotes the zero of S. In the first case we call K *globally idempotent,* in the second case K is called *nilpotent.* By [8], every non-trivial subdirectly irreducible semigroup has a core.

A semigroup is called a *homogroup* if it contains a kernel which is a group. By [8], every subdirectly irreducible homogroup without zero is a group. We note that, by [8], a commutative group is subdirectly irreducible if and only if it is a subgroup of a quasicyclic p-group ([3]), p is a prime.

Theorem 2.1. *A semigroup is subdirectly irreducible completely symmetrical with globally idempotent core if and only if it is isomorphic to either G or G^0 or F, where G is a non-trivial subgroup of a quasicyclic p-group, p is a prime, and F is a two-element semilattice.*

Proof. Let S be a subdirectly irreducible completely symmetrical semigroup with globally idempotent core K. Assume that S has no zero element. Then K is simple. As K is completely symmetrical, it follows that K is a commutative group (Lemma 1.2). So S is a subdirectly irreducible homogroup without zero. Then S is a subdirectly irreducible commutative group. Consequently S is isomorphic to a non-trivial subgroup of a quasicyclic p-group, p is a prime. Consider the case when S has a zero element. Then K is 0-simple. By Lemma 1.1, S is a semilattice of archimedean completely symmetrical semigroups. Let S_0 denote the semilattice component of S containing the zero 0 of S. There are two cases.

Assume $|\, S_0 \,| = 1$. We show that $S - \{0\}$ is a subsemigroup of S. Assume, on the contrary, that $ab = 0$ for some $a, b \in S - \{0\}$. Then

$$X = \{x \in S : xb = 0\}$$

is a non-trivial ideal of S, because S_0 is a semilattice component of S. Thus $K \subseteq X$ and so $Kb = \{0\}$. We can prove, in a similar way, that

$$Y = \{y \in S : Ky = \{0\}\}$$

is a non-trivial ideal of S. Thus $K \subseteq Y$ and so $K^2 = \{0\}$, contradicting the assumption that K is globally idempotent. So $S - \{0\}$ is a subsemigroup of S.

Consider the case when $| S_0 |> 1$. As S_0 is a non-trivial ideal of S, $K \subseteq S_0$. By Lemma 1.2, K is a commutative group with 0-adjoined. Then, for every non-zero element k of K and every positive integer n, $k^n \neq 0$, contradicting the fact that S_0 is an archimedean semigroup. Consequently $S - \{0\}$ is a subdirectly irreducible completely symmetrical semigroup. If $S - \{0\}$ is a one-element semigroup, then S is a two-element semilattice.

Assume that $S - \{0\}$ has at least two elements. It is evident that $S - \{0\}$ has no zero element. Assume, on the contrary, that $S - \{0\}$ has a zero element $0'$. Then $\{0, 0'\}$ is an ideal of S. Let α denote the congruence on S defined by the partition $S - \{0\}$, $\{0\}$ of S and let β denote the Rees congruence on S defined by the ideal $\{0, 0'\}$. Then $\alpha \cap \beta = id_S$ (the equality relation on S). As S is subdirectly irreducible, $\alpha = id_S$ or $\beta = id_S$. If $\alpha = id_S$, then $S - \{0\}$ is a one-element semigroup which is a contradiction. If $\beta = id_S$, then we have $0 = 0'$ which is also a contradiction. Consequently $S - \{0\}$ is a subdirectly irreducible completely symmetrical semigroup with globally idempotent core. So it is isomorphic to a non-trivial subgroup of a quasicyclic p-group G, p is a prime. So S is isomorphic to G^0. Thus the direct part of the theorem is proved.

As G, G^0 and F are subdirectly irreducible commutative semigroups with globally idempotent core (G is a subgroup of a quasicyclic p-group, p is a prime, and F is a two-element semilattice), the theorem is proved. ∎

Next we deal with subdirectly irreducible completely symmetrical semigroups with nilpotent core. Before formulating the next theorem, consider some definitions which will be needed later. If S is a semigroup with zero 0, then let A_S denote the *annihilator* of S, that is $A_S = \{a \in S : (\forall s \in S)\ as = sa = 0\}$. It is evident that A_S is an ideal of S. Following [8], an element s of a semigroup S is called a *disjunctive element* if the congruence $C_{\{s\}}$ on S defined by

$$C_{\{s\}} = \{(a,b) \in S \times S : (\forall x, y \in S^1)\ xay = s \ \Leftrightarrow\ xby = s\}$$

equals id_S.

Theorem 2.2. *A completely symmetrical semigroup S with zero and $| A_S |> 1$ is subdirectly irreducible if and only if it has a non-zero disjunctive element.*

Proof. Let S be a subdirectly irreducible completely symmetrical semigroup with zero 0 such that $| A_S |> 1$. Denoting the core of S by K, $K \subseteq A_S$. So $K^2 = \{0\}$. Let a and b be arbitrary non-zero elements in A_S. Then the subsets $\{a, 0\}$ and $\{b, 0\}$ are ideals of S. As S is subdirectly irreducible, $a = b$. Consequently $| A_S |= 2$ and so $| K |= 2$. Let k denote the non-zero element of K. Since $\{k\}$ is a $C_{\{k\}}$-class of S, it follows that $\rho_K \cap C_{\{k\}} = id_S$, where ρ_K denotes the Rees congruence on S determined by the ideal K. As S is subdirectly irreducible and $| K |= 2$, it follows that $C_{\{k\}} = id_S$, that is k is a disjunctive element of S.

Conversely, assume that S is a completely symmetrical semigroup with zero such that $| A_S |> 1$ and S has a non-zero disjunctive element k. Since $r(k) = \{s \in S : (\forall x, y \in S^1) xsy \neq k\}$ is a $C_{\{k\}}$-class and an ideal of S, it follows that $r(k) = \{0\}$ (because k is disjunctive). Let I be an arbitrary non-trivial ideal of S. Then, for every non-zero element a of I, there are elements $x, y \in S^1$ such that $xay = k$ (because $r(k) = \{0\}$). So $k \in I$. Consequently S has a core K and $k \in K$.

Let g be an arbitrary element of A_S. Then, for all $u \in S$, $ug = gu = 0$. So $g \in r(k) \cup \{k\}$. As $r(k) = \{0\}$, it follows that $g \in K$. Consequently $A_S \subseteq K$ and so $A_S = K$. Let a be an arbitrary non-zero element of A_S. Since $\{a, 0\}$ is an ideal of S, it follows that $A_S = K \subseteq \{a, 0\} \subseteq A_S$ which implies that K has exactly two elements (that is K is primitive [8]). By Corollary 3.7.1 of [8], it remains to show that the zero of S is disjunctive. Let e and f be arbitrary elements of S with $e \neq f$. As k is a disjunctive element of S, we have $(e, f) \notin C_{\{k\}}$. So there are elements $x, y \in S^1$ such that, for example, $xey = k$ and $xfy \neq k$. If $xfy = 0$, then $(e, f) \notin C_{\{0\}}$. If $xfy \neq 0$, then there are elements $u, v \in S^1$ such that $uxfyv = k$ (using $r(k) = \{0\}$). As $xfy \neq k$, we have $u \neq 1$ or $v \neq 1$. So $ukv = 0$ (because $k \in A_S$) from which we get $uxeyv = ukv = 0$. This and $uxfyv = k$ (see above) together imply that $(e, f) \notin C_{\{0\}}$. Consequently $(e, f) \notin C_{\{0\}}$ for any elements $e, f \in S$ with $e \neq f$. So the zero of S is disjunctive. Thus the theorem is proved. ■

Example. Let S be a semigroup defined by the following Cayley-table:

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

It can be easily verified that S is completely symmetrical. The element a is the zero of S and $A_S = \{a, b\}$. Moreover b is a non-zero disjunctive element. By Theorem 2.2, S is subdirectly irreducible.

Theorem 2.3. *If S is a subdirectly irreducible completely symmetrical semigroup with zero such that $|A_S| = 1$ and the core of S is nilpotent, then S is commutative (and so is described by Schein in [8]).*

Proof. Let S be a completely symmetrical subdirectly irreducible semigroup such that $|A_S| = 1$ and the core K of S is nilpotent. Let $k \neq 0$ be an arbitrary element of K. As $k \notin A_S$, $Sk \cup kS \cup SkS$ is a non-trivial ideal of S. So $Sk \cup kS \cup SkS = K$. Thus $k = e_1 k$ or $k = ke_2$ or $k = e_3 k e_4$ for some $e_1, e_2, e_3, e_4 \in S$. We show that the third case implies the second case. Assume that $k = e_3 k e_4$ for some $e_3, e_4 \in S$. Then $e_3 k = e_3 e_3 k e_4 = e_4 k e_3 e_3 = e_3 k e_4 e_3 = ke_3$ and so $k = e_3 k e_4 = k e_3 e_4$. Consequently we may consider only the first two cases.

Assume $k = e_1 k$ for some $e_1 \in S$. We show that

$$Z = \{z \in S : z = e_1 z\}$$

is an ideal of S. As $k = e_1 k$, Z contains at least two elements of S. It is evident that Z is a right ideal. Let $z \in Z$ and $s \in S$ be arbitrary elements. Then $e_1 z = z$. As $e_1 sz = e_1 s e_1 z = e_1 z s e_1 = sz e_1 e_1 = s e_1 e_1 z = sz$, we get $sz \in Z$. So Z is also a left ideal of S. Thus Z is a non-trivial ideal of S. Then

$K \subseteq Z$, that is $e_1k = k$ for all $k \in K$. Let α denote the relation on S defined as follows:

$$\alpha = \{(a,b) \in S \times S : (\exists n, m \in N)\ e_1^n a = e_1^m b\},$$

where N is the set of all positive integers. It can be easily verified that α is a right congruence on S. We show that α is also left compatible. Let $(a,b) \in \alpha$ and $x \in S$ be arbitrary elements ($a, b \in S$). Then $e_1^n a = e_1^m b$ for some $n, m \in N$. So

$$e_1^{n+1} xa = e_1^n e_1 xa = e_1^n axe_1 = e_1^m bxe_1 = e_1^{m+1} xb,$$

that is $(xa, xb) \in \alpha$. Consequently α is left compatible and so it is a congruence on S. Since $e_1k = k$ for all $k \in K$ (see above), the restriction of α to K equals id_K. As S is subdirectly irreducible, K is a dense ideal ([6]) of S. Thus $\alpha = id_S$. So $e_1e_1^2 = e_1^2e_1$ (that is $(e_1^2, e_1) \in \alpha$) implies $e_1^2 = e_1$. Consequently, for all $s \in S$,

$$e_1s = e_1e_1s = se_1e_1 = se_1.$$

As $e_1s = e_1e_1s$ (that is $(s, e_1s) \in \alpha$) for all $s \in S$, we get $e_1s = s$ for all $s \in S$. Consequently, for all $s \in S$, $e_1s = se_1 = s$ which means that e_1 is a two-sided identity element of S. So, for all $s, t \in S$, $st = se_1t = te_1s = ts$, because S is completely symmetrical. Thus S is commutative.

In case $k = ke_2$, $e_2 \in S$, we can prove the commutativity of S in a similar way. Thus the theorem is proved. ■

References

[1] Chrislock, J.L., On medial semigroups, Journal of Algebra **12** (1969), 1–9.

[2] Clifford, A. H. and G. B. Preston, The Algebraic Theory of Semigroups, Amer. Math. Soc., Providence, Volume I (1961), Volume II (1967).

[3] Kurosh, A. G., The Theory of Groups, Chelsea, New York, 1960.

[4] Nagy, A., The least separative congruence on a completely symmetrical semigroup, Notes on Semigroups VI., Dept. of Math. Karl Marx Univ. of Economics, Budapest, 1980-4, 13–17.

[5] Nordahl, T., Semigroups satisfying $(xy)^m = x^m y^m$, Semigroup Forum **8** (1974), 332–346.

[6] Petrich, M., Lectures in Semigroups, Akademie-Verlag, Berlin, 1977.

[7] Putcha, M. S., Semilattice decomposition of semigroups, Semigroup Forum **6** (1973), 12-34.

[8] Schein, B.M ., Homomorphisms and subdirect decompositions of semigroups, Pacific Journal of Mathematics, **17** (1966), 529–547.

Department of Mathematics
Transport Engineering Faculty
Technical Univ. of Budapest
1111 Budapest, Egry J. u. 20-22
HUNGARY

Received May 5, 1990
and in final form September 27, 1991

Semigroup Forum Vol. 45 (1992) 272–282

RESEARCH ARTICLE

On the ranks of certain semigroups of order-preserving transformations

Gracinda M. S. Gomes* and John M. Howie*

Communicated by R. B. McFadden

1. Introduction

Let $X_n = \{1, 2, \ldots n\}$, let T_n be the full transformation semigroup on X_n, and let $\mathrm{Sing}_n = \{\alpha \in T_n : |\operatorname{im}\alpha| \leq n-1\}$ be the semigroup of all *singular* selfmaps of X_n. Let

$$\mathcal{O}_n = \{\alpha \in \mathrm{Sing}_n : (\forall x, y \in X_n)\, x \leq y \Rightarrow x\alpha \leq y\alpha\} \tag{1.1}$$

be the subsemigroup of Sing_n consisting of all *order-preserving* singular selfmaps of X_n. This last semigroup was studied in [3], where it was shown that $|\mathcal{O}_n| = \binom{2n-1}{n-1} - 1$. More interestingly, it was shown that E, the set of idempotents of $\mathcal{O}_n$, has cardinality $f_{2n} - 1$, where f_{2n} is the $2n$th Fibonacci number. The main algebraic result of [3] is that $\mathcal{O}_n$ is generated by the set E_1 of idempotents of defect 1. (The *defect* of an element α of T_n is defined as $n - |\operatorname{im}\alpha|$.)

As usual, the *rank* of a finite semigroup S is defined by

$$\operatorname{rank} S = \min\{|A| : A \subseteq S,\ \langle A \rangle = S\}.$$

If S is generated by its set E of idempotents, then the *idempotent rank* of S is defined by

$$\operatorname{idrank} S = \min\{|A| : A \subseteq E,\ \langle A \rangle = S\}.$$

In this paper we investigate the rank and the idempotent rank of $\mathcal{O}_n$. We also look at the larger semigroup

$$\mathcal{PO}_n = \mathcal{O}_n \cup \{\alpha : \operatorname{dom}\alpha \subset X_n,\ (\forall x, y \in \operatorname{dom}\alpha)\, x \leq y \Rightarrow x\alpha \leq y\alpha\} \tag{1.2}$$

of all *partial* order-preserving transformations of X_n (excluding the identity map), and obtain values for its rank and its idempotent rank. In the final section we turn our attention to the semigroup $\mathcal{SPO}_n = \mathcal{PO}_n \backslash \mathcal{O}_n$ of *strictly partial* order-preserving maps of X_n. This semigroup is not idempotent-generated and so the question of its idempotent rank does not arise. We show that its rank is $2n-2$.

Related questions have been considered in recent years. Gomes and Howie [2] investigated Sing_n and found that the rank and the idempotent rank are equal to $n(n-1)/2$. This was generalized by Howie and McFadden [5], who considered the semigroup $K(n,r) = \{\alpha \in T_n : |\operatorname{im}\alpha| \leq r\}$ $(2 \leq r \leq n-1)$ and showed that both the rank and the idempotent rank are equal to $S(n,r)$,

* The authors acknowledge financial support from the Junta Nacional de Investigação Científica e Tecnológia and the Instituto Nacional de Investigação Científica

the Stirling number of the second kind. Garba [1] considered the semigroup P_n of all partial transformations of X_n and showed that for the semigroup $K'(n,r) = \{\alpha \in P_n : |\operatorname{im}\alpha| \le r\}$ both the rank and the idempotent rank are equal to $S(n+1, r+1)$. Gomes and Howie [2] examined the symmetric inverse semigroup I_n $(= \mathcal{I}(X_n))$ consisting of all partial one-one maps of X_n and showed that the rank (as an inverse semigroup) of the inverse semigroup $SP_n = \{\alpha \in I_n : |\operatorname{im}\alpha| \le n-1\}$ is $n+1$. Garba generalized this by showing that for $r = 3, \ldots, n-1$ the rank of $L(n,r) = \{\alpha \in I_n : |\operatorname{im}\alpha| \le r\}$ is $\binom{n}{r} + 1$.

2. Order-preserving transformations

It is clear that $\mathcal{O}_n$ is a regular subsemigroup of T_n. Hence by [4, Proposition II.4.5 and Exercise II.10] we have that in $\mathcal{O}_n$

$$\alpha\,\mathcal{L}\beta \text{ if and only if } \operatorname{im}\alpha = \operatorname{im}\beta,$$

$$\alpha\mathcal{R}\beta \text{ if and only if } \ker\alpha = \ker\beta,$$

$$\alpha\mathcal{J}\beta \text{ if and only if } |\operatorname{im}\alpha| = |\operatorname{im}\beta|.$$

Thus $\mathcal{O}_n$, like T_n itself, is the union of $\mathcal{J}$-classes $J_1, J_2, \ldots, J_{n-1}$, where $J_r = \{\alpha \in \mathcal{O}_n : |\operatorname{im}\alpha| = r\}$. We shall want to pay particular attention to the $\mathcal{J}$-class J_{n-1} at the top of the semigroup.

By contrast with Sing_n, the semigroup $\mathcal{O}_n$ is *aperiodic* (i.e., has trivial $\mathcal{H}$-classes); for once we fix $\operatorname{im}\alpha$ and $\ker\alpha$ there is precisely one order-preserving map having the given image and kernel. It is easy to see that the $(\ker\alpha)$-classes are *convex* subsets C of X_n, in the sense that $x, y \in C$ and $x \le z \le y$ together imply that $z \in C$. Thus, denoting by $|i,j|$ the equivalence on X_n whose sole non-singleton class is $\{i,j\}$, we see that the $\mathcal{R}$-classes in the $\mathcal{J}$-class J_{n-1} are indexed by the $n-1$ equivalences

$$|1,2|,\ |2,3|,\ \ldots,\ |n-1,n|.$$

The $\mathcal{L}$-classes correspond to the n possible images

$$X_n\backslash\{1\},\ X_n\backslash\{2\},\ \ldots,\ X_n\backslash\{n\}.$$

Thus $|J_{n-1}| = n(n-1)$. The (unique) element in the $\mathcal{H}$-class determined by the equivalence $|i,i+1|$ and the subset $X_n\backslash\{k\}$ is idempotent if and only if $X_n\backslash\{k\}$ is a transversal (or cross-section) of $|i,i+1|$, i.e., if and only if $k = i$ or $k = i+1$.

The element with kernel $|i,i+1|$ and image $X_n\backslash\{k\}$ is

$$\alpha = \begin{pmatrix} 1 & \ldots & i & i+1 & \ldots & k & k+1 & \ldots & n \\ 1 & \ldots & i & i & \ldots & k-1 & k+1 & \ldots & n \end{pmatrix}$$

if $k \ge i+1$, and is

$$\alpha = \begin{pmatrix} 1 & \ldots & k-1 & k & \ldots & i & i+1 & \ldots & n \\ 1 & \ldots & k-1 & k+1 & \ldots & i+1 & i+1 & \ldots & n \end{pmatrix}$$

if $k \le i$. In the former case we say that α is a *decreasing* element and write

$$\alpha = [k \to k-1 \to \cdots \to i];$$

in the latter case we say that α is an *increasing* element and write

$$\alpha = [k \to k+1 \to \cdots \to i+1].$$

There are $n-1$ decreasing idempotents $[i \to i-1]$ $(i = 2, 3, \ldots, n)$ and $n-1$ increasing idempotents $[i \to i+1]$ $(i = 1, 2, \ldots, n-1)$. Thus if

$$E_1 = \{\alpha \in \mathcal{O}_n : |\operatorname{im}\alpha| = n-1,\ \alpha^2 = \alpha\} \tag{2.1}$$

is the set of all idempotents in J_{n-1}, then $|E_1| = 2n-2$.

From the result in [3] that $\mathcal{O}_n = \langle E_1 \rangle$ we immediately deduce

Lemma 2.2. *Let $\mathcal{O}_n$ be as defined in* (1.1). *Then* $\operatorname{rank}\mathcal{O}_n \leq 2n-2$ *and* $\operatorname{idrank}\mathcal{O}_n \leq 2n-2$. ∎

From Lemma 1 in [5] we deduce that the rank of $\mathcal{O}_n$ must be at least as large as the number of $\mathcal{L}$-classes in J_{n-1}. Thus we have

Lemma 2.3. $\operatorname{rank}\mathcal{O}_n \geq n$. ∎

For $i = 1, 2, \ldots, n-1$ let us denote the increasing idempotent $[i \to i+1]$ by ϵ_i. Denote the decreasing (non-idempotent) element $[n \to n-1 \to \cdots \to 1]$ of $\mathcal{O}_n$ by γ, and let

$$A = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_{n-1}, \gamma\}. \tag{2.4}$$

Then we have

Lemma 2.5. *All increasing elements of J_{n-1} belong to $\langle A \rangle$.*

Proof. It is easy to verify that

$$[k \to k+1 \to \cdots \to i+1] = [i \to i+1][i-1 \to i]\ldots[k \to k+1],$$

a product of increasing idempotents. ∎

Next, we have

Lemma 2.6. *All decreasing idempotents in J_{n-1} belong to $\langle A \rangle$.*

Proof. If we multiply the element $\gamma = [n \to n-1 \to \cdots \to 1]$ by the product $\epsilon_{n-1}\epsilon_{n-2}\ldots\epsilon_i$ of increasing idempotents we obtain the element

$$[i \to i-1 \to \cdots \to 1].$$

Then if we multiply this by the increasing element $[1 \to 2 \to \cdots \to i-1]$ (which is in $\langle A \rangle$ by Lemma 2.5) we obtain the decreasing idempotent $[i \to i-1]$. ∎

We may now use the argument of Lemma 2.5 to show that all decreasing elements of J_{n-1} are in $\langle A \rangle$. Thus $J_{n-1} \subseteq \langle A \rangle$, and so

$$\mathcal{O}_n = \langle E_1 \rangle \subseteq \langle J_{n-1} \rangle \subseteq \langle A \rangle \subseteq \mathcal{O}_n.$$

We have shown that $\langle A \rangle = \mathcal{O}_n$, with $|A| = n$. This, together with Lemma 2.3, gives us

Theorem 2.7. *Let $n \geq 2$ and let $\mathcal{O}_n$ be as defined in* (1.1). *Then* $\operatorname{rank}\mathcal{O}_n = n$. ∎

By contrast, we now establish

Theorem 2.8. *Let $n \geq 2$ and let $\mathcal{O}_n$ be as defined in* (1.1). *Then* $\operatorname{idrank}\mathcal{O}_n = 2n-2$.

Proof. Let F be a set of idempotents in J_{n-1} with the property that $\langle F \rangle = \mathcal{O}_n$. Then in particular $J_{n-1} \subseteq \langle F \rangle$. Consider a product

$$\eta_1\eta_2\ldots\eta_m \in J_{n-1}, \tag{2.8}$$

with $\eta_1, \eta_2, \ldots, \eta_m$ in F. Suppose that $\eta_k = [i \to i+1]$, an increasing idempotent, and that the next idempotent η_{k+1} is decreasing. For the product $\eta_k\eta_{k+1}$ to

Theorem 3.1. *$|\mathcal{PO}_n| + 1$ is the coefficient of x^n in the series expansion of*

$$(1+x)^n(1-x)^{-n}. \qquad \blacksquare$$

The number $|\mathcal{PO}_n|$ grows quite rapidly with n. Some values for small n are given in the table:

n	2	3	4	5
$\lvert\mathcal{PO}_n\rvert$	7	37	191	1001

Our next main result is

Theorem 3.2. *Let $n \geq 2$ and let $\mathcal{PO}_n$ be as defined in* (1.2). *Then* $\operatorname{rank}\mathcal{PO}_n = 2n - 1$.

Proof. In $\mathcal{PO}_n$, and indeed in the larger semigroup P_n of *all* partial transformations of X_n, it is convenient to refer to an element α as being *of type* (r, s), or as *belonging to the set* $[r, s]$ if $|\operatorname{dom}\alpha| = r$ and $|\operatorname{im}\alpha| = s$. We must of course have $0 \leq s \leq r \leq n$.

The $\mathcal{J}$-class $J_{n-1} = \{\alpha \in \mathcal{PO}_n : |\operatorname{im}\alpha| = n-1\}$ is the union of $[n, n-1]$ and $[n-1, n-1]$. Within $[n, n-1]$ there are $n - 1$ $\mathcal{R}$-classes, indexed by the equivalences

$$|1,2|,\ |2,3|,\ \ldots,\ |n-1, n|;$$

within $[n-1, n-1]$, which consists of one-one partial maps, there are n $\mathcal{R}$-classes, indexed by the domains $X_n\backslash\{1\}$, $X_n\backslash\{2\}$, $\ldots$, $X_n\backslash\{n\}$. As in the argument leading to Lemma 2.2, we can be sure that a generating set for $\mathcal{PO}_n$ covers the $\mathcal{R}$-classes in J_{n-1}. So we have

Lemma 3.3. $\operatorname{rank}\mathcal{PO}_n \geq 2n - 1$. $\blacksquare$

To show the opposite inequality we begin by establishing

Lemma 3.4. *For $r = 1, 2, \ldots, n-2$,* $\quad [r, r] \subseteq [r+1, r+1][r+1, r+1]$.

Proof. Consider a typical element

$$\alpha = \begin{pmatrix} a_1 & a_2 & \ldots & a_r \\ b_1 & b_2 & \ldots & b_r \end{pmatrix} \tag{3.5}$$

of $[r, r]$, with $a_1 < a_2 < \cdots < a_r$ and $b_1 < b_2 < \cdots < b_r$. The ordered r-tuple $(a_1, a_2, \ldots, a_r)$ will be said to have a *gap in position* i if $a_i \geq a_{i-1} + 2$. By convention, we shall interpret a gap in position 1 as occurring when $a_1 \geq 2$, and a gap in position $r+1$ as occurring when $a_r \leq n - 1$. Similar definitions apply to $(b_1, b_2, \ldots, b_r)$. Since $r \leq n - 2$ there must be at least one gap in each case.

Suppose now that the element α given by (3.5) is such that there is a gap in position i in $(a_1, a_2, \ldots, a_r)$ and a gap in position j in $(b_1, b_2, \ldots, b_r)$. Let x, y be such that $a_{i-1} < x < a_i$, $b_{j-1} < y < b_j$. (Notice that, for example, $x = 1$ if $i = 1$ and $x = n$ if $i = r + 1$.)

We now distinguish two cases, according to whether $i \leq j$ or $i > j$. Suppose first that $i \leq j$. Then let ξ map $\{a_1, a_2, \ldots a_r\} \cup \{x\}$ onto $\{1, 2, \ldots, r+2\}\backslash\{j+1\}$ in an order-preserving way. Thus $\xi \in [r+1, r+1]$ and

$$a_k\xi = \begin{cases} k & \text{if } k \leq i-1, \\ k+1 & \text{if } i \leq k \leq j-1, \\ k+2 & \text{if } k \geq j, \end{cases} \qquad x\xi = i. \tag{3.6}$$

be in J_{n-1} we require that the image $X_n \setminus \{i\}$ of η_k be a transversal of $\ker \eta_{k+1}$. Thus

$$\eta_{k+1} = [i+1 \to i] \text{ or } \eta_{k+1} = [i \to i+1].$$

In the former case $\eta_k \eta_{k+1} = \eta_{k+1}$, while in the latter case $\eta_k \eta_{k+1} = \eta_k$. Similar considerations apply when η_k is decreasing and η_{k+1} is increasing. Thus the product (2.8) can step by step be reduced, and we deduce that every element α in J_{n-1} is expressible as a reduced product $\eta'_1 \eta'_2 \dots \eta'_p$ in which either every η'_i is increasing (if α is increasing) or every η'_i is decreasing (if α is decreasing). Now, by Lemma 1 in [5], if $\alpha = \eta'_1 \eta'_2 \dots \eta'_p$ then $\alpha \, \mathcal{R} \eta'_1$. It follows that each $\mathcal{R}$-class in J_{n-1}, determined (say) by the equivalence $|i, i+1|$, must be represented in F both by the increasing idempotent $[i \to i+1]$ and by the decreasing idempotent $[i+1 \to i]$. Hence $F = E_1$, the set of *all* idempotents in J_{n-1}. The required result now follows by Lemma 2.2. ∎

3. Partial transformations

We begin with some combinatorial ideas which will lead to an expression for the order of the semigroup $\mathcal{PO}_n$. Let us refer to an equivalence ρ on the set X_m as *convex* if its classes are convex subsets of X_m, and let us say that ρ is of *weight* r if $|X_m/\rho| = r$. A convex equivalence of weight r is determined by the insertion of $r-1$ 'boundaries' in the $m-1$ spaces between $1, 2, \dots, m$. (Thus, for example, the convex equivalence of weight 3 on X_6 whose classes are $\{1,2,3\}$, $\{4\}$ and $\{5,6\}$ is determined by inserting two boundaries, between 3 and 4 and between 4 and 5.) We deduce that the number of convex equivalences of weight r on X_m is $\binom{m-1}{r-1}$. Now, for $r = 1, 2, \dots, n$, let $W_r = \{\alpha \in \mathcal{PO}_n : |\operatorname{dom} \alpha| = r\}$. Once $\operatorname{dom} \alpha$ is chosen, $\operatorname{im} \alpha$ is necessarily a subset of X_n with $|\operatorname{im} \alpha| = s \le r$, and $\ker \alpha$ is a convex equivalence on $\operatorname{dom} \alpha$ with weight s. So

$$|W_r| = \binom{n}{r} \sum_{s=1}^{r} \binom{n}{s} \binom{r-1}{s-1}.$$

Now

$$\sum_{s=1}^{r} \binom{n}{s} \binom{r-1}{s-1}$$

is the coefficient of x^r in $(1+x)^n (1+x)^{r-1} = (1+x)^{n+r-1}$; so

$$\sum_{s=1}^{r} \binom{n}{s} \binom{r-1}{s-1} = \binom{n+r-1}{r}.$$

It follows that

$$|\mathcal{PO}_n| + 1 = \sum_{r=1}^{n} \binom{n}{r} \binom{n+r-1}{r}.$$

(The $+1$ on the left arises from the fact that our counting has included the identity map of X_n, which we are excluding from $\mathcal{PO}_n$.)

Finally, from the binomial series

$$(1-x)^{-n} = \sum_{r=0}^{\infty} \binom{n+r-1}{r} x^r \quad (|x| < 1)$$

we obtain

Let η map $\{1,2,\ldots,r+2\}\backslash\{i\}$ onto $\{b_1,b_2,\ldots,b_r\}\cup\{y\}$ in an order-preserving way. Again $\eta\in[r+1,r+1]$, and

$$k\eta=\begin{cases} b_k & \text{if } k\le i-1,\\ b_{k-1} & \text{if } i+1\le k\le j,\\ y & \text{if } k=j+1,\\ b_{k-2} & \text{if } k\ge j+2.\end{cases} \tag{3.7}$$

From (3.6) and (3.7) it is easy to verify that $\xi\eta=\alpha$.

Suppose now that $i>j$. Let $\xi\colon\{a_1,a_2,\ldots a_r\}\cup\{x\}\to\{1,2,\ldots,r+2\}\backslash\{j\}$ be an order-preserving surjection, and let η similarly map $\{1,2,\ldots,r+2\}\backslash\{i+1\}$ onto $\{b_1,b_2,\ldots,b_r\}\cup\{y\}$. Then

$$a_k\xi=\begin{cases} k & \text{if } k\le j-1,\\ k+1 & \text{if } j\le k\le i-1,\\ k+2 & \text{if } k\ge i,\end{cases}$$
$$x\xi=i+1,$$

while

$$k\eta=\begin{cases} b_k & \text{if } k\le j-1,\\ y & \text{if } k=j,\\ b_{k-1} & \text{if } j+1\le k\le i,\\ b_{k-2} & \text{if } k\ge i+2,\end{cases}$$

and again it is easy to verify that $\xi\eta=\alpha$. ∎

Next, we have

Lemma 3.8. *If $1\le s<r\le n-1$ then $[r,s]\subseteq[r,r][n,s]$.*

Proof. Consider an element α in $[r,s]$, with domain $A=\{a_1,a_2,\ldots a_r\}$, where $a_1<a_2<\cdots<a_r$, and denote $a_i\alpha$ by b_i for $i=1,2,\ldots,r$. Define ξ to be the identity on A. Define η as follows: if $x\in A$ let $x\eta=x\alpha$; if $x<a_1$ let $x\eta=b_1$; if $a_i<x<a_{i+1}$ let $x\eta=b_i$; if $x>a_r$ let $x\eta=b_r$. Then $\xi\in[r,r]$, $\eta\in[n,s]$ and $\xi\eta=\alpha$. ∎

By the main result in [3] we have that (for all $s\le n-1$) $[n,s]\subseteq\langle[n,n-1]\rangle$. This, together with Lemmas 3.4 and 3.8 and the observation already made that $J_{n-1}=[n,n-1]\cup[n-1,n-1]$, enables us to deduce

Lemma 3.9. $\langle J_{n-1}\rangle=\mathcal{PO}_n$. ∎

Now, for $i=2,3,\ldots,n$ let us denote by α_i the element of $[n-1,n-1]$ that maps $X_n\backslash\{i\}$ onto $X_n\backslash\{i-1\}$ in an order-preserving way; and let α_1 similarly map $X_n\backslash\{1\}$ onto $X_n\backslash\{n\}$. Thus, for example, if $n=4$ then

$$\alpha_1=\begin{pmatrix}2&3&4\\1&2&3\end{pmatrix},\quad \alpha_2=\begin{pmatrix}1&3&4\\2&3&4\end{pmatrix},$$
$$\alpha_3=\begin{pmatrix}1&2&4\\1&3&4\end{pmatrix},\quad \alpha_4=\begin{pmatrix}1&2&3\\1&2&4\end{pmatrix}.$$

Let

$$B_P=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}\quad B_T=\{\epsilon_1,\epsilon_2,\ldots,\epsilon_{n-1}\}, \tag{3.10}$$

where as before $\epsilon_i=[i\to i+1]\in[n,n-1]$. Let $B=B_P\cup B_T$. We shall show that B, consisting of $2n-1$ elements, is a generating set for $\mathcal{PO}_n$. More precisely, we shall show that $J_{n-1}\subseteq\langle B\rangle$, and this, together with Lemma 3.9, gives the required result. First, we show

Lemma 3.11. $[n, n-1] \subseteq \langle B \rangle$.

Proof. From the proof of Theorem 2.7 we know that A, defined by (2.4), generates $[n, n-1]$. The sets A and B overlap in $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_{n-1}\}$, and the previously considered element

$$\gamma = [n \to n-1 \to \cdots \to 1]$$

of A, being expressible as $\epsilon_1 \alpha_1$, is in $\langle B \rangle$. The result follows. ■

Next, we have

Lemma 3.12. $[n-1, n-1] \subseteq \langle B_P \rangle \subseteq \langle B \rangle$.

Proof. Consider a typical element β in $[n-1, n-1]$, with $\operatorname{dom}\beta = X_n \setminus \{i\}$, $\operatorname{im}\beta = X_n \setminus \{j\}$. If $i > j$ then $\beta = \alpha_i \alpha_{i-1} \ldots \alpha_{j+1}$; if $i \leq j$ then

$$\beta = \alpha_i \alpha_{i-1} \ldots \alpha_1 \alpha_n \ldots \alpha_{j+1}.$$ ■

From Lemmas 3.9, 3.11 and 3.12 we now immediately deduce that B generates $\mathcal{PO}_n$. Thus the proof of Theorem 3.2 is complete. ■

Next, we have

Theorem 3.13. *The semigroup $\mathcal{PO}_n$ is idempotent-generated. Its idempotent rank is 3n-2.*

Proof. Consider the set E_1 defined in (2.1). For $i = 1, 2, \ldots, n$ let δ_i be the identity map of $X_n \setminus \{i\}$, let $D = \{\delta_1, \delta_2, \ldots \delta_n\}$, and let $F = E_1 \cup D$, a set of idempotents with $|F| = 3n-2$. Since $\langle E_1 \rangle = \mathcal{O}_n$ by [3] we certainly have $\mathcal{O}_n \subseteq \langle F \rangle$. Also, for $i = 2, 3, \ldots, n$,

$$\alpha_i = \delta_i \epsilon_{i-1} \in DE_1 \subseteq \langle F \rangle, \quad \alpha_1 = \delta_1 \gamma \in D\mathcal{O}_n \subseteq \langle F \rangle.$$

Thus $B_P \subseteq \langle F \rangle$, and hence by Lemma 3.12 every element of $[n-1, n-1]$ is in $\langle F \rangle$. Hence we have

Lemma 3.14. $\langle F \rangle = \mathcal{PO}_n$. ■

It follows that $\operatorname{idrank} \mathcal{PO}_n \leq 3n-2$. To show that no set of idempotents of smaller cardinality will generate $\mathcal{PO}_n$, notice first that a generating set must cover the n $\mathcal{R}$-classes in $[n-1, n-1]$. The idempotents $\delta_1, \delta_2, \ldots, \delta_n$ do have this property, and in fact are the *only* idempotents in $[n-1, n-1]$. It is clear therefore that any idempotent generating set for $\mathcal{PO}_n$ must contain D.

We saw earlier that no proper subset of E_1 would generate $\mathcal{O}_n$. Suppose now that $\alpha \in [n, n-1]$ and that

$$\alpha = \zeta_1 \zeta_2 \ldots \zeta_k, \tag{3.15}$$

where $\zeta_1, \zeta_2, \ldots, \zeta_k \in F = E_1 \cup D$. We must in fact have $\zeta_1 \in E_1$, since otherwise $|\operatorname{dom}\alpha| = n-1$. Suppose that not all of $\zeta_1, \zeta_2, \ldots, \zeta_k$ are in E_1, and that ζ_j (with $j \geq 2$) is the first ζ not in E_1. Then $\operatorname{dom}\zeta_j = \operatorname{im}(\zeta_1 \zeta_2 \ldots \zeta_{j-1})$, for otherwise

$$|\operatorname{im}\alpha| \leq |\operatorname{im}(\zeta_1 \zeta_2 \ldots \zeta_j)| \leq n-2.$$

In fact ζ_j is then the identity map on $\operatorname{im}(\zeta_1 \zeta_2 \ldots \zeta_{j-1})$ and so

$$\zeta_1 \zeta_2 \ldots \zeta_j = \zeta_1 \zeta_2 \ldots \zeta_{j-1}.$$

Continuing this argument, we see that the elements of D in the product (3.15) are in fact superfluous. Hence the product can be replaced by a reduced product $\alpha = \zeta_1' \zeta_2' \ldots \zeta_l'$ in which $\zeta_1', \zeta_2', \ldots, \zeta_l' \in E_1$. It follows that any generating set consisting of idempotents in J_{n-1} (i.e., any subset of F) must contain enough elements of E_1 to generate all the elements of $\mathcal{O}_n$. As we saw earlier, in the proof of Theorem 2.7, this means that all of E_1 must be present. ■

4. Strictly partial transformations

The subset $\mathcal{SPO}_n = \mathcal{PO}_n \backslash \mathcal{O}_n$ of *strictly partial* order-preserving maps of X_n is a subsemigroup (indeed a right ideal) of $\mathcal{PO}_n$. We shall prove

Theorem 4.1. *Let $\mathcal{SPO}_n$ be the semigroup of strictly partial order-preserving maps of X_n. Then* $\operatorname{rank} \mathcal{SPO}_n = 2n-2$.

Proof. We begin by showing that every generating set G of $\mathcal{SPO}_n$ must contain at least $2n-2$ elements. The top $\mathcal{J}$-class here is $[n-1, n-1]$, and since this consists entirely of one-one maps it is clear that it does not generate $\mathcal{SPO}_n$.

On the other hand, it is clear that in generating the elements of $[n-1, n-1]$ only elements of $[n-1, n-1]$ may be used. The number of generators must be at least as large as the number of $\mathcal{R}$-classes in $[n-1, n-1]$, i.e., as the number of subsets of X_n of cardinality $n-1$. Thus

$$\left|G \cap [n-1, n-1]\right| \geq n. \tag{4.2}$$

Notice also that, since no non-idempotent element in $[n-1, n-1]$ can be expressed as a product of idempotents, the semigroup $\mathcal{SPO}_n$ is not idempotent-generated.

Now consider a typical element

$$\alpha = \begin{pmatrix} a_1 & a_2 & \dots & a_{n-1} \\ b_1 & b_2 & \dots & b_{n-1} \end{pmatrix} \in [n-1, n-2]. \tag{4.3}$$

Here $a_1 < a_2 < \dots < a_{n-1}$, $b_1 \leq b_2 \leq \dots \leq b_{n-1}$. In fact all but one of the inequalities between the b's are strict; we say that α is *of kernel type* i and write $K(\alpha) = i$ if $b_i = b_{i+1}$. The possible values for $K(\alpha)$ are $1, 2, \dots, n-2$.

Suppose, then, that we have an α given by (4.3), and that $K(\alpha) = i$. Suppose now that $\alpha = \gamma_1\gamma_2\dots\gamma_m$, with $\gamma_1, \gamma_2, \dots, \gamma_m \in G$. Then each γ_j is either in $[n-1, n-1]$ or in $[n-1, n-2]$, and

$$\operatorname{dom}\gamma_1 = \operatorname{dom}\alpha, \quad \operatorname{dom}\gamma_r \supseteq \operatorname{im}(\gamma_1\gamma_2\dots\gamma_{r-1}) \tag{4.4}$$

for $r = 2, 3, \dots, m$. At least one γ_j must be in $[n-1, n-2]$. Suppose that γ_t is the *first* such γ; that is, suppose that

$$\gamma_1, \dots, \gamma_{t-1} \in [n-1, n-1], \ \gamma_t \in [n-1, n-2].$$

Then

$$\gamma_1 \dots \gamma_{t-1} = \begin{pmatrix} a_1 & a_2 & \dots & a_{n-1} \\ c_1 & c_2 & \dots & c_{n-1} \end{pmatrix} \in [n-1, n-1],$$

with $c_1 < c_2 < \dots < c_{n-1}$, and by virtue of (4.4) we must have

$$\gamma_t = \begin{pmatrix} c_1 & c_2 & \dots & c_{n-1} \\ d_1 & d_2 & \dots & d_{n-1} \end{pmatrix}.$$

There is exactly one p for which $d_p = d_{p+1}$. But then it follows that

$$a_p\gamma_1 \dots \gamma_t = a_{p+1}\gamma_1 \dots \gamma_t$$

and hence that $a_p\alpha = a_{p+1}\alpha$. Since $K(\alpha) = i$ we must therefore have that $p = i$.

The conclusion of this argument is that $G \cap [n-1, n-2]$ must contain at least one element of each of the $n-2$ possible kernel types. This, together with (4.2), gives us

$$|G| \geq n + (n-2) = 2n - 2.$$

To prove the theorem we now require to produce a set G of generators of $\mathcal{SPO}_n$ such that $|G| = 2n-2$. For $i = 1, 2, \ldots, n-2$ let β_i (belonging to $[n-1, n-2]$) have domain $\{1, 2, \ldots, n-1\}$ and let

$$i\beta_i = i+1,\ j\beta_i = j\ (j \neq i).$$

Let

$$G = B_P \cup \{\beta_1, \beta_2, \ldots, \beta_{n-2}\},$$

where B_P is the subset $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of $[n-1, n-1]$ given by (3.10). Then certainly $|G| = 2n-2$. We shall show that $\langle G \rangle = \mathcal{SPO}_n$.

From Lemma 3.12 we know that

$$[n-1, n-1] \subseteq \langle B_P \rangle \subseteq \langle G \rangle,$$

and from Lemma 3.4 it follows that

$$[r, r] \subseteq \langle G \rangle \tag{4.5}$$

for $r = 1, 2, \ldots, n-1$.

Next, we have

Lemma 4.6. $[n-1, n-2] \subseteq [n-1, n-1]G[n-2, n-2]$.

Proof. Consider a typical element

$$\alpha = \begin{pmatrix} a_1 & \ldots & a_i & a_{i+1} & a_{i+2} & \ldots & a_{n-1} \\ a_1 & \ldots & b_i & b_i & b_{i+1} & \ldots & b_{n-2} \end{pmatrix}$$

of $[n-1, n-2]$, where $a_1 < a_2 < \cdots < a_{n-1}$ and $b_1 < b_2 < \cdots < b_{n-2}$. Let

$$\xi = \begin{pmatrix} a_1 & a_2 & \ldots & a_{n-1} \\ 1 & 2 & \ldots & n-1 \end{pmatrix} \in [n-1, n-1],$$

$$\beta_i = \begin{pmatrix} 1 & \ldots & i-1 & i & i+1 & \ldots & n-1 \\ 1 & \ldots & i-1 & i+1 & i+1 & \ldots & n-1 \end{pmatrix} \in G,$$

$$\eta = \begin{pmatrix} 1 & \ldots & i-1 & i+1 & \ldots & n-1 \\ b_1 & \ldots & b_{i-1} & b_i & \ldots & b_{n-2} \end{pmatrix} \in [n-2, n-2].$$

Then $\alpha = \xi\beta_i\eta$. ∎

Next, we have

Lemma 4.7. *For* $r = 2, 3, \ldots, n-2$,

$$[r, r-1] \subseteq [r+1, r][r-1, r-1] \cup [r+1, r+1][r+1, r].$$

Proof. Consider a typical element

$$\alpha = \begin{pmatrix} a_1 & \ldots & a_i & a_{i+1} & \ldots & a_r \\ b_1 & \ldots & b_i & b_i & \ldots & b_{r-1} \end{pmatrix}$$

of $[r, r-1]$, where $i \leq r-1$. Suppose that $(a_1, a_2, \ldots, a_r)$ has a gap in position j, and let x be such that $a_{j-1} < x < a_j$. We distinguish several cases.

Case 1. $j > i+1$. Let

$$\xi = \begin{pmatrix} a_1 & \ldots & a_i & a_{i+1} & \ldots & a_{j-1} & x & a_j & \ldots & a_r \\ 1 & \ldots & i & i & \ldots & j-2 & j-1 & j & \ldots & r \end{pmatrix},$$

$$\eta = \begin{pmatrix} 1 & \ldots & j-2 & j & \ldots & r \\ b_1 & \ldots & b_{j-2} & b_{j-1} & \ldots & b_{r-1} \end{pmatrix};$$

then $\xi \in [r+1, r]$, $\eta \in [r-1, r-1]$ and $\xi\eta = \alpha$.

Case 2. $j \leq i$. Let

$$\xi = \begin{pmatrix} a_1 & \ldots & a_{j-1} & x & a_j & \ldots & a_i & a_{i+1} & \ldots & a_r \\ 1 & \ldots & j-1 & j & j+1 & \ldots & i+1 & i+1 & \ldots & r \end{pmatrix},$$

$$\eta = \begin{pmatrix} 1 & \ldots & j-1 & j+1 & \ldots & r \\ b_1 & \ldots & b_{j-1} & b_j & \ldots & b_{r-1} \end{pmatrix};$$

then $\xi \in [r+1, r]$, $\eta \in [r-1, r-1]$ and $\xi\eta = \alpha$.

Case 3. $j = i+1$. Suppose that $(b_1, b_2, \ldots, b_{r-1})$ has a gap in position l, and let y be such that $b_{l-1} < y < b_l$. If $i < l$, let

$$\xi = \begin{pmatrix} a_1 & \ldots & a_i & x & a_{i+1} & \ldots & a_l & a_{l+1} & \ldots & a_r \\ 1 & \ldots & i & i+1 & i+2 & \ldots & l+1 & l+3 & \ldots & r+2 \end{pmatrix},$$

$$\eta = \begin{pmatrix} 1 & \ldots & i & i+2 & \ldots & l+1 & l+2 & l+3 & \ldots & r+2 \\ b_1 & \ldots & b_i & b_i & \ldots & b_{l-1} & y & b_l & \ldots & b_{r-1} \end{pmatrix};$$

then $\xi \in [r+1, r+1]$, $\eta \in [r+1, r]$, and $\xi\eta = \alpha$. If $i \geq l$, let

$$\xi = \begin{pmatrix} a_1 & \ldots & a_{l-1} & a_l & \ldots & a_i & x & a_{i+1} & \ldots & a_r \\ 1 & \ldots & l-1 & l+1 & \ldots & i+1 & i+2 & i+3 & \ldots & r+2 \end{pmatrix},$$

$$\eta = \begin{pmatrix} 1 & \ldots & l-1 & l & l+1 & \ldots & i+1 & i+3 & \ldots & r+2 \\ b_1 & \ldots & b_{l-1} & y & b_l & \ldots & b_i & b_i & \ldots & b_{r-1} \end{pmatrix};$$

then $\xi \in [r+1, r+1]$, $\eta \in [r+1, r]$, and $\xi\eta = \alpha$. ∎

Our final lemma is

Lemma 4.8. *For* $2 \leq r \leq n-1$ *and* $1 \leq s \leq r-2$,

$$[r, s] \subseteq [r, r-1][r-1, s].$$

Proof. Let

$$\alpha = \begin{pmatrix} a_1 & a_2 & \ldots & a_r \\ b_1 & b_2 & \ldots & b_r \end{pmatrix} \in [r, s],$$

where $a_1 < a_2 < \cdots < a_r$ and $b_1 \leq b_2 \leq \cdots \leq b_r$, and suppose that $b_i = b_{i+1}$. Let

$$\xi = \begin{pmatrix} a_1 & \ldots & a_i & a_{i+1} & a_{i+2} & \ldots & a_r \\ 1 & \ldots & i & i & i+1 & \ldots & r-1 \end{pmatrix},$$

$$\eta = \begin{pmatrix} 1 & \ldots & i & i+1 & \ldots & r-1 \\ b_1 & \ldots & b_i & b_{i+2} & \ldots & b_r \end{pmatrix};$$

then $\xi \in [r, r-1]$, $\eta \in [r-1, s]$, and $\alpha = \xi\eta$. ∎

To prove our assertion that $\langle G\rangle = \mathcal{SPO}_n$ we need to prove that $[r,s] \subseteq \langle G\rangle$ for $1 \leq r \leq n-1$ and $1 \leq s \leq r$. From (4.5) we have that $[r,r] \subseteq \langle G\rangle$ for $1 \leq r \leq n-1$, and from Lemma 4.6 it follows that $[n-1, n-2] \subseteq \langle G\rangle$. From Lemma 4.7 we then deduce that $[r, r-1] \subseteq \langle G\rangle$ for $2 \leq r \leq n-1$. Then finally by repeated use of Lemma 4.8 we have

$$[r,s] \subseteq [r,r-1][r-1,r-2]\ldots[s+1,s]$$

and so the proof is complete. ∎

References

[1] Garba, G. U., *Idempotents in partial transformation semigroups*, Proc. Royal Soc. Edinburgh A **116** (1990), 359-366.

[2] Gomes, Gracinda M. S., and John M. Howie, *On the ranks of certain finite semigroups of transformations*, Math. Proc. Cambridge Phil. Soc. **101** (1987), 395–403.

[3] Howie, John M., *Products of idempotents in certain semigroups of transformations*, Proc. Edinburgh Math. Soc. (2) **17** (1971), 223–236.

[4] Howie, John M., "An introduction to semigroup theory", Academic Press, 1976.

[5] Howie, John M. and Robert B. McFadden, *Idempotent rank in finite full transformation semigroups*, Proc. Royal Soc. Edinburgh A **114** (1990), 161–167.

Departamento de Matemática
Faculdade de Ciências
Universidade de Lisboa
1700 Lisboa, Portugal

Mathematical Institute
University of St. Andrews
St. Andrews, U.K.

Received July 24, 1990
and in final form January 30,1991

Semigroup Forum Vol. 45 (1992) 283–287

RESEARCH ARTICLE

On Some Problems of Petrich Concerning Bruck and Reilly Semigroups

Azeeza Ali Ismaeel*

Communicated by T. E. Hall

Abstract. For an inverse semigroup S with its idempotents dually well-ordered, we prove that S is isomorphic to the semigroup of all one-to-one partial right translations of S. Also, we prove for a Bruck semigroup $S = B(T,\alpha)$ that S is E-unitary if and only if T is E-unitary and α is an idempotent pure homomorphism. Moreover, we characterize all E-unitary covers of $B(T,\alpha)$, where T is a finite chain of groups.

Introduction

A semigroup is called a Reilly semigroup if it is bisimple, regular, and the set of its idempotents is dually well-ordered of type ω^*. Reilly gave a description of such inverse semigroups [3], and characterized the homomorphisms between two Reilly semigroups [3].

In this paper we solve three problems which arose in Petrich's book [2].

In section 1 we prove that if S is an inverse semigroup such that its idempotents are dually well-ordered then S is isomorphic to $\widehat{S}$. The semigroup $\widehat{S}$ was introduced by McAlister [1] (see Definition 2 below). This immediately gives a solution of Problem V.2.12 in Petrich's book [2].

In Section 2, we give necessary and sufficient conditions for a Bruck semigroup $B(T,\alpha)$ to be E-unitary [2, Problem XI.4.7].

In Section 3, we give a characterization of all E-unitary covers of $S = B(T,\alpha)$ where T is a finite chain of groups. This gives a characterization of all E-unitary covers of Reilly semigroups [2, Problem VII.4.17].

For undefined notions and notations the reader is referred to [2].

Section 1

First we recall some definitions.

Definition 1. A one-to-one partial right translation of a semigroup S is a one-to-one partial transformation ρ, the domain of which is a left ideal of S, such that $(xy)\rho = x \cdot (y)\rho$ for ech $x \in S$, $y \in \operatorname{dom}\rho$.

Definition 2. $\widehat{S}$ is the set of all one-to-one partial right translations of S.

$\widehat{S}$ is a semigroup, and if S is an inverse semigroup, then so also is $\widehat{S}$.

---

* I wish to express my thanks to Dr. G. Pollák and Dr. Mária B. Szendrei for their valuable advice.

1980 Mathematics subject classification (Amer. Math. Soc.): 20 M 10

Lemma 1. *Let S be a regular semigroup with its set of idempotents E_S dually well-ordered. Then S is an inverse semigroup in which every left ideal is principal.*

The proof is clear.

Lemma 2. *Let S be a regular semigroup with its set of idempotents E_S dually well-ordered. If $\rho_1, \rho_2 \in \widehat{S}$ are such that $\operatorname{dom} \rho_1 = \operatorname{dom} \rho_2 = Se$ for some $e \in E_S$ and $e\rho_1 = e\rho_2$, then $\rho_1 = \rho_2$.*

Proof. Let $x \in Se$, i.e. $x = ye$ for some $y \in S$. Then $x\rho_1 = (ye)\rho_1 = y \cdot e\rho_1 = (ye)\rho_2 = x\rho_2$. Hence $\rho_1 = \rho_2$. ∎

Theorem 1. *If S is a regular semigroup with its set of idempotents E_S dually well-ordered, then S is isomorphic to $\widehat{S}$.*

Proof. For each $a \in S$, let $\rho_a \in \widehat{S}$ be given by $x\rho_a = xa$ for each $x \in \operatorname{dom} \rho_a = Saa^{-1}$. Define a map $\theta : S \to \widehat{S}$ by $a\theta = \rho_a$. By [2, Theorem IV.1.6], θ is a monomorphism. We will check that θ is surjective. Suppose $\rho \in \widehat{S}$. Since E_S is dually well-ordered, there is a unique $e \in E_S$ such that $\operatorname{dom} \rho = Se$. Let $a = e\rho \in S$. Then

$$\begin{aligned} e\rho &= e\rho \cdot (e\rho)^{-1} \cdot e\rho = (e\rho \cdot (e\rho)^{-1} \cdot e)\rho = (e \cdot e\rho \cdot (e\rho)^{-1})\rho \\ &= (e\rho \cdot (e\rho)^{-1})\rho \end{aligned}$$

and since ρ is one-to-one then

$$e = e\rho \cdot (e\rho)^{-1} = aa^{-1} .$$

But $\operatorname{dom} \rho_a = Saa^{-1} = Se = \operatorname{dom} \rho$ while $e\rho_a = ea = e \cdot e\rho = e\rho$, so by Lemma 2, $\rho = \rho_a$.[1) ∎

Corollary. *If S is a Reilly semigroup, then S is isomorphic to $\widehat{S}$.*

Section 2

Let T be a monoid, α be a homomorphism of T into its group of units and $\mathbb{N}$ be the set of all nonnegative integers. On the set $S = \mathbb{N} \times T \times \mathbb{N}$ define a multiplication by

$$(m, a, n)(p, b, q) = \big(m + p - r, (a\alpha^{p-r}) \cdot (b\alpha^{n-r}), n + q - r\big)$$

where $r = \min\{n, p\}$, α^0 is the identity mapping on T. Then $S = B(T, \alpha)$ is called the Bruck semigroup defined by T and α.

We have

[1)] I am greatly indebted to P. G. Trotter for this short proof.

Theorem 2. *$S = B(T, \alpha)$ is E-unitary if and only if T is E-unitary and α is an idempotent pure homomorphism.*

Proof. Let S be an E-unitary semigroup. First we show that if $a \in T$ and $a\alpha = 1$, then $a \in E_T$. Indeed, if $a\alpha = 1$, then

$$(0, a, 0)(1, a\alpha, 1) = (1, a\alpha, 1) \in E_S ,$$

and hence $(0, a, 0) \in E_S$ since S is E-unitary. This implies that $a \in E_T$, whence α is an idempotent pure homomorphism. Since $T \cong (0, T, 0) \subseteq S$, T is an E-unitary semigroup.

Conversely, suppose T is E-unitary and α is idempotent pure. Let $(m, a, n) \in S$ and (p, e, p), $(q, f, q) \in E_S$ such that $(m, a, n)(p, e, p) = (q, f, q)$. Then $m + p - r = q$ and $n + p - r = q$ where $r = \min\{n, p\}$, and this implies $n = m$. Furthermore $a\alpha^{p-r} \cdot e\alpha^{n-r} = f$. Now $e\alpha^{n-r}$ is an idempotent; since T is E-unitary this implies $a\alpha^{p-r} \in E_T$. Hence a is idempotent because α is idempotent pure. Thus $(m, a, n) \in E_S$. So S is E-unitary. ∎

Section 3

Recall the definition of an E-unitary cover.

Definition 3. Let S, T be inverse semigroups. T is said to be an E-unitary cover of Sif T is an E-unitary inverse semigroup and there exists an idempotent separating homomorphism from T onto S.

Let T be an inverse semigroup with identity. For a homomorphism α from T into its group of units, and an arbitrary element $x \in T$, denote $x_0 = xx^{-1}$, $x_m = x \cdot x\alpha \cdot \ldots \cdot x\alpha^{m-1}$.

Theorem 3. *Let $S = B(T, \alpha)$ be a Bruck semigroup over a finite chain of groups T. Let T' be an E-unitary cover of T with an idempotent separating homomorphism θ from T' onto T, and let α' be an idempotent pure homomorphism from T' into its group of units such that there exists an element z which belongs to the group of units of T with $\alpha'\theta\varepsilon_z = \theta\alpha$, where $a\varepsilon_z = z^{-1}az$ for all $a \in T$.*

Define $\eta : B(T', \alpha') \to B(T, \alpha)$ by

$$\eta : (m, a, n) \mapsto (m, z_m^{-1} \cdot a\theta \cdot z_n, n) .$$

Then η is an idempotent separating homomorphism from $B(T', \alpha')$ onto $B(T, \alpha)$, and $B(T', \alpha')$ is an E-unitary cover of $B(T, \alpha)$.

Conversely, every E-unitary cover of S can be constructed in this way uniquely up to isomorphism.

Proof. Since α' is idempotent pure and T' is an E-unitary inverse semigroup, the Bruck semigroup $B(T', \alpha')$ is an E-unitary inverse semigroup (by Theorem 2). Now to prove η is onto, we let $u = (m, t, n) \in B(T, \alpha)$. Since θ is onto, there exist elements $a, b \in T'$ such that $a\theta = t$ and $b\theta = z$. First we prove by induction that

$$b{\alpha'}^m\theta = z_{m+1} \cdot z_m^{-1} .$$

For $m = 0$ this is trivial. Suppose that $b\alpha'^r\theta = z_{r+1} \cdot z_r^{-1}$. Then

$$\begin{aligned}
b\alpha'^{r+1}\theta &= (b\alpha'^r)\alpha'\theta \\
&= z \cdot \big((b\alpha'^r)\theta\alpha\big) \cdot z^{-1} = z \cdot \big((z_{r+1} \cdot z_r^{-1})\alpha\big) \cdot z^{-1} \\
&= z \cdot \big((z \cdot z\alpha \cdot \ldots \cdot z\alpha^r \cdot (z\alpha^{r-1}) \cdot \ldots \cdot z^{-1})\alpha\big) \cdot z^{-1} \\
&= z \cdot \big(z\alpha \cdot z\alpha^2 \cdot \ldots \cdot z\alpha^{r+1} \cdot (z\alpha^r)^{-1} \cdot \ldots \cdot (z\alpha)^{-1}\big) \cdot z^{-1} = z_{r+2} \cdot z_{r+1}^{-1} .
\end{aligned}$$

Now applying this we shall have $z_m = (b^{-1})_m^{-1}\theta$, by

$$\begin{aligned}
z_m &= z_m \cdot \left(z_{m-1}^{-1} \cdot z_{m-1}\right)\left(z_{m-2}^{-1} \cdot z_{m-2}\right) \ldots \left(z^{-1} \cdot z\right) \\
&= b\alpha'^{m-1}\theta \cdot b\alpha'^{m-2}\theta \cdot \ldots \cdot b\alpha'\theta \cdot b\theta \\
&= (b\alpha'^{m-1} \cdot b\alpha'^{m-2} \cdot \ldots \cdot b\alpha' \cdot b)\theta \\
&= (b^{-1} \cdot b^{-1}\alpha' \cdot \ldots \cdot b^{-1}\alpha'^{m-1})^{-1}\theta \\
&= \left((b^{-1})_m^{-1}\right)\theta .
\end{aligned}$$

Let $v = \left(m, (b^{-1})_m^{-1} \cdot a \cdot (b^{-1})_n, n\right) \in B(T', \alpha')$. Then

$$\begin{aligned}
v\eta &= \left(m, z_m^{-1} \cdot \left((b^{-1})_m^{-1} \cdot a \cdot (b^{-1})_n\right)\theta \cdot z_n, n\right) = (m, a\theta, n) \\
&= (m, t, n) = u .
\end{aligned}$$

Hence η is onto.

The fact that η is a homomorphism can be proved analogously to the case of Reilly semigroups [2; Theorem XI.1.1].

Now to prove η is idempotent separating, let $e, f \in E_{B(T',\alpha')}$ such that $e\eta = f\eta$. If $e = (m, e', m)$ and $f = (n, f', n)$ for some $e', f' \in E_{T'}$, then $m = n$ and $z_m^{-1} \cdot e'\theta \cdot z_m = z_m^{-1} \cdot f'\theta \cdot z_m$. This implies $e'\theta = f'\theta$, for $z_m^{-1} \cdot z_m$ is the identity element of T. Since θ is idempotent separating, we have $e' = f'$, thus $e = f$ and $B(T', \alpha')$ is an E-unitary cover of S.

Conversely, let S' be an E-unitary cover of $S = B(T, \alpha)$. This means that there exists an idempotent separating homomorphism η from S' onto S. First, to prove S' is simple, we let e, f be two different idempotents in S'. Then $e\eta \neq f\eta$ which implies that there exists $x \in S$ such that $x \cdot x^{-1} = e\eta$ and $x^{-1} \cdot x \leqq f\eta$ (for S is simple). Since η is onto then there exists $y \in S'$ such that $y\eta = x$, therefore $(yy^{-1})\eta = xx^{-1} = e\eta$ and $(y^{-1}y)\eta = x^{-1}x \cdot f\eta = (y^{-1}y)\eta \cdot f\eta = (y^{-1}y \cdot f)\eta$.

Since η is idempotent separating, we have $yy^{-1} = e$, $y^{-1}y = y^{-1}yf \leqq f$. Hence S' is simple.

As $\eta\big|_{E_{S'}}$ is an isomorphism onto E_S it is clear that $E_{S'}$ is dually well-ordered of type ω^*. Thus $S' \cong B(T', \alpha')$ for some finite chain of groups T' and a homomorphism α' from T' into its group of units. Since S' is an E-unitary inverse semigroup, T' is E-unitary and α' is an idempotent pure homomorphism (by Theorem 2). Now as in the case of Reilly semigroups [2, Theorem XI.1.1] and because η is an idempotent separating homomorphism, one can show easily that there exists an idempotent separating homomorphism θ from T' onto T, and an element z in the group of units of T such that $\alpha'\theta\varepsilon_z = \theta\alpha$ and T' is an E-unitary cover of T, and η is defined by

$$\eta(m, s, n) = (m, z_m^{-1} \cdot s\theta \cdot z_n, n) .$$

This completes the proof of Theorem 3. ■

Corollary. *Let $S = B(G,\alpha)$ be a Reilly semigroup, and let G' be any group, $\alpha' : G' \to G'$ a monomorphism, and $\theta : G' \to G$ an epimorphism such that there exists $z \in G$ with $\alpha'\theta\varepsilon_z = \theta\alpha$. Define $\eta : B(G',\alpha') \to B(G,\alpha)$ by*

$$\eta(m,a,n) \mapsto (m, z_m^{-1} \cdot a\theta \cdot z_n, n) \ .$$

Then η is an idempotent separating homomorphism from $B(G',\alpha')$ onto $B(G,\alpha)$, and $B(G',\alpha')$ is an E-unitary cover of $B(G,\alpha)$.

Conversely, every E-unitary cover of S can be constructed in this way uniquely up to isomorphism. ■

References

[1] McAlister, D. B., *Some covering and embedding theorems for inverse semigroups*, J. Austr. Math. Soc. **22** (1976), 188–211.

[2] Petrich, M., *Inverse Semigroups*, John Wiley & Sons, Inc. New York (1984).

[3] Reilly, N. R., *Bisimple ω-semigroups*, Proc. Glasgow Math. **7** (1965-66), 160–167.

c/o G. Pollák
Bolyai Institute
Aradi vértanuk tere 1.
H-6720 Szeged
Hungary

Received September 23, 1990
and in final form April 3, 1991

Semigroup Forum Vol. 45 (1992) 288–301

RESEARCH ARTICLE

An Associative Operator on the Lattice of Varieties of Inverse Semigroups

David Cowan

Communicated by N. R. Reilly

Introduction

Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of inverse semigroups. Define their product $Wr(\mathcal{U}, \mathcal{V})$ to be the variety generated by wreath products of semigroups in $\mathcal{U}$ with semigroups in $\mathcal{V}$. Then Wr is a binary operator on the lattice $\mathcal{L}(\mathcal{I})$ of varieties of inverse semigroups. In this note we show that the operator Wr is associative and so $\mathcal{L}(\mathcal{I})$ is a semigroup (in fact, a monoid) under the operation Wr. Using the associativity of this operation we show that, for any variety $\mathcal{V}$ of inverse semigroups, the variety generated by wreath products of semilattices and semigroups in $\mathcal{V}$ is the largest variety (denoted $\mathcal{V}^K$) satisfying the same idempotent laws as $\mathcal{V}$. This variety coincides with the variety generated by the wreath product of the two element semilattice and the $\mathcal{V}$-free semigroup $F\mathcal{V}(X)$ on a countably infinite set X. Consequently, we have that if the variety $\mathcal{V}$ has decidable equational theory then so does $\mathcal{V}^K$.

Section one is devoted to preliminary material. The second section deals exclusively with showing that the operator Wr is associative. In section three we show that $\mathcal{V}^K$ equals $Wr(\mathcal{S}, \mathcal{V})$ and consider some consequences of this result.

§1. Preliminaries

We will follow the notation and terminology of Petrich [6] to which the reader is referred for basic information on inverse semigroups and varieties of inverse semigroups. The following notation will be used consistently:

$\mathcal{I}$	—	the variety of all inverse semigroups
$\mathcal{T}$	—	the trivial variety
$\mathcal{G}$	—	the variety of groups
$\mathcal{S}$	—	the variety of semilattices
$\mathcal{B}$	—	the variety generated by the five element combinatorial Brandt semigroup
$\mathcal{B}^1$	—	the variety generated by the five element combinatorial Brandt semigroup with an identity adjoined
$\mathcal{L}(\mathcal{V})$	—	the lattice of varieties of inverse semigroups contained in $\mathcal{V}$
$\langle \mathcal{C} \rangle$	—	the variety of inverse semigroups generated by the class $\mathcal{C}$ of inverse semigroups; if $\mathcal{C}$ consists of the single semigroup S, we write $\langle S \rangle$ instead of $\langle \mathcal{C} \rangle$
$F\mathcal{V}(X)$	—	the relatively free object on X in $\mathcal{V}$
$\rho(\mathcal{V})$	—	the fully invariant congruence on $F\mathcal{I}(X)$ corresponding to $\mathcal{V}$, where X is a countably infinite set
X^+	—	the free semigroup on X

X^{-1} — the set of formal inverses of the elements of X and in one-to-one correspondence with X via $x \leftrightarrow x^{-1}$.

Let S and T be inverse semigroups and suppose that T is an inverse subsemigroup of $\mathcal{I}(I)$, the symmetric inverse semigroup on I. Let ${}^IS$ denote the set of functions (written on the right) from subsets of I into S. For any $\psi \in {}^IS$, denote the domain of ψ by $\mathbf{d}\psi$. Define a multiplication on ${}^IS$ by

$$i(\psi \cdot \psi') = (i\psi) \cdot (i\psi') \qquad [i \in \mathbf{d}\psi \cap \mathbf{d}\psi'] .$$

For any $\beta \in \mathcal{I}(I)$ and $\psi \in {}^IS$, we define a mapping ${}^{\beta}\psi$ by

$$i({}^{\beta}\psi) = (i\beta)\psi \qquad [i \in \mathbf{d}\beta, i\beta \in \mathbf{d}\psi] .$$

The *(right) wreath product of S and T* is the set

$$S\ wr\ T = \{(\psi, \beta) \in {}^IS \times T : \mathbf{d}\psi = \mathbf{d}\beta\}$$

with multiplication given by

$$(\psi, \beta) \cdot (\psi', \beta') = (\psi\, {}^{\beta}\psi', \beta\beta') .$$

If T is an inverse subsemigroup of $\mathcal{I}(I)$, we will sometimes write (T, I) for T if we wish to emphasize the set I on which T acts.

This definition of wreath product was introduced by Houghton [4]. In [4] the wreath product $W(S,T)$ of inverse semigroups S and T is, in our notation, $S\ wr\ (T,T)$ where T is given the Wagner representation by partial right translations. Our notation follows Petrich [6; V.4]. It is shown in [4] (and in [3]) that if S and (T,I) are inverse semigroups then so is $S\ wr\ (T,I)$.

We find it convenient in our investigations to make use of a graphical representation of semigroups introduced by Stephen [7].

Let $P = (X; R)$ be a fixed presentation of the inverse semigroup S with τ the corresponding congruence on $F\mathcal{I}(X)$, the free inverse semigroup on X. Let $w \in S$ and R_w be the $\mathcal{R}$-class of w in S. The *Schützenberger graph of R_w with respect to P* is the labelled digraph $\Gamma(w)$, where

$$V(\Gamma(w)) = R_w$$
$$E(\Gamma(w)) = \{(v_1, x, v_2) : v_1, v_2 \in R_w, x \in X \cup X^{-1} \text{ and } v_1(x\tau) = v_2\}.$$

The *Schützenberger representation of w* (with respect to P) is the birooted labelled digraph $(ww^{-1}, \Gamma(w), w)$, where ww^{-1} is the *start* vertex and w is the *end* or *terminal* vertex. The Schützenberger representation of the semigroup S is the family of birooted graphs $\{(ww^{-1}, \Gamma(w), w) : w \in S\}$. A homomorphism of one graph into another is a map on the vertices (that induces a map on the edges) which preserves incidence, orientation and labelling. Graph isomorphisms and embeddings are defined in the obvious way. We remark that if S is a group, then for any $w \in S$, $\Gamma(w)$ is the Cayley graph of S (see [7; 3.7]).

Schützenberger graphs enjoy the following properties:
Let $v \in S$, $\Gamma(v)$ be its Schützenberger graph with respect to P, $x \in X \cup X^{-1}$, $v_1, v_2, v_3 \in R_v$ and $w \in (X \cup X^{-1})^+$ (see [7]).

a) if (v_1, x, v_2) is an edge in $\Gamma(v)$ then (v_2, x^{-1}, v_1) is also an edge in $\Gamma(v)$;
b) if (v_1, x, v_2) and (v_1, x, v_3) are edges in $\Gamma(v)$ then $v_2 = v_3$;

c) if (v_2, x, v_1) and (v_3, x, v_1) are edges in $\Gamma(v)$ then $v_2 = v_3$;
d) $v_1(w\tau) = v_2$ if and only if w labels a $v_1 - v_2$ walk;
e) $(w\tau) \geq v$ if and only if w labels a $vv^{-1} - v$ walk;
f) $v_1 \ \mathcal{D} \ v_2$ if and only if $\Gamma(v_1)$ is isomorphic to $\Gamma(v_2)$;
g) $v_1 \ \mathcal{R} \ v_2$ if and only if there exists an isomorphism from $\Gamma(v_1)$ to $\Gamma(v_2)$ such that $v_1v_1^{-1}$ is mapped to $v_2v_2^{-1}$;
h) $v_1 \ \mathcal{L} \ v_2$ if and only if there exists an isomorphism from $\Gamma(v_1)$ to $\Gamma(v_2)$ such that v_1 is mapped to v_2.

We will only be considering Schützenberger graphs of the $\mathcal{V}$-free inverse semigroup on (countably infinite) X with respect to the presentation $P = (X; \rho(\mathcal{V}))$, for $\mathcal{V}$ a variety of inverse semigroups. For further properties and a detailed discussion of Schützenberger graphs we refer the reader to Stephen [7].

Let $\mathcal{V}$ be a variety of inverse semigroups and ρ the fully invariant congruence on $F\mathcal{I}(X)$ corresponding to $\mathcal{V}$. Let $w \in (X \cup X^{-1})^+$ and let $\Gamma_{\mathcal{V}}(w)$ be the Schützenberger graph of w in the $\mathcal{V}$-free inverse semigroup on S. Let Y be a countably infinite set and Y^{-1} a set disjoint from Y and in 1-1 correspondence with Y via $y \leftrightarrow y^{-1}$. Assume that $X \cup X^{-1}$ and $Y \cup Y^{-1}$ are disjoint. From $\Gamma_{\mathcal{V}}(w)$ we obtain the *doubly labelled Schützenberger graph* $\overline{\Gamma_{\mathcal{V}}(w)}$ *of* w *relative to* $\mathcal{V}$, as follows:

$$\overline{\Gamma_{\mathcal{V}}(w)} = (\Gamma_{\mathcal{V}}(w), \lambda_w)$$

where

$$\lambda_w : E(\Gamma_{\mathcal{V}}(w)) \to Y \cup Y^{-1}$$

satisfies

(i) $(v_1, x, v_2) \in E(\Gamma_{\mathcal{V}}(w))$ and $x \in X$ implies that $\lambda_w(v_1, x, v_2) \in Y$;
(ii) $\lambda_w(v_2, x^{-1}, v_1) = [\lambda_w(v_1, x, v_2)]^{-1}$;
(iii) $\lambda_w(v_1, x, v_2) = \lambda_w(v_3, z, v_4)$ implies that $v_1 = v_3$, $v_2 = v_4$, and $x = z$.

We call x the *primary label* and $\lambda_w(v_1, x, v_2)$ the *secondary label* of the edge (v_1, x, v_2). Condition (iii) says that no two distinct edges have the same secondary label, condition (ii) says that inverse edges have inverse secondary labels and condition (i) is just convenient. Thus, the doubly labelled basic graph of w is just $\Gamma_{\mathcal{V}}(w)$ with a secondary label attached to each edge such that inverse edges have inverse secondary labels and no two distinct edges have the same secondary label.

We define the *derived word* $d_{\mathcal{V}}(w)$ of w relative to $\mathcal{V}$ as follows:

Let v_1 and v_2 be the start and end vertices respectively, of the Schützenber graph $\Gamma_{\mathcal{V}}(w)$ of w relative to $\mathcal{V}$. Then w labels a $v_1 - v_2$ walk in $\overline{\Gamma_{\mathcal{V}}(w)}$ by primary labels. Let $e_1, \ldots e_n$ be the edge sequence corresponding to this walk. Define $d_{\mathcal{V}}(w) = \lambda_w(e_1)\lambda_w(e_2)\ldots\lambda_w(e_n) \in (Y \cup Y^{-1})^+$. That is, $d_{\mathcal{V}}(w)$ is just the word obtained by taking the secondary labels from each edge in our $v_1 - v_2$ walk.

Note that if $w = a_1 \ldots a_k$, $a_i \in X \cup X^{-1}$ for $i = 1, \ldots, k$, and $d_{\mathcal{V}}(w) = b_1 \ldots b_m$, $b_i \in Y \cup Y^{-1}$, then $m = k$ and if e is the edge corresponding to a_i in the $v_1 - v_2$ walk labelled by w in $\Gamma_{\mathcal{V}}(w)$, then $b_i = \lambda_w(e)$ is the secondary label of e in $\overline{\Gamma_{\mathcal{V}}(w)}$.

We now present some preliminary results which are required in the sequel.

Proposition 1.1 ([3] Prop. 2.2; [2], Prop. 4.1.3). *Let $\mathcal{V}$ be a variety of inverse semigroups and let $v, w \in (X \cup X^{-1})^+$.*

a) *If* $w\,\rho(\mathcal{V})\,w^2$ *then* $d_{\mathcal{V}}(w^2) = [d_{\mathcal{V}}(w)]^2$;
b) *If* ρ *is the Wagner congruence, then* $v\,\rho\,w$ *if and only if* $d_{\mathcal{V}}(v)\,\rho\,d_{\mathcal{V}}(w)$.

Throughout this note we will write $w \leq_{\mathcal{V}} v$ as a shorthand notation for $w\rho(\mathcal{V}) \leq v\rho(\mathcal{V})$, where $\leq$ is the natural partial order on $F\mathcal{I}(X)/\rho(\mathcal{V})$.

Lemma 1.2 ([3]; Lemma 2.3). *Let* $w = a_1 \dots a_k$ *and* $v = d_1 \dots d_m$ *with* $w\,\rho(\mathcal{V})\,v$. *Set* $d_{\mathcal{V}}(w) = b_1 \dots b_k$ *and* $d_{\mathcal{V}}(v) = c_1 \dots c_m$, *where we construct both* $d_{\mathcal{V}}(w)$ *and* $d_{\mathcal{V}}(v)$ *from the same doubly labelled Schützenberger graph* $\overline{\Gamma_{\mathcal{V}}(w)}$. *Then*

$$\begin{array}{lll}
\text{a)} & b_i = c_j & \Leftrightarrow \quad w \leq_{\mathcal{V}} a_1 \dots a_{i-1} d_j \dots d_m \text{ and } a_i = d_j ; \\
\text{b)} & b_i = c_j^{-1} & \Leftrightarrow \quad w \leq_{\mathcal{V}} a_1 \dots a_i d_j \dots d_m \text{ and } a_i = d_j^{-1} \\
 & & \Leftrightarrow \quad w \leq_{\mathcal{V}} a_1 \dots a_{i-1} d_{j+1} \dots d_m \text{ and } a_i = d_j^{-1} .
\end{array}$$

Remark. If we take $w = v$ in Lemma 1.2, we obtain

$$\begin{array}{lll}
\text{a)} & b_i = b_j & \Leftrightarrow \quad w \leq_{\mathcal{V}} a_1 \dots a_{i-1} a_j \dots a_m \text{ and } a_i = a_j ; \\
\text{b)} & b_i = b_j^{-1} & \Leftrightarrow \quad w \leq_{\mathcal{V}} a_1 \dots a_i a_j \dots a_m \text{ and } a_i = a_j^{-1} ; \\
 & & \Leftrightarrow \quad w \leq_{\mathcal{V}} a_1 \dots a_{i-1} a_{j+1} \dots a_m \text{ and } a_i = a_j^{-1} .
\end{array}$$

Definition. Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of inverse semigroups. Define the product variety of $\mathcal{U}$ and $\mathcal{V}$ by $Wr(\mathcal{U},\mathcal{V}) = \langle T\,wr\,(F,I) : T \in \mathcal{U}, F \in \mathcal{V}\rangle$.

Note that if $\mathcal{U}$ and $\mathcal{V}$ are varieties of inverse semigroups then $\mathcal{U} \subseteq Wr(\mathcal{U},\mathcal{V})$ and $\mathcal{V} \subseteq Wr(\mathcal{U},\mathcal{V})$.

Lemma 1.3 ([3]; Theorem 2.5). *Let* $\mathcal{U}$ *and* $\mathcal{V}$ *be varieties of inverse semigroups. Then for any words* $u, w \in F\mathcal{I}(X)$,

$$u\,\rho(Wr(\mathcal{U},\mathcal{V}))\,w \Leftrightarrow u\,\rho(\mathcal{V})\,w \text{ and } d_{\mathcal{V}}(u)\,\rho(\mathcal{U})\,d_{\mathcal{V}}(w) ,$$

where $d_{\mathcal{V}}(u)$ *and* $d_{\mathcal{V}}(w)$ *are both obtained from the same doubly labelled Schützenberger graph* $\overline{\Gamma_{\mathcal{V}}(w)}$.

Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of inverse semigroups. The *Mal'cev product* of $\mathcal{U}$ and $\mathcal{V}$, denoted by $\mathcal{U} \circ \mathcal{V}$, is the collection of those inverse semigroups S for which there exists a congruence ρ on S with the property that $e\rho \in \mathcal{U}$ for all $e \in E_S$ and $S/\rho \in \mathcal{V}$; we say that ρ *witnesses* that $S \in \mathcal{U} \circ \mathcal{V}$.

In general, $\mathcal{U} \circ \mathcal{V}$ is not a variety. However, when $\mathcal{U}$ is a variety of groups, $\mathcal{U} \circ \mathcal{V}$ is a variety [see [6; XII 8.3] or [1]].

It is always true that $Wr(\mathcal{U},\mathcal{V}) \subseteq \langle \mathcal{U} \circ \mathcal{V}\rangle$ since if $S\,wr\,(T,I)$ is any generator of $Wr(\mathcal{U},\mathcal{V})$ then the projection map onto T witnesses that $S\,wr\,(T,I)$ belongs to $\mathcal{U} \circ \mathcal{V}$. In general the two varieties are not equal. For example, the five element combinatorial Brandt semigroup with an identity adjoined belongs to $\mathcal{B} \circ \mathcal{B}$, as witnessed by the Rees congruence determined by the ideal of all nonidentity elements. On the other hand, this semigroup does not belong to $Wr(\mathcal{B},\mathcal{B}) = \mathcal{B}$ (see Theorem 1.5 below). However, we do have the following result.

Theorem 1.4 ([3]; Theorem 3.4). *Let* $\mathcal{V}$ *be a variety of inverse semigroups and let* $\mathcal{U}$ *be a variety of groups. Then* $Wr(\mathcal{U},\mathcal{V}) = \mathcal{U} \circ \mathcal{V}$, *the Mal'cev product of* $\mathcal{U}$ *and* $\mathcal{V}$. *Thus, if* $\mathcal{U}$ *and* $\mathcal{V}$ *are both varieties of groups, the (group variety) product of* $\mathcal{U}$ *and* $\mathcal{V}$ *coincides with* $Wr(\mathcal{U},\mathcal{V})$.

Theorem 1.5 ([3]; Theorem 5.1.5). *Let $\mathcal{V}$ be a variety of inverse semigroups. Then*

$$Wr(\mathcal{V},\mathcal{S}) = \mathcal{V} \vee \mathcal{S}, Wr(\mathcal{V},\mathcal{B}) = \mathcal{V} \vee \mathcal{B}, Wr(\mathcal{V},\mathcal{T}) = Wr(\mathcal{T},\mathcal{V}) = \mathcal{V}$$

and

$$Wr(\mathcal{I},\mathcal{V}) = Wr(\mathcal{V},\mathcal{I}) = \mathcal{I} \ .$$

Notation. *If $\mathcal{V}$ is a variety of groups we denote by $\mathcal{V}^K$ the variety of inverse semigroups $[u^2 = u : u^2 = u$ is a law in $\mathcal{V}]$.*

2. The Associativity of Wr

The binary operator Wr on the lattice of varieties of inverse semigroups is, in fact, an associative operator and so $(\mathcal{L}(\mathcal{I}), Wr)$ is a semigroup. The proof of this makes use of Theorem 1.3, the description of the fully invariant congruence on the free inverse semigroup corresponding to $Wr(\mathcal{U},\mathcal{V})$, for any pair of varieties $\mathcal{U}$ and $\mathcal{V}$ of inverse semigroups, and Lemma 1.2, the description of the derived word relative to the variety $\mathcal{V}$.

We say that the two equations $u_1 = u_2$ and $v_1 = v_2$ are *equivalent* if each is a consequence of the other. Another way of saying this is that $u_1 = u_2$ and $v_1 = v_2$ are equivalent if and only if, for any variety $\mathcal{U}$ of inverse semigroups, $\mathcal{U}$ satisfies the equation $u_1 = u_2$ if and only if $\mathcal{U}$ satisfies the equation $v_1 = v_2$.

Lemma 2.1. *Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of inverse semigroups and let $v, w \in (X \cup X^{-1})^+$ be such that $w\,\rho(\mathcal{U})\,v$ and $d_{\mathcal{U}}(w)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(v)$ (or, equivalently, $w\,\rho(Wr(\mathcal{V},\mathcal{U})\,v)$. Set*

$w = a_1 \dots a_n,\ v = b_1 \dots b_m,$ *where $a_i, b_j \in X \cup X^{-1}$, $i = 1,\dots,n,\ j = 1,\dots,m$.*

$d_{\mathcal{U}}(w) = c_1 \dots c_n,\ d_{\mathcal{U}}(v) = d_1 \dots d_m,$ *where both words are constructed from the same doubly labelled Schützenberger graph $\overline{\Gamma_{\mathcal{U}}(w)}$.*

$d_{\mathcal{V}}(d_{\mathcal{U}}(w)) = e_1 \dots e_n,\ d_{\mathcal{V}}(d_{\mathcal{U}}(v)) = f_1 \dots f_m,$ *where both words are constructed from the same doubly labelled Schützenberger graph $\overline{\Gamma_{\mathcal{V}}(d_{\mathcal{U}}(w))}$.*

$d_{Wr(\mathcal{V},\mathcal{U})}(w) = g_1 \dots g_n,\ d_{Wr(\mathcal{V},\mathcal{U})}(v) = h_1 \dots h_m,$ *where both words are constructed from the same doubly labelled Schützenberger graph $\overline{\Gamma_{Wr(\mathcal{V},\mathcal{U})}(w)}$.*

Then the equations $d_{\mathcal{V}}(d_{\mathcal{U}}(w)) = d_{\mathcal{V}}(d_{\mathcal{U}}(v))$ and $d_{Wr(\mathcal{V},\mathcal{U})}(w) = d_{Wr(\mathcal{V},\mathcal{U})}(v)$ are equivalent.

Proof. We prove the stronger statement that each of the two equations is a one-to-one relabelling of the other. That is, we prove the following statements. For all i, j,

1) $g_i = h_j \ \Leftrightarrow e_i = f_j$;
2) $g_i = h_j^{-1} \ \Leftrightarrow e_i = f_j^{-1}$;
3) $g_i = g_j \ \Leftrightarrow e_i = e_j$;
4) $g_i = g_j^{-1} \ \Leftrightarrow e_i = e_j^{-1}$;

5) $h_i = h_j \quad \Leftrightarrow f_i = f_j$;
6) $h_i = h_j^{-1} \quad \Leftrightarrow f_i = f_j^{-1}$.

1) First of all, observe that

$$\begin{aligned} g_i = h_j \quad & \Leftrightarrow \quad w \le_{Wr(\mathcal{V},\mathcal{U})} a_1 \dots a_{i-1} b_j \dots b_m \text{ and } a_i = b_j \\ & \qquad\quad \text{(Lemma 1.2 since } w\,\rho(Wr(\mathcal{V},\mathcal{U}))v) \\ & \Leftrightarrow \quad w\,\rho(Wr(\mathcal{V},\mathcal{U}))\,ww^{-1}a_1 \dots a_{i-1} b_j \dots b_m \text{ and } a_i = b_j \\ & \Leftrightarrow \quad w\,\rho(\mathcal{U}) ww^{-1} a_1 \dots a_{i-1} b_j \dots b_m , \\ & \qquad d_{\mathcal{U}}(w)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(ww^{-1}a_1 \dots a_{i-1} b_j \dots b_m) \text{ and } a_i = b_j \\ & \qquad\quad \text{(Theorem 1.3).} \end{aligned}$$

Next, observe that

$$\begin{aligned} e_i = f_j \quad & \Leftrightarrow \quad d_{\mathcal{U}}(w) \le_{\mathcal{V}} c_1 \dots c_{i-1} d_j \dots d_m \text{ and } c_i = d_j \\ & \qquad\quad \text{(Lemma 1.2 since } d_{\mathcal{U}}(w)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(v)) \\ & \Leftrightarrow \quad (c_1 \dots c_n)\,\rho(\mathcal{V})\,(c_1 \dots c_n)(c_1 \dots c_n)^{-1} c_1 \dots c_{i-1} d_j \dots d_m \\ & \qquad\quad \text{and } c_i = d_j; \\ & \Leftrightarrow \quad (c_1 \dots c_n)\,\rho(\mathcal{V})\,(c_1 \dots c_n)(c_1 \dots c_n)^{-1} c_1 \dots c_{i-1} d_j \dots d_m \\ & \qquad w \le_{\mathcal{U}} a_1 \dots a_{i-1} b_j \dots b_m \text{ and } a_i = b_j \\ & \qquad\quad \text{(Lemma 1.2 since } w\,\rho(\mathcal{U})\,v) \\ & \Leftrightarrow \quad (c_1 \dots c_n)\,\rho(\mathcal{V})\,(c_1 \dots c_n)(c_1 \dots c_n)^{-1} c_1 \dots c_{i-1} d_j \dots d_m , \\ & \qquad w\,\rho(\mathcal{U})\,ww^{-1} a_1 \dots a_{i-1} b_j \dots b_m \text{ and } a_i = b_j . \end{aligned}$$

If $w\,\rho(\mathcal{U})\,ww^{-1}a_1 \dots a_{i-1}b_j \dots b_m$ then w, v and $ww^{-1}a_1 \dots a_{i-1}b_j \dots b_m$ each label $s-e$ paths in $\overline{\Gamma_{\mathcal{U}}(w)}$. The path starting at s labelled $ww^{-1}a_1 \dots a_{i-1}b_j \dots b_m$ must have secondary labels $(c_1 \dots c_n)(c_1 \dots c_n)^{-1}c_1 \dots c_{i-1}$, since $\overline{\Gamma_{\mathcal{U}}(w)}$ is deterministic. Likewise, the path labelled $b_j \dots b_m$ ending at e must have secondary labels $d_j \dots d_m$ since the path labelled v ending at e has secondary labels $d_1 \dots d_m$. It then follows that

$$d_{\mathcal{U}}(ww^{-1}a_1 \dots a_{i-1}b_j \dots b_m) = (c_1 \dots c_n)(c_1 \dots c_n)^{-1}c_1 \dots c_{i-1}d_j \dots d_m ,$$

if $v\,\rho(\mathcal{U})\,w\rho(\mathcal{U})\,ww^{-1}a_1 \dots a_{i-1}b_j \dots b_m$. Consequently, $g_i = h_j$ if and only if $e_i = f_j$.

2) We proceed in a similar manner:

$$\begin{aligned} g_i = h_j^{-1} \quad & \Leftrightarrow \quad w \le_{Wr(\mathcal{V},\mathcal{U})} a_1 \dots a_i b_j \dots b_m \text{ and } a_i = b_j^{-1}; \\ & \qquad\quad \text{(Lemma 1.2 since } w\,\rho(Wr(\mathcal{V},\mathcal{U}))v) \\ & \Leftrightarrow \quad w\,\rho(Wr(\mathcal{V},\mathcal{U}))\,ww^{-1}a_1 \dots a_i b_j \dots b_m \text{ and } a_i = b_j^{-1}; \\ & \Leftrightarrow \quad w\,\rho(\mathcal{U}) ww^{-1} a_1 \dots a_i b_j \dots b_m , \\ & \qquad d_{\mathcal{U}}(w)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(ww^{-1}a_1 \dots a_i b_j \dots b_m) \text{ and } a_i = b_j^{-1} \\ & \qquad\quad \text{(Theorem 1.3).} \end{aligned}$$

Also,

$$\begin{aligned} e_i = f_j^{-1} \quad & \Leftrightarrow \quad d_{\mathcal{U}}(w) \le_{\mathcal{V}} c_1 \dots c_i d_j \dots d_m \text{ and } c_i = d_j^{-1} \\ & \qquad\quad \text{(Lemma 1.2 since } d_{\mathcal{U}}(w)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(v)) \\ & \Leftrightarrow \quad (c_1 \dots c_n)\,\rho(\mathcal{V})\,(c_1 \dots c_n)(c_1 \dots c_n)^{-1} c_1 \dots c_i d_j \dots d_m \\ & \qquad\quad \text{and } c_i = d_j^{-1} \end{aligned}$$

$$\begin{aligned}&\Leftrightarrow \quad (c_1\dots c_n)\,\rho(\mathcal{V})\,(c_1\dots c_n)(c_1\dots c_n)^{-1}c_1\dots c_id_j\dots d_m\\ &\quad\ w\le_{\mathcal{U}} a_1\dots a_ib_j\dots b_m \text{ and } a_i=b_j^{-1}\\ &\qquad\quad \text{(Lemma 1.2 since } w\,\rho(\mathcal{U})\,v\text{)}\\ &\Leftrightarrow \quad (c_1\dots c_n)\,\rho(\mathcal{V})\,(c_1\dots c_n)(c_1\dots c_n)^{-1}c_1\dots c_id_j\dots d_m,\\ &\quad\ w\,\rho(\mathcal{U})\,ww^{-1}a_1\dots a_ib_j\dots b_m \text{ and } a_i=b_j^{-1}.\end{aligned}$$

We have by hypothesis that $w\,\rho(\mathcal{U})\,v$ and so if $w\,\rho(\mathcal{U})\,ww^{-1}a_1\dots a_ib_j\dots b_m$ we must have, as in 1), that $d_{\mathcal{U}}(ww^{-1}a_1\dots a_ib_j\dots b_m)=(c_1\dots c_n)(c_1\dots c_n)^{-1}c_1\dots c_id_j\dots d_m$. It follows that $g_i=h_j^{-1}$ if and only if $e_i=f_j^{-1}$.

The proofs of 3), 4), 5) and 6) are similar, noting the remark immediately following Lemma 1.2. ■

Theorem 2.2. *The operator Wr is associative.*

Proof. Let $\mathcal{W}$, $\mathcal{V}$ and $\mathcal{U}$ be varieties of inverse semigroups. For any $v,\ w\in F\mathcal{I}(X)$,

$$\begin{aligned} v\,\rho(Wr(\mathcal{W},Wr(\mathcal{V},\mathcal{U})))\,w \ &\Leftrightarrow v\,\rho(Wr(\mathcal{V},\mathcal{U}))\,w \text{ and}\\ &\quad\ d_{Wr(\mathcal{V},\mathcal{U})}(v)\,\rho(\mathcal{W})d_{Wr(\mathcal{V},\mathcal{U})}(w)\\ &\quad\ \text{(Theorem 1.3)}\\ &\Leftrightarrow v\,\rho(\mathcal{U})\,w, d_{\mathcal{U}}(v)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(w) \text{ and}\\ &\quad\ d_{Wr(\mathcal{V},\mathcal{U})}(v)\,\rho(\mathcal{W})d_{Wr(\mathcal{V},\mathcal{U})}(w)\\ &\quad\ \text{(Theorem 1.3)}\\ &\Leftrightarrow v\,\rho(\mathcal{U})\,w, d_{\mathcal{U}}(v)\,\rho(\mathcal{V})\,d_{\mathcal{U}}(w) \text{ and}\\ &\quad\ d_{\mathcal{V}}(d_{\mathcal{U}}(v))\,\rho(\mathcal{W})\,d_{\mathcal{V}}(d_{\mathcal{U}}(w))\\ &\quad\ \text{(Lemma 2.1)}\\ &\Leftrightarrow v\,\rho(\mathcal{U})\,w \text{ and}\\ &\quad\ d_{\mathcal{U}}(v)\,\rho(Wr(\mathcal{W},\mathcal{V}))\,d_{\mathcal{U}}(w)\\ &\quad\ \text{(Theorem 1.3)}\\ &\Leftrightarrow v\,\rho(Wr(Wr(\mathcal{W},\mathcal{V}),\mathcal{U}))\,w\,.\end{aligned}$$

Therefore, $\rho(Wr(\mathcal{W},Wr(\mathcal{V},\mathcal{U})))=\rho(Wr(Wr(\mathcal{W},\mathcal{V}),\mathcal{U}))$ and as a consequence, $Wr(\mathcal{W},Wr(\mathcal{V},\mathcal{U}))=Wr(Wr(\mathcal{W},\mathcal{V}),\mathcal{U})$. Thus, the operator Wr is associative. ■

Theorem 2.3. *$\mathcal{L}(\mathcal{I})$ is a monoid with zero under the operation Wr.*

Proof. That $(\mathcal{L}(\mathcal{I}),Wr)$ is a semigroup is a consequence of Theorem 2.2. The trivial variety $\mathcal{T}$ is an identity for $(\mathcal{L}(\mathcal{I}),Wr)$ and so $(\mathcal{L}(\mathcal{I}),Wr)$ is a monoid. The variety $\mathcal{I}$ of all inverse semigroups is a zero of the monoid $(\mathcal{L}(\mathcal{I}),Wr)$. ■

In light of the results obtained thus far, instead of writing $Wr(\mathcal{U},\mathcal{V})$ we will denote this product by $\mathcal{U}\mathcal{V}$.

We conclude this section with some results concerning generators of varieties of the form $\mathcal{U}\mathcal{V}$.

Theorem 2.4. *$F(\mathcal{UV})(X)$ can be embedded in $F(\mathcal{U})\, wr\, (F(\mathcal{V}), F(\mathcal{V}))$.*

Proof. Set $\rho = \rho(\mathcal{UV})$. Let $X = \{x_i : i \in \omega\}$ and $U = \cup\{x_{i_u} : u \in F(\mathcal{V})x_i^{-1}\rho(\mathcal{V})\}$, where the union is over all $i \in \omega$. Define

$$\Theta : F(\mathcal{UV})(X) \to F(\mathcal{U})(Y)\, wr\, (F(\mathcal{V})(X), F(\mathcal{V})(X))$$

as follows: for each $i \in \omega$, set

$$(x_i\rho)\Theta = (\psi_i, \beta_i)\ .$$

Here β_i corresponds to $x_i\rho(\mathcal{V})$; that is,

$$\begin{aligned} \mathbf{d}\beta_i &= F(\mathcal{V})(x_i^{-1}\rho(\mathcal{V})), & \\ u\beta_i &= ux_i\rho(\mathcal{V}) & (u \in \mathbf{d}\beta_i), \\ u\psi_i &= x_{i_u}\rho(\mathcal{U}) & (u \in \mathbf{d}\beta_i). \end{aligned}$$

It is immediate that Θ maps $\{x_i\rho : i \in \omega\}$ into $F(\mathcal{U})(Y)\, wr\, (F(\mathcal{V})(X), F(\mathcal{V})(X))$. Since $F(\mathcal{U})(Y)\, wr\, (F(\mathcal{V})(X), F(\mathcal{V})(X))$ is a member of $\mathcal{UV}$ and $F(\mathcal{UV})(X)$ is $\mathcal{UV}$-free, we let Θ be the unique extension of Θ thus far defined, to $F(\mathcal{UV})$. The tedious and arduous task of verifying that Θ is indeed an embedding is left to the reader. ∎

Corollary 2.5. *For any pair of varieties $\mathcal{U}$ and $\mathcal{V}$ of inverse semigroups, the variety $\mathcal{UV}$ is generated by $F(\mathcal{U})\, wr\, (F(\mathcal{V}), F(\mathcal{V}))$.*

Proof. This is an immediate consequence of Theorem 2.4 and the definition of the product operator Wr. ∎

In fact, as was shown in [3; 2.8], if S generates the variety $\mathcal{U}$, then $\mathcal{UV}$ is generated by $S\, wr\, (F(\mathcal{V}), F(\mathcal{V}))$.

The following result will be used in the next section.

Proposition 2.6. *Let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be varieties of inverse semigroups. Then $(\mathcal{U} \vee \mathcal{V})\mathcal{W} = \mathcal{UW} \vee \mathcal{VW}$.*

Proof. Set $\rho = \rho((\mathcal{U} \vee \mathcal{V})\mathcal{W})$. Then for any $u, v \in (X \cup X^{-1})^+$ we have

$$\begin{aligned} u\,\rho\, v &\Leftrightarrow u\,\rho(\mathcal{W})\, v \text{ and } d_{\mathcal{W}}(u)\,\rho(\mathcal{U} \vee \mathcal{V})\, d_{\mathcal{W}}(v) \quad \text{(Theorem 1.3)} \\ &\Leftrightarrow u\,\rho(\mathcal{W})\, v \text{ and } d_{\mathcal{W}}(u)\,\rho(\mathcal{U}) \cap \rho(\mathcal{V})\, d_{\mathcal{W}}(v) \\ &\Leftrightarrow u\,\rho(\mathcal{W})\, v \text{ and } d_{\mathcal{W}}(u)\,\rho(\mathcal{U})\, d_{\mathcal{W}}(v) \text{ and } d_{\mathcal{W}}(u)\rho(\mathcal{V})\, d_{\mathcal{W}}(v) \\ &\Leftrightarrow u\,\rho(\mathcal{UW})\, v \text{ and } u\,\rho(\mathcal{VW})\, v \quad \text{(Theorem 1.3)} \\ &\Leftrightarrow u\,\rho(\mathcal{UW}) \cap \rho(\mathcal{VW})\, v \\ &\Leftrightarrow u\,\rho\,((\mathcal{UW}) \vee (\mathcal{VW}))\, v\ . \end{aligned}$$

∎

3. The variety $\mathcal{SV}$

The principal result of this section is the following. For any variety $\mathcal{V}$ of inverse semigroups, $\mathcal{SV}$ is the largest variety of inverse semigroups which satisfies the equations $w = w^2$ whenever $\mathcal{V}$ satisfies $w = w^2$. Throughout this section we will use the following convention. If $w \in (X \cup X^{-1})^+$ and $\mathcal{V}$ is a variety of inverse semigroups, we will write $w_{\mathcal{V}}$ to denote $w\rho(\mathcal{V})$.

Theorem 3.1. *Let $\mathcal{U} \subseteq \mathcal{V}$ be varieties of inverse semigroups and let ρ be the congruence on $F\mathcal{V}(X)$ such that $F\mathcal{V}(X)/\rho \cong F\mathcal{U}(X)$. Then ρ is idempotent pure if and only if for every $w \in (X \cup X^{-1})^+$, $\Gamma_{\mathcal{V}}(w)$ is embeddable in $\Gamma_{\mathcal{U}}(w)$ (and this embedding sends roots to roots).*

Proof. Suppose that ρ is idempotent pure and let $w \in (X \cup X^{-1})^+$. Define a map ϕ on $\mathcal{R}_{w_\mathcal{V}}$, the set of vertices of $\Gamma_\mathcal{V}(w)$, by setting $v\phi = v\rho$. Green's relation $\mathcal{R}$ is preseved under homomorphism so ϕ maps $\mathcal{R}_{w_\mathcal{V}}$ into $\mathcal{R}_{w_\mathcal{V}\rho}$, which is the vertex set of $\Gamma_\mathcal{U}(w)$ since, for any $v \in (X \cup X^{-1})^+$, $v_\mathcal{V}\rho = v_\mathcal{U}$. If (v_1, x, v_2) is an edge in $\Gamma_\mathcal{V}(w)$ then $v_1 x_\mathcal{V} = v_2$ and so $(v_1\rho)(x_\mathcal{V}\rho) = (v_2\rho)$. But this means that $(v_1\phi)x_\mathcal{U} = v_2\phi$ from which it follows that $(v_1\phi, x, v_2\phi)$ is an edge in $\Gamma_\mathcal{U}(w)$. Therefore, ϕ is a graph homomorphism. Also, ϕ maps $ww_\mathcal{V}^{-1}$ to $ww_\mathcal{U}^{-1}$ and $w_\mathcal{V}$ to $w_\mathcal{U}$ and so ϕ maps the roots of $\Gamma_\mathcal{V}(w)$ to the roots of $\Gamma_\mathcal{U}(w)$.

Suppose that $v_1\phi = v_2\phi$ for some $v_1, v_2 \in \mathcal{R}_{w_\mathcal{V}}$. Then $v_1 \,\rho\, v_2$ and so $v_1 = v_2$ since $\rho \cap \mathcal{R} = \varepsilon$ whenever ρ is idempotent pure. Thus, ϕ is an embedding of $\Gamma_\mathcal{V}(w)$ into $\Gamma_\mathcal{U}(w)$.

Conversely, suppose that $\Gamma_\mathcal{V}(w)$ is embeddable in $\Gamma_\mathcal{U}(w)$ for every $w \in (X \cup X^{-1})^+$. Let $e, a \in (X \cup X^{-1})^+$ be such that $e_\mathcal{V} = e_\mathcal{V}^2$ and $e_\mathcal{V} \,\rho\, a_\mathcal{V}$. Then $a_\mathcal{U} = aa_\mathcal{U}^{-1}$. If ϕ is the embedding of $\Gamma_\mathcal{V}(a)$ into $\Gamma_\mathcal{U}(a)$ then $a_\mathcal{V}\phi = a_\mathcal{U}$ and $aa_\mathcal{V}^{-1}\phi = aa_\mathcal{U}^{-1}$, since ϕ maps roots to roots. Since ϕ is one-to-one on the vertices of $\Gamma_\mathcal{V}(a)$ and $a_\mathcal{U} = aa_\mathcal{U}^{-1}$, we must have that $a_\mathcal{V} = aa_\mathcal{V}^{-1}$ and so ρ is idempotent pure. ■

Lemma 3.2. *Let $\mathcal{U} \subseteq \mathcal{V}$ be varieties of inverse semigroups and suppose that $\mathcal{U}^K = \mathcal{V}^K$. If ρ is the congruence on $F\mathcal{V}(X)$ such that $F\mathcal{V}(X)/\rho \cong F\mathcal{U}(X)$ then ρ is idempotent pure.*

Proof. Let $w, a \in (X \cup X^{-1})^+$ be such that $w_\mathcal{V} = w_\mathcal{V}^2$ and $w_\mathcal{V} \,\rho\, a_\mathcal{V}$. Then $a_\mathcal{V} \,\rho\, a_\mathcal{V}^2$; that is, $a_\mathcal{U} = a_\mathcal{U}^2$. Hence $a_\mathcal{V} = a_\mathcal{V}^2$, since $\ker \rho(\mathcal{U}) = \ker \rho(\mathcal{V})$. Consequently, ρ is idempotent pure. ■

Theorem 3.3. *Let $\mathcal{V}$ be a variety of inverse semigroups. Then $\mathcal{S}\mathcal{V} = \mathcal{V}^K = \langle \mathcal{S} \circ \mathcal{V} \rangle$.*

Proof. First of all, for any $w \in (X \cup X^{-1})^+$, $w \,\rho(\mathcal{S}\mathcal{V})\, w^2$ if and only if $w \,\rho(\mathcal{V})\, w^2$ and $d_\mathcal{V}(w) \,\rho(\mathcal{S})\, d_\mathcal{V}(w^2) = d_\mathcal{V}(w)^2$, by Theorem 1.3 with the last equality by Proposition 1.1 (a). Thus, $w \,\rho(\mathcal{S}\mathcal{V})\, w^2$ if and only if $w \,\rho(\mathcal{V})\, w^2$ and so $\mathcal{S}\mathcal{V} \subseteq \mathcal{V}^K$.

Let ρ_1, ρ_2 be the congruences on $F\mathcal{V}^K(X)$ such that

$$F\mathcal{V}^K(X)/\rho_1 \cong F\mathcal{S}\mathcal{V}(X)$$
$$F\mathcal{V}^K(X)/\rho_2 \cong F\mathcal{V}(X)$$

and let ρ_3 be the congruence on $F\mathcal{S}\mathcal{V}(X)$ such that

$$F\mathcal{S}\mathcal{V}(X)/\rho_3 \cong F\mathcal{V}(X)\,.$$

From the preceding lemma we obtain that ρ_1, ρ_2 and ρ_3 are idempotent pure and so, by the theorem above, for all $w \in (X \cup X^{-1})^+$, $\Gamma_{\mathcal{V}^K}(w)$ is embedable in $\Gamma_{\mathcal{S}\mathcal{V}}(w)$ which in turn is embeddable in $\Gamma_\mathcal{V}(w)$. Let $w, u \in (X \cup X^{-1})^+$ be such that $w_{\mathcal{V}^K} = (w_{\mathcal{V}^K})^2$, $u_{\mathcal{V}^K} = (u_{\mathcal{V}^K})^2$ and $w_{\mathcal{V}^K} \,\rho_1\, u_{\mathcal{V}^K}$. Then $w \,\rho(\mathcal{S}\mathcal{U})\, u$ and, by Theorem 1.3, $w_\mathcal{V} = u_\mathcal{V}$ and $c(d_\mathcal{V}(w)) = c(d_\mathcal{V}(u))$. It follows that $\Gamma_\mathcal{V}(w) = \Gamma_\mathcal{V}(u)$

and both u and w label $ww_{\mathcal{V}}^{-1} - w_{\mathcal{V}}$ paths in $\Gamma_{\mathcal{V}}(w)$. Moreover, the $ww_{\mathcal{V}}^{-1} - w_{\mathcal{V}}$ path labelled by u in $\Gamma_{\mathcal{V}}(w)$ uses only the edges in the $ww_{\mathcal{V}}^{-1} - w_{\mathcal{V}}$ path in $\Gamma_{\mathcal{V}}(w)$ labelled by w. Thus, u labels a $ww_{\mathcal{V}}^{-1} - w_{\mathcal{V}}$ path in the subgraph of $\Gamma_{\mathcal{V}}(w)$ consisting of the $ww_{\mathcal{V}}^{-1} - w_{\mathcal{V}}$ path labelled by w. Since $\Gamma_{\mathcal{V}^K}(w)$ is embeddable in $\Gamma_{\mathcal{V}}(w)$, this subgraph is embeddable in $\Gamma_{\mathcal{V}^K}(w)$ and so u labels a $ww_{\mathcal{V}^K}^{-1} - w_{\mathcal{V}^K}$ path in $\Gamma_{\mathcal{V}^K}(w)$. By the definition of Schützenberger graph, we have that $ww^{-1}u_{\mathcal{V}^K} = w_{\mathcal{V}^K}$ and hence that $u_{\mathcal{V}^K} \geq w_{\mathcal{V}^K}$. In a similar fashion we may demonstrate that $w_{\mathcal{V}^K} \geq u_{\mathcal{V}^K}$ and so obtain that $w_{\mathcal{V}^K} = u_{\mathcal{V}^K}$. As a consequence, we have that ρ_1 is an idempotent separating congruence. But the only idempotent pure and idempotent separating congruence on any inverse semigroup is the identical relation ε. Thus, $\rho_1 = \varepsilon$ and $F\mathcal{V}^K(X) \cong F\mathcal{S}\mathcal{V}(X)$. Therefore,

$$\mathcal{S}\mathcal{V} = \langle F\mathcal{S}\mathcal{V}(X)\rangle = \langle F\mathcal{V}^K(X)\rangle = \mathcal{V}^K .$$

Now let $S \in \mathcal{S} \circ \mathcal{V}$. Then there exists a congruence ρ on S such that $S/\rho \in \mathcal{V}$ and $e\rho \in \mathcal{S}$, for all idempotents e in S. Thus, $\ker\rho = E_S$ and $\rho$is idempotent pure. Suppose that $\mathcal{V}$ satisfies $w = w^2$. Then S/ρ satisfies $w = w^2$. Thus, if w is a word in the variables $x_1, \dots, x_n$, then for an arbitrary substitution $s_1, \dots, s_n$ of variables from S, the resulting element $w[s_1, \dots, s_n]$ belongs to $\ker\rho = E_S$. Therefore, S also satisfies the equation $w = w^2$, from which it follows that $S \in \mathcal{V}^K$. Consequently, $\mathcal{S} \circ \mathcal{V} \subseteq \mathcal{V}^K$ and so $\langle \mathcal{S} \circ \mathcal{V}\rangle \subseteq \mathcal{V}^K$. But $\mathcal{V}^K = \mathcal{S}\mathcal{V} \subseteq \langle \mathcal{S} \circ \mathcal{V}\rangle$, and so

$$\langle \mathcal{S} \circ \mathcal{V}\rangle = \mathcal{V}^K = \mathcal{S}\mathcal{V} .$$ ∎

Thus, in light of the results in the previous section, if Y is the two element semilattice, then $\mathcal{V}^K$ is generated by $Y\,wr\,(F\mathcal{V}(X), F\mathcal{V}(X))$.

The next result is a consequence of Theorem 3.3 and makes use of Theorem 3.6 of [3] which states that if $\mathcal{U}$ and $\mathcal{V}$ are varieties of inverse semigroups whose free objects on countably infinite X have solvable word problem, then so does the free object on X in $\mathcal{U}\mathcal{V}$.

Theorem 3.4. *If the variety $\mathcal{V}$ has decidable equational theory then so does $\mathcal{V}^K$. That is, if the free object in $\mathcal{V}$ on countably infinite X has solvable word problem, then so does the free object on X in $\mathcal{V}^K$.*

Proof. $F\mathcal{S}(X)$ has solvable word problem. By Theorem 3.6 of [2], $F\mathcal{S}\mathcal{V}(X)$ has solvable word problem if $F\mathcal{V}(X)$ does. The result now follows by Theorem 3.3. ∎

A variety $\mathcal{V}$ is *locally finite* if and only if every finitely generated member of $\mathcal{V}$ is finite. Equivalently, $\mathcal{V}$ is locally finite if and only if $F\mathcal{V}(X)$ is finite whenever X is finite. The following result is a consequence of Theorem 3.3 and a special case of Theorem 3.8 of [3] which states that the product variety $\mathcal{U}\mathcal{V}$ is locally finite if and only if both $\mathcal{U}$ and $\mathcal{V}$ are locally finite.

Theorem 3.5. *Let $\mathcal{V}$ be a variety of inverse semigroups. Then $\mathcal{V}$ is locally finite if and only if $\mathcal{V}^K$ is locally finite.*

Proof. Since $\mathcal{S}$ is locally finite, by Theorem 3.8 of [3], $\mathcal{S}\mathcal{V}$ is locally finite. The result now follows by Theorem 3.3. ∎

Proposition 3.6. *Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of inverse semigroups. Then*

a) $(\mathcal{U}\mathcal{V})^K = \mathcal{U}^K\mathcal{V}$;

b) *If* $\mathcal{U} = \mathcal{U}^K$ *then* $(\mathcal{U}\mathcal{V})^K = \mathcal{U}\mathcal{V}$;

c) $\mathcal{U}\mathcal{V}^K = (\mathcal{U} \vee \mathcal{S})\mathcal{V} = \mathcal{U}\mathcal{V} \vee \mathcal{V}^K$;

d) *If* $\mathcal{U}$ *is not a variety of groups, then* $\mathcal{U}\mathcal{V}^K = \mathcal{U}\mathcal{V}$;

e) $(\mathcal{U}\mathcal{V})^K = \mathcal{U}^K\mathcal{V}^K$.

Proof. a) $(\mathcal{U}\mathcal{V})^K = \mathcal{S}(\mathcal{U}\mathcal{V})$ by Theorem 3.3. Since Wr is associative, we have $\mathcal{S}(\mathcal{U}\mathcal{V}) = (\mathcal{S}\mathcal{U})\mathcal{V} = U^K\mathcal{V}$, where the second equality follows from Theorem 3.3.

b) If $\mathcal{U} = \mathcal{U}^K$, then $(\mathcal{U}\mathcal{V})^K = \mathcal{U}^K\mathcal{V}$, by part a) and $\mathcal{U}^K\mathcal{V} = \mathcal{U}\mathcal{V}$ by our assumption.

c) $\mathcal{U}\mathcal{V}^K = \mathcal{U}(\mathcal{S}\mathcal{V})$ by Theorem 3.3. By the associativity of Wr we have that $\mathcal{U}(\mathcal{S}\mathcal{V}) = (\mathcal{U}\mathcal{S})\mathcal{V}$ and $(\mathcal{U}\mathcal{S})\mathcal{V} = (\mathcal{U} \vee \mathcal{S})\mathcal{V}$ by Theorem 1.5. By Proposition 2.6, $(\mathcal{U} \vee \mathcal{S}) = \mathcal{U}\mathcal{V} \vee \mathcal{S}\mathcal{V} = \mathcal{U}\mathcal{V} \vee \mathcal{V}^K$.

d) If $\mathcal{U}$ is not a variety of groups, then $\mathcal{U} \vee \mathcal{S} = \mathcal{U}$. By part c) above, $\mathcal{U}\mathcal{V}^K = \mathcal{U}\mathcal{V}$.

e) For any variety $\mathcal{U}$ of inverse semigroups, $\mathcal{U}^K$ is not a variety of groups. By part d), $\mathcal{U}^K\mathcal{V}^K = \mathcal{U}^K\mathcal{V}$ and so, by part a), $(\mathcal{U}\mathcal{V})^K = \mathcal{U}^K\mathcal{V}^K$. ∎

If we let $\mathcal{U}$ and $\mathcal{V}$ be varieties of groups in Proposition 3.6 (e), then we have that $(\mathcal{U} \circ \mathcal{V})^K = \mathcal{U}^K\mathcal{V}^K$ (see Theorem 1.4). Thus, if the variety $\mathcal{X}$ has E-unitary covers over $\mathcal{U}$ and the variety $\mathcal{W}$ has E-unitary covers over $\mathcal{V}$ then $\mathcal{X} \subseteq \mathcal{U}^K$ and $\mathcal{W} \subseteq \mathcal{V}^K$ and so $\mathcal{X}\mathcal{W} \subseteq \mathcal{U}^K\mathcal{V}^K = (\mathcal{U} \circ \mathcal{V})^K$. Consequently, we have the following result which was proved in [3].

Theorem 3.7. *Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of groups and let $\mathcal{W}$ and $\mathcal{X}$ be varieties of inverse semigroups. If $\mathcal{X}$ has E-unitary covers over $\mathcal{U}$ and $\mathcal{W}$ has E-unitary covers over $\mathcal{V}$ then $\mathcal{X}\mathcal{W}$ has E-unitary covers over $\mathcal{U} \circ \mathcal{V} = \mathcal{U}\mathcal{V}$.*

Proposition 3.8. *Let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be varieties of inverse semigroups. If $\mathcal{W}$ is not a variety of groups then $\mathcal{U}^K = \mathcal{V}^K$ implies that $\mathcal{W}\mathcal{U} = \mathcal{W}\mathcal{V}$. If $\mathcal{W}$ is any variety of inverse semigroups then $\mathcal{U}^K = \mathcal{V}^K$ implies that $(\mathcal{W}\mathcal{U})^K = (\mathcal{W}\mathcal{V})^K$.*

Proof. If $\mathcal{U}^K = \mathcal{V}^K$ then $\mathcal{W}\mathcal{U}^K = \mathcal{W}\mathcal{V}^K$ and so, by Theorem 3.3 and the associativity of Wr, $(\mathcal{W}\mathcal{S})\mathcal{U} = (\mathcal{W}\mathcal{S})\mathcal{V}$. By Theorem 1.5, $\mathcal{W}\mathcal{S} = W \vee \mathcal{S} = \mathcal{W}$ since $\mathcal{W}$ is not a variety of groups. Therefore, $\mathcal{W}\mathcal{U} = \mathcal{W}\mathcal{V}$.

In light of the first statement of the Lemma, to prove the second statement we need only show it is true when $\mathcal{W}$ is a group variety. Now, by Theorem 3.3 and the associativity of Wr, $(\mathcal{W}\mathcal{U}^K)^K = \mathcal{W}^K\mathcal{U}^K = (\mathcal{W}\mathcal{U})^K$, where this last equality follows from Proposition 3.6 (e). As a result, $\mathcal{U}^K = \mathcal{V}^K$ implies that $(\mathcal{W}\mathcal{U})^K = (\mathcal{W}\mathcal{U}^K)^K = (\mathcal{W}\mathcal{V}^K)^K = (\mathcal{W}\mathcal{V})^K$. ∎

A consequence of Proposition 3.8 is that the semigroup of varieties of inverse semigroups does not possess unique factorization (unlike the semigroup of varieties of groups – see [5]).

Proposition 3.9. *Let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be varieties of inverse semigroups. If $\mathcal{U}^K = \mathcal{V}^K$ then $(\mathcal{U}\mathcal{W})^K = (\mathcal{V}\mathcal{W})^K$, but the converse is not true.*

Proof. By Proposition 3.6 (a), $(\mathcal{U}\mathcal{W})^K = \mathcal{U}^K\mathcal{W}$ and $\mathcal{U}^K\mathcal{W} = \mathcal{V}^K\mathcal{W}$ by the hypothesis. Again by Proposition 3.6 (a), $\mathcal{V}^K\mathcal{W} = (\mathcal{V}\mathcal{W})^K$, and so $(\mathcal{U}\mathcal{W})^K = (\mathcal{V}\mathcal{W})^K$. As for the converse, consider the variety $\mathcal{C}_2 = [x^2 = x^3]$. If S and (T, I) are inverse semigroups that satisfy the equation $x^2 = x^3$, then the inverse semigroup $S\,wr\,(T,I)$ satisfies $x^2 = x^3$, as a simple calculation will show. Thus, $\mathcal{C}_2$ is closed under the formation of wreath products and consequently, for any variety $\mathcal{V}$ contained in $\mathcal{C}_2$, the variety $\mathcal{V}\mathcal{C}_2 = \mathcal{C}_2$. Now $\mathcal{B}$ and $\mathcal{B}^1$ are both contained in $\mathcal{C}_2$, so that $\mathcal{B}\mathcal{C}_2 = \mathcal{C}_2 = \mathcal{B}^1\mathcal{C}_2$ but $\mathcal{B}^K = \mathcal{B} \neq (\mathcal{B}^1)^K$. ■

Thus, even the right ideal of $(\mathcal{L}(\mathcal{I}), Wr)$ consisting of varieties of the form $\mathcal{V}^K$ does not possess unique factorization. For example, $\mathcal{B}\mathcal{C}_2 = \mathcal{C}_2 = (\mathcal{B}^1)^K\mathcal{C}_2$, but $\mathcal{B}^K = \mathcal{B} \neq (\mathcal{B}^1)^K$.

Theorem 3.10. *Let $\mathcal{S}$ denote the variety of semilattices and let $\mathcal{B}$ denote the variety generated by the five element combinatorial Brandt semigroup. Define the relations Φ_1 and Φ_2 on $\mathcal{L}(\mathcal{I})$ as follows:*

$$\mathcal{U}\,\Phi_1\,\mathcal{V} \;\Leftrightarrow\; \mathcal{S}\mathcal{U} = \mathcal{S}\mathcal{V} \qquad (\mathcal{U},\mathcal{V} \in \mathcal{L}(\mathcal{I}))$$

$$\mathcal{U}\,\Phi_2\,\mathcal{V} \;\Leftrightarrow\; \mathcal{B}\mathcal{U} = \mathcal{B}\mathcal{V} \qquad (\mathcal{U},\mathcal{V} \in \mathcal{L}(\mathcal{I})).$$

Then Φ_1 and Φ_2 are semigroup congruences on $\mathcal{L}(\mathcal{I})$.

Proof. Clearly, both Φ_1 and Φ_2 are right semigroup congruences by the associativity of Wr. The relation Φ_1 is a left congruence by Proposition 3.8 and so it is a congruence.

Let $\mathcal{U}$ and $\mathcal{W}$ be varieties of inverse semigroups. If $\mathcal{W} \supseteq \mathcal{B}$, then $\mathcal{W}\vee\mathcal{B} = \mathcal{W}$, and so $\mathcal{B}\mathcal{W}\mathcal{U} = \mathcal{B}(\mathcal{W}\vee\mathcal{B})\mathcal{U}$. If $\mathcal{W}$ does not contain $\mathcal{B}$, then $\mathcal{W}$ is a variety of semilattices of groups and so $\mathcal{S}\mathcal{W} = \mathcal{S}(\mathcal{W}\vee\mathcal{B})$ and hence $\mathcal{S}\mathcal{W}\mathcal{U} = \mathcal{S}(\mathcal{W}\vee\mathcal{B})\mathcal{U}$. Thus, $\mathcal{B}\mathcal{S}\mathcal{W}\mathcal{U} = \mathcal{B}\mathcal{S}(\mathcal{W}\vee\mathcal{B})\mathcal{U}$, whence $\mathcal{B}\mathcal{W}\mathcal{U} = \mathcal{B}(\mathcal{W}\vee\mathcal{B})\mathcal{U}$, since $\mathcal{B}\mathcal{S} = \mathcal{B}$. Thus, in either case, we have by Theorem 1.5 that $\mathcal{B}\mathcal{W}\mathcal{U} = \mathcal{B}\mathcal{W}\mathcal{B}\mathcal{U}$. Now, if $\mathcal{U}\,\Phi_2\,\mathcal{V}$ then $\mathcal{B}\mathcal{U} = \mathcal{B}\mathcal{V}$ and therefore, $\mathcal{B}\mathcal{W}\mathcal{U} = \mathcal{B}\mathcal{W}\mathcal{B}\mathcal{U} = \mathcal{B}\mathcal{W}\mathcal{B}\mathcal{V} = \mathcal{B}\mathcal{W}\mathcal{V}$. It follows that Φ_2 is a semigroup congruence on $\mathcal{L}(\mathcal{I})$. ■

Corollary 3.11. *The following maps are semigroup homomorphisms of $\mathcal{L}(\mathcal{I})$ into $\mathcal{L}(\mathcal{I})$.*

$$\mathcal{U}\Phi_1 = \mathcal{S}\mathcal{U} \qquad (\mathcal{U} \in \mathcal{L}(\mathcal{I}))\ ,$$

$$\mathcal{U}\Phi_2 = \mathcal{B}\mathcal{U} \qquad (\mathcal{U} \in \mathcal{L}(\mathcal{I}))\ .$$

Proof. For any varieties $\mathcal{U}$ and $\mathcal{V}$ of inverse semigroups, $\mathcal{V}\,\Phi_1\,\mathcal{S}\mathcal{V}$ and so by Theorem 3.10, $\mathcal{U}\mathcal{V}\,\Phi_1\,\mathcal{U}\mathcal{S}\mathcal{V}$. Therefore, $\mathcal{S}\mathcal{U}\mathcal{V} = \mathcal{S}\mathcal{U}\mathcal{S}\mathcal{V}$ and the map Φ_1 is a homomorphism which quite clearly induces the congruence Φ_1 of Theorem 3.10. In the same way one shows that the map Φ_2 is a semigroup homomorphism which induces the congruence Φ_2 of Theorem 3.10. ■

Thus, the right ideal of $(\mathcal{L}(\mathcal{I}), Wr)$ consisting of varieties of the form $\mathcal{V}^K$ and the right ideal of $(\mathcal{L}(\mathcal{I}), Wr)$ consisting of varieties of the form $\mathcal{B}\mathcal{V}$ are also homomorphic images of the semigroup of varieties of inverse semigroups.

Theorem 3.3 deals with varieties which satisfy the same 'kernel identities'; that is, identities of the form $w = w^2$. We conclude with some companion results to Theorem 3.3 that deal with 'trace identities'.

Theorem 3.12. *Let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be varieties of inverse semigroups. If $\operatorname{tr}\rho(\mathcal{U}) = \operatorname{tr}\rho(\mathcal{V})$ then $\operatorname{tr}\rho(\mathcal{U}\mathcal{W}) = \operatorname{tr}\rho(\mathcal{V}\mathcal{W})$.*

Proof. Let v and w be idempotents in $F\mathcal{I}(X)$ and suppose that $v\,\rho(\mathcal{U}\mathcal{V})\,w$. Then, by Theorem 1.3, $v\,\rho(\mathcal{W})\,w$ and $d_{\mathcal{W}}(v)\,\rho(\mathcal{U})\,d_{\mathcal{W}}(w)$. By Proposition 1.1 (b), both $d_{\mathcal{W}}(v)$ and $d_{\mathcal{W}}(w)$ are idempotents of $F\mathcal{I}(Y)$. Consequently, $v\,\rho(\mathcal{W})\,w$ and $d_{\mathcal{W}}(v)\,\rho(\mathcal{V})\,d_{\mathcal{W}}(w)$, and so $v\,\rho(\mathcal{V}\mathcal{W})\,w$. Similarly, $v\,\rho(\mathcal{V}\mathcal{W})\,w$ implies that $v\,\rho(\mathcal{U}\mathcal{W})\,w$, and the result follows. ■

Corollary 3.13. *For any varieties $\mathcal{U}$ and $\mathcal{V}$ of inverse semigroups,*

$$\operatorname{tr}\rho(\mathcal{U}\mathcal{V}) = \operatorname{tr}\rho((\mathcal{U}\vee\mathcal{G})\mathcal{V}) \ ,$$

and

$$\mathcal{U}\mathcal{V}\vee\mathcal{G} = (\mathcal{U}\vee\mathcal{G})\mathcal{V} \ .$$

Proof. The first statement is a consequence of Theorem 3.12, since $\operatorname{tr}\rho(\mathcal{U}) = \operatorname{tr}\rho(\mathcal{U}\vee\mathcal{G})$ for any variety $\mathcal{U}$ of inverse semigroups [6; XII.2.2]. It then follows that $\mathcal{U}\mathcal{V}\vee\mathcal{G} = (\mathcal{U}\vee\mathcal{G})\mathcal{V}\vee\mathcal{G}$, by [6; XII.2.2]. Since $\mathcal{G}\subseteq(\mathcal{U}\vee\mathcal{G})\mathcal{V}$, we have that $(\mathcal{U}\vee\mathcal{G})\mathcal{V}\vee\mathcal{G} = (\mathcal{U}\vee\mathcal{G})\mathcal{V}$, and consequently that $\mathcal{U}\mathcal{V}\vee\mathcal{G} = (\mathcal{U}\vee\mathcal{G})\mathcal{V}$. ∎

As a result of Theorem 3.12, we have that tr is a right semigroup congruence on $\mathcal{L}(\mathcal{I})$. It is not a congruence however, as $\operatorname{tr}\rho(\mathcal{G}) = \operatorname{tr}\rho(\mathcal{T})$, but $\mathcal{I} = \mathcal{S}\mathcal{G} \neq \mathcal{S}\mathcal{T} = \mathcal{S}$.

For any variety $\mathcal{V}$ of inverse semigroups, we denote by $\mathcal{V}^T$ the largest variety satisfying the same trace identities as $\mathcal{V}$. That is, $\mathcal{V}^T$ is the largest variety $\mathcal{U}$ for which $\operatorname{tr}\rho(\mathcal{U}) = \operatorname{tr}\rho(\mathcal{V})$ and $\mathcal{V}^T = [uu^{-1} = vv^{-1} : \mathcal{V} \text{ satisfies } uu^{-1} = vv^{-1}]$.

Theorem 3.14. *Let $\mathcal{V}$ be a variety of inverse semigroups. Then*

$$\mathcal{V}^T = \mathcal{G}\mathcal{V} = \mathcal{G}\vee\mathcal{V} \ .$$

Proof. By Theorem 1.4, $\mathcal{G}\mathcal{V} = \mathcal{G}\circ\mathcal{V}$ and, by results found in [6; XII], $\mathcal{G}\circ\mathcal{V} = \mathcal{G}\vee\mathcal{V} = \mathcal{V}^T$. ∎

Let $\mathcal{U}$ and $\mathcal{V}$ be varieties of inverse semigroups. By Theorem 3.3, $\mathcal{S}\mathcal{U} = \mathcal{S}\mathcal{V}$ implies that $\ker\rho(\mathcal{U}) = \ker\rho(\mathcal{V})$, and by Corollary 3.13, $\mathcal{G}\mathcal{U} = \mathcal{G}\mathcal{V}$ implies that $\operatorname{tr}\rho(\mathcal{U}) = \operatorname{tr}\rho(\mathcal{V})$. Since a congruence is completely determined by its trace and kernel, we have that $\mathcal{S}\mathcal{U} = \mathcal{S}\mathcal{V}$ and $\mathcal{G}\mathcal{U} = \mathcal{G}\mathcal{V}$ implies that $\rho(\mathcal{U}) = \rho(\mathcal{V})$. Therefore, in the language of the semigroup of varieties of inverse semigroups, another way of saying that a variety is completely determined by its kernel and trace identities is that $\mathcal{U} = \mathcal{V}$ if and only if $\mathcal{G}\mathcal{U} = \mathcal{G}\mathcal{V}$ and $\mathcal{S}\mathcal{U} = \mathcal{S}\mathcal{V}$.

References

[1] Bales, J. L., *On product varieties of inverse semigroups*, J. Austr. Math. Soc. **28** (1979), 107–119.

[2] Cowan, D. F., *Wreath Products and Varieties of Inverse Semigroups*, Doctoral Thesis, Simon Fraser University, 1989.

[3] Cowan, D. F., *A class of varieties of inverse semigroups*, J. Alg. **141** (1991), 115–142.

[4] Houghton, C. H., *Embedding inverse semigroups in wreath products*, Glasgow Math. J. **17** (1976), 77–82.

[5] Neumann, H., "Varieties of Groups", Springer-Verlag, New York, 1967.

[6] Petrich, M., "Inverse semigroups", Wiley, New York, 1984.

[7] Stephen, J. B., *Presentations of inverse monids*, J. of Pure and Appl. Alg. **63** (1990), 81–112.

Department of Mathematics
and Statistics
Simon Fraser University
Burnaby, British Columbia
Canada, V5A 1S6

Received September 12, 1990
and in final form January 16, 1991

Semigroup Forum Vol. 45 (1992) 302–321

RESEARCH ARTICLE

Non autonomous evolution operators of hyperbolic type

Giuseppe Da Prato & Eugenio Sinestrari*

Communicated by R. Nagel

Abstract

In this paper we consider a Cauchy problem in a Banach space E : $u'(t) = A(t)u(t) + f(t)$, $t \in [t_0, T]$, $u(t_0) = u_0$, where $A(\cdot)$ is a family of linear operators in E which satisfy all the requirements of Kato's semigroup approach to the non autonomous hyperbolic equations except for the density of the common domains of $A(t)$. An application is given to a hyperbolic partial differential equation with discontinuous coefficients.

1. Introduction

Let E be a Banach space with norm $\|\cdot\|$. We will use in this paper the following Banach spaces of functions (endowed with their usual norms):

- $C(t_0, T; E)$: the space of all continuous functions $u : [t_0, T] \to E$.
- $C^n(t_0, T; E)$: the space of all n times continuously differentiable functions $u : [t_0, T] \to E$.
- $L^p(t_0, T; E)$: the space of all Bochner measurable functions $u : [t_0, T] \to E$ such that $\|u(\cdot)\| \in L^p(t_0, T)$; $1 \leq p < \infty$.
- $W^{1,p}(t_0, T; E)$: the set of all functions $u : [t_0, T] \to E$ such that there exists $v \in L^p(t_0, T; E)$, such that
$$u(t) = u(t_0) + \int_{t_0}^{t} v(s)ds, \ t \in [t_0, T].$$

If X is a space of functions $f : [t_0, T] \to \mathbb{R}$ then we set
$$X_{\#} := \{f \in X \mid f(t_0) = f(T)\}.$$

Let us consider the Cauchy problem
$$\begin{cases} u'(t) = Au(t) + f(t), \ t \in [t_0, T] \\ u(t_0) = u_0, \end{cases} \tag{1.1}$$
where $A : D \subset E \to E$ is a linear operator and $f : [t_0, T] \to E, u_0 \in E$ are given.

It is known (see e.g [5]) that the classical semigroup theory ensures the well posedness of problem (1.1) when A is the infinitesimal generator of a strongly continuous semigroup of bounded linear operators in E or, equivalently (by virtue of the Hille-Yosida theorem) when:

* Partially supported by the Italian National Project M. P. I. "Equazioni di Evoluzione e Applicazioni Fisico-Matematiche"

(i) $\overline{D} = E$,

(ii) *There exists* $M, \omega \in \mathbb{R}$ *such that if* $\lambda > \omega$, $R(\lambda, A) = (\lambda - A)^{-1} \in \mathcal{L}(E)$ *and* $\|(\lambda - \omega)^n R^n(\lambda, A)\|_{\mathcal{L}(E)} \leq M, \ \forall n \in \mathbb{N}$.

Here $\mathcal{L}(E)$ is the Banach space of all linear bounded operators in E. In this case one can prove (see e.g. [8]) the existence and uniqueness of a strict solution of (1.1) when $u_0 \in D$ and $f \in C^1(t_0, T; E)$ or $f, Af \in C(t_0, T; E)$. But in some applications to partial differential equations (see [4]) the operator A does not generate a semigroup because only hypothesis (ii) is satisfied: this fact led us to extend the above result by proving (see [3] and [4]) that also in this case, problem (1.1) has a strict solution when $f \in W^{1,1}(t_0, T; E)$, $u_0 \in D$ and $Au_0 + f(t_0) \in \overline{D}$ or when $f \in C(t_0, T; E), Af \in L^1(t_0, T; E), u_0 \in D$ and $Au_0 \in \overline{D}$ (note that some of these conditions are necessary for the existence of a strict solution and when $f = 0$ the results can be obtained by using the classical semigroup theory because A generates a strongly continuous semigroup in the space $\overline{D}$). In addition, estimates in the uniform norm for u, u' and Au as well as the existence and uniqueness of a weak solution (in the sense of Friedrichs) and of an integral solution were proved. Later some of these results have been demonstrated in different ways by T. Kato, M. G. Crandall (personal communications), H. Thieme (see [9]) and by W. Arendt, H. Kellermann and M. Hieber (see [1], [7]).

The method we used in [4] was based on an interpretation of (1.1) as a functional equation in the space $L^p(t_0, T; E)$ and on the use of the Yosida approximations of the time derivative considered as an operator in that space. In the present paper we will show that this approach turns out to be effective also in the time-dependent version of problem (1.1):

$$\begin{cases} u'(t) = A(t)u(t) + f(t), \ t \in [t_0, T] \\ u(t_0) = u_0, \end{cases} \tag{1.2}$$

which will be studied under the following hypotheses (which in the dense domain case are those of [6]):

Hypothesis 1.1. *For all* $t \in [t_0, T]$, $A(t) : D \to E$ *is a linear operator between the Banach spaces* D *(norm* $\|\cdot\|_D$*) and* E *(norm* $\|\cdot\|$*).*

Hypothesis 1.2. $D \subset E$ *and there exists* $c_0 > 0$ *such that for all* $t \in [t_0, T]$ *and* $x \in D$

$$c_0^{-1}\|x\|_D \leq \|x\| + \|A(t)x\| \leq c_0\|x\|_D.$$

Hypothesis 1.3. *There exists* $\omega, M \in \mathbb{R}$ *such that* $\rho(A(t)) \supset]\omega, +\infty[$, $t \in [t_0, T]$ *and for each* $n \in \mathbb{N}$ *we have*

$$\|R(\lambda, t_1, ..., t_n)\|_{\mathcal{L}(E)} \leq \frac{M}{(\lambda - \omega)^n}$$

when $t_0 \leq t_n \leq ... \leq t_1 \leq T$ *and* $\lambda > \omega$.

Here we denote the resolvent set of $A(t) : D \subset E \to E$ by $\rho(A(t))$ and set

$$R(\lambda; t_1, ..., t_n) = R(\lambda, A(t_1)) \cdots R(\lambda, A(t_n)).$$

Finally we shall (at least) suppose that

$$A(.) \in C([0,T]; \mathcal{L}(D; E)), \tag{1.3}$$

where $\mathcal{L}(D; E)$ is the Banach space of all linear bounded operators from D to E. But to prove the existence of the solutions we shall need some regularity properties for A (see Theorems 3.5, 4.2 and 5.1): these can be verified in the applications to differential problems under mild assumptions on the coefficients as we shall see in the example at the end of this paper (and in forthcoming papers devoted to applications of this theory).

Let us define the different kind of solutions of (1.2) which will be considered in this paper.

- A *strict solution* in L^p of (1.2) is a function $u \in W^{1,p}(t_0, T; E) \cap L^p(t_0, T; D)$ verifying (1.2) a.e. in $[t_0, T]$.

- A function $u \in L^p(t_0, T; E)$ is called an *F-solution* in L^p of (1.2) if, for each $k \in N$ there is $u_k \in W^{1,p}(t_0, T; E) \cap L^p(t_0, T; D)$ such that

$$\lim_{k \to \infty} [\|u_k(0) - u_0\| + \|u - u_k\|_p + \|u_k' - A(\cdot)u_k - f\|_p] = 0, \tag{1.4}$$

  where $\|\cdot\|_p = \|\cdot\|_{L^p(t_0,T;E)}$.

- A *strict solution* in C of (1.2) is a function $u \in C^1(t_0, T; E) \cap C(t_0, T; D)$ verifying (1.2) in $[t_0, T]$.

- A function $u \in C(t_0, T; E)$ is called an *F-solution* in C of (1.2) if, for each $k \in \mathbb{N}$ there is $u_k \in C^1(t_0, T; E) \cap C(t_0, T; D)$ such that

$$\lim_{k \to \infty} [\|u_k(0) - u_0\| + \|u - u_k\|_0 + \|u_k' - A(\cdot)u_k - f\|_0] = 0, \tag{1.5}$$

  where $\|\cdot\|_0 = \|\cdot\|_{C(t_0,T;E)}$.

Let us observe that a strict solution in C is also a strict solution in L^p and the same is true for the F-solutions. Moreover a strict solution is also an F- solution. Later we shall prove that an F- solution in L^p is continuous and $u(t_0) = u_0$ (see Theorem 2.5). Finally, let us observe that $u_0 \in \overline{D}$ is a necessary condition for the existence of any solution of the above type.

2. The approximating problem and the a priori estimates

We want to write the problem (1.2) as a functional equation in the space $L^p(t_0, T; E)$ by considering the time derivative as an operator in $L^p(t_0, T; E)$ with a domain which takes into account the initial condition $u(t_0) = u_0$ and which is also dense in $L^p(t_0, T; E)$: more precisely, (1.2) will be written as an equation in $L^p(t_0, T; E)$:

$$B(u - u_0) + \Lambda u + f = 0, \quad u \in D(B), \tag{2.1}$$

where $B : D(B) \subset L^p(t_0,T;E) \to L^p(t_0,T;E)$ and $\Lambda : D(\Lambda) \subset L^p(t_0,T;E) \to L^p(t_0,T;E)$ are defined as

$$\begin{cases} D(B) = W_0^{1,p}(t_0,T;E) = \{u \in W^{1,p}(t_0,T;E) \mid u(t_0) = 0\} \\ Bu = -u' \end{cases} \tag{2.2}$$

and

$$\begin{cases} D(\Lambda) & = L^p(t_0,T;D) \\ & = \{u \in L^p(t_0,T;E) \mid u(t) \in D \text{ a.e and } A(\cdot)u(\cdot) \in L^p(t_0,T;E)\} \\ (\Lambda u)(t) = A(t)u(t), & t \in [t_0,T] \text{ a.e..} \end{cases} \tag{2.3}$$

Here (and later) we will denote by u_0 the function $u(t) = u_0$ for $t \in [t_0,T]$. We have proved in Section 3 of [4] that B is the generator of a semigroup of contractions in $L^p(t_0,T;E)$, i.e.

$$\|\lambda R(\lambda, B)\|_{\mathcal{L}(L^p(t_0,T;E))} \leq 1 \text{ for } \lambda > 0,$$

$\rho(B) = \mathbb{C}$ and for each $\lambda \in \mathbb{C}$ and $u \in L^p(t_0,T;E)$

$$(R(\lambda, B)u)(t) = \int_{t_0}^{t} e^{-\lambda(t-s)}u(s)ds, \ t \in [t_0,T]. \tag{2.4}$$

As $p < \infty$ we have $\overline{D(B)} = L^p(t_0,T;E)$; hence the Yosida approximations B_n of B, $B_n = n^2R(n,B) - nI = nBR(n,B)$ verify the property

$$\lim_{n\to\infty} \|B_nu - Bu\|_{L^p(t_0,T;E)} = 0, \ u \in L^p(t_0,T;E). \tag{2.5}$$

For each $u \in L^p(t_0,T;E)$ and $t \in [t_0,T]$ a.e. we have

$$(B_nu)(t) = n^2 \int_{t_0}^{t} e^{-n(t-s)}u(s)ds - nu(t), \tag{2.6}$$

and for each $u \in D(B)$

$$\begin{cases} B_nu = -nR(n,B)u' = -n(R(n,B)u)', \\ (B_nu)' = B_nu'. \end{cases} \tag{2.7}$$

If $u \in C(t_0,T;E)$, then $B_nu \in C(t_0,T;E)$ and for each $t \in]t_0,T]$ we have

$$\|(nR(n,B)u))(t)\| \leq \|u\|_{C(t_0,t;E)}. \tag{2.8}$$

Let us observe that (1.3) implies the following: if $\lambda > \omega$ then $\lambda \in \rho(\Lambda)$ and for each $u \in L^p(t_0,T;E)$

$$(R(\lambda,\Lambda)u)(t) = R(\lambda, A(t))u(t), \ t \in [t_0,T] \text{ a.e.} \tag{2.9}$$

Now we want to solve a problem which approximates (1.2) and to give an estimate for its solutions: this will enable us to give also an a-priori estimate for the solutions of (1.2). To this end we need a result on problem (1.2) when $A(t)$ are bounded operators; its proof can be found e.g. in Chapter 3 of [2].

Lemma 2.1. *If $H \in C(t_0,T;\mathcal{L}(E)), \varphi \in L^p(t_0,T;E)$ and $z_0 \in E$, then the problem*

$$\begin{cases} z'(t) = H(t)z(t) + \varphi(t), \ t \in [t_0,T] \\ z(t_0) = z_0 \end{cases} \tag{2.10}$$

has a unique solution $z \in W^{1,p}(t_0,T;E)$ given by the formula

$$z(t) = U(t,t_0)z_0 + \int_{t_0}^{t} U(t,s)\varphi(s)ds, \ t \in [t_0,T], \tag{2.11}$$

where $U(t,s) \in \mathcal{L}(E)$ is defined for $t_0 \le s \le t \le T$ as

$$U(t,s) = I_E + \sum_{k=1}^{\infty} \int_{\Delta_k(s,t)} H(t_k)...H(t_1)dt_1...dt_k \tag{2.12}$$

with

$$\Delta_k(s,t) = \{(t_1,...,t_k) \in \mathbb{R}^k \mid \ s \le t_1... \le t_k \le t\}. \tag{2.13}$$

In addition, if $\varphi \in C(t_0,T;E)$ then $z \in C^1(t_0,T;E)$.

Moreover, when $H \in C^1(t_0,T;E)$, from $\varphi \in W^{1,p}(t_0,T;E)$, (respectively, $\varphi \in C^1(t_0,T;E)$), we deduce $z \in W^{2,p}(t_0,T;E)$, (respectively, $z \in C^2(t_0,T;E)$). Finally, if $H, \varphi \in C^k(t_0,T;E)$ then $z \in C^{k+1}(t_0,T;E), k = 1,2,\dots$.

Now we can solve a problem which approximates problem (2.1).

Theorem 2.2. *Given $f \in L^p(t_0,T;E)$ and $u_0 \in E$ there exists for each $n > \omega$ a unique $v_n \in L^p(t_0,T;D)$ being a solution of*

$$B_n(v_n - u_0) + \Lambda v_n + f = 0 \tag{2.14}$$

and the following estimate holds for $t \in [t_0,T]$ a.e.:

$$\|v_n(t)\| \le \frac{Mn}{n-\omega}\|u_0\|e^{\frac{n\omega}{n-\omega}(t-t_0)} + \frac{Mn^2}{(n-\omega)^2}\int_{t_0}^{t} e^{\frac{n\omega}{n-\omega}(t-s)}\|f(s)\|ds + \frac{M}{n-\omega}\|f(t)\|. \tag{2.15}$$

In addition, if $f \in C(t_0,T;E)$ then $v_n \in C(t_0,T;D)$, (2.15) *holds for each $t \in [t_0,T]$ and*

$$v_n(t_0) = R(n,A(t_0))(nu_0 + f(t_0)). \tag{2.16}$$

Finally, we have the following regularity results

$$A \in C^1(t_0,T;\mathcal{L}(D,E)), f \in W^{1,p}(t_0,T;E) \Rightarrow v_n \in W^{1,p}(t_0,T;D) \tag{2.17}$$

and

$$A \in C^k(t_0,T;\mathcal{L}(D,E)), f \in C^k(t_0,T;E) \Rightarrow v_n \in C^k(t_0,T;D) \tag{2.18}$$

for all $k \in \mathbb{N}$.

Proof. If $v_n \in L^p(t_0,T;D)$ is a solution of (2.14), by virtue of (2.6) we have for $t \in [t_0,T]$ a.e.

$$n^2 \int_{t_0}^t e^{-n(t-s)} v_n(s)ds + ne^{-n(t-t_0)}u_0 - (n - A(t))v_n(t) + f(t) = 0. \tag{2.19}$$

By applying $R(n, A(t))$ and setting

$$w_n(t) = \int_{t_0}^t e^{ns} v_n(s)ds, \; t \in [t_0, T] \tag{2.20}$$

we have that $w_n \in W^{1,p}(t_0,T;D)$ and w_n is a solution of (2.10), where

$$\begin{aligned} H(t) &= n^2 R(n, A(t)), \; z_0 = 0 \\ \varphi(t) &= R(n, A(t))(ne^{nt_0}u_0 + e^{nt}f(t)). \end{aligned} \tag{2.21}$$

By virtue of (2.11) we must have for $t \in [t_0,T]$

$$w_n(t) = \int_{t_0}^t U(t,s)R(n,A(s)) \left[ne^{nt_0}u_0 + e^{ns}f(s)\right] ds,$$

where

$$U(t,s) = I + \sum_{k=1}^{\infty} \int_{\Delta_k(s,t)} n^{2k} R(n, \sigma_k, ...\sigma_1) d\sigma_1...d\sigma_k$$

and $\Delta_k(s,t)$ is defined in (2.13). It follows that

$$\begin{aligned} w_n(t) \;\; = \;\; & \int_{t_0}^t \left\{ R(n,A(s)) + \sum_{k=1}^{\infty} \int_{\Delta_k(s,t)} n^{2k} R(n; \sigma_k, ...\sigma_1, s) d\sigma_1...d\sigma_k \right\} \\ & \cdot \left[ne^{nt_0}u_0 + e^{ns}f(s)\right] ds. \end{aligned} \tag{2.22}$$

So from (2.19) and (2.20) we obtain for $t \in [t_0, T]$ a. e.

$$\begin{aligned} v_n(t) \;\; &= \;\; R(n,A(t)) \left\{ n^2 e^{-nt} w_n(t) + ne^{-n(t-t_0)}u_0 + f(t) \right\} \\ &= \;\; R(n,A(t)) \left\{ ne^{-n(t-t_0)}u_0 + f(t) \right\} + n^2 e^{-nt} \\ & \quad \cdot \int_{t_0}^t \left\{ R(n;t,s) + \sum_{k=1}^{\infty} \int_{\Delta_k(s,t)} n^{2k} R(n; t, \sigma_k, ...\sigma_1, s) d\sigma_1...d\sigma_k) \right\} \\ & \quad \cdot [ne^{nt_0}u_0 + e^{ns}f(s)]ds. \end{aligned}$$

Since the measure of $\Delta_k(s,t)$ is $\frac{(t-s)^k}{k!}$ we deduce from (1.3)

$$\begin{aligned} \|v_n(t)\| \;\; \le \;\; & \frac{Mn}{n-\omega} e^{-n(t-t_0)} \|u_0\| + \frac{M}{n-\omega} \|f(t)\| + Mn^2 e^{-nt} \cdot \\ & \int_{t_0}^t \left\{ \left[\frac{1}{(n-\omega)^2} + \sum_{k=1}^{\infty} \frac{(t-s)^k}{k!} \frac{n^{2k}}{(n-\omega)^{k+2}} \right] \left(ne^{nt_0}\|u_0\| + e^{ns}\|f(s)\| \right) \right\} ds \end{aligned}$$

and (2.15) follows. Conversely, given $u_0 \in E$ and $f \in L^p(t_0,T;E)$ we get from Lemma 2.1 a solution $w_n \in W^{1,p}(t_0,T;E)$ of (2.10) with H, φ and u_0 given by (2.21), hence $v_n(t) = e^{-nt}w_n'(t)$ is a solution of (2.19).

The second part of the theorem is a consequence of the last part of Lemma 2.1; in particular when $v_n \in C(t_0,T;D)$ setting $t = t_0$ in (2.19) we obtain (2.16). ∎

It will be useful to deduce from (2.15) some estimates for v_n, which are uniform with respect to n.

Corollary 2.3. *Under the assumptions of Theorem 2.2 there are* $C_1 = C_1(\omega, M, T - t_0, p)$, $C_2 = C_2(\omega, M, T - t_0)$ *and* $C_3 = C_3(\omega, M, T - t_0)$ *such that for each* $n > \omega$

$$\|v_n\|_{L^p(t_0,T;E)} \le C_1(\|u_0\| + \|f\|_{L^p(t_0,T;E)}), \tag{2.23}$$

$$\|v_n(t)\| \le C_2 \left(\|u_0\| + \int_{t_0}^t \|f(s)\| ds + \frac{\|f(t)\|}{n} \right), \quad t \in [t_0, T], a.e. \tag{2.24}$$

If in addition $f \in C(t_0, T; E)$ *then we also have*

$$\|v_n(t)\| \le C_3(\|u_0\| + \|f\|_{C(t_0,T;E)}), \quad t \in [t_0, T]. \tag{2.25}$$

A first proof of the uniqueness of the solutions of (1.2) can be deduced from the fact that any F-solution in $L^p(t_0, T; E)$ of (1.2) is the limit in $L^p(t_0, T; E)$ of v_n as the following theorem shows. We omit its proof because it is similar to that of Theorem 4.2 of [4].

Theorem 2.4. *Given* $f \in L^p(t_0, T; E)$ *and* $u_0 \in E$, *let* v_n *be the solution of the approximating problem* (2.14). *If* u *is an* F*-solution in* $L^p(t_0, T; E)$ *of* (1.2) *we have*

$$\lim_{n\to\infty} \|u - v_n\|_{L^p(t_0,T;E)} = 0. \tag{2.26}$$

Another proof of uniqueness can be obtained from the following theorem which gives the a-priori estimate for the solutions of (1.2) and which is the main tool of our approach to the problem.

Theorem 2.5. *Let* u *be an* F*-solution of* (1.2). *After a possible modification of* u *on a set of measure* 0, *we have that* $u \in C(t_0, T; E)$, $u(t) \in \overline{D}$ *for each* $t \in [t_0, T]$, $u(t_0) = u_0$; *moreover if* $\{u_k\}$ *verifies* (1.4) *then* $\{u_k\}$ *converges to* u *in* $C(t_0, T; E)$. *Finally, the following a-priori estimates hold for each* $t \in [t_0, T]$

$$\|u(t)\| \le M \left\{ e^{\omega(t-t_0)} \|u_0\| + \int_{t_0}^t e^{\omega(t-s)} \|f(s)\| ds \right\} \tag{2.27}$$

and there exists $C_4 = C_4(\omega, M, T - t_0, p)$ *such that*

$$\|u(t)\| \le C_4 \left(\|u_0\| + \|f\|_{L^p(t_0,T;E)} \right). \tag{2.28}$$

Proof. Let us first suppose that u also is a strict solution in $L^p(t_0, T; E)$. Since we can write

$$B_n(u - u_0) + \Lambda u = u' - f + B_n(u - u_0)$$

(2.15) implies for $t \in [t_0, T]$ a.e.,

$$\begin{aligned} \|u(t)\| \le {} & \frac{Mn}{n - \omega} \|u_0\| e^{\frac{n\omega}{n-\omega}(t-t_0)} + \\ & \frac{Mn^2}{(n - \omega)^2} \int_{t_0}^t e^{\frac{n\omega}{n-\omega}(t-s)} \|u'(s) - f(s) + B_n(u - u_0)(s)\| ds + \\ & \frac{M}{n - \omega} \|u'(t) - f(t) + B_n(u - u_0)(t))\|. \end{aligned} \tag{2.29}$$

As $u-u_0 \in D(B)$ we deduce that $B_n(u-u_0)$ converges to $B(u-u_0)$ in $L^p(t_0,T;E)$ (see (2.5)) ; and so we can suppose that $\lim_{n\to\infty} B_n(u-u_0)(t) = -u'(t)$ for $t \in [t_0,T]$ a.e.. Now for $n \to \infty$ estimate (2.29) implies (2.27). Let us suppose that u is an F-solution and let $u_k \in W^{1,p}(t_0,T;E) \cap L^p(t_0,T;D)$ be such that (1.4) holds. Setting $f_k = u_k' - \Lambda u_k$ and $u_{0k} = u_k(0)$ we deduce from the first part of the proof that for $h,k \in \mathbb{N}$ and $t \in [t_0,T]$ we have

$$\|u_k(t) - u_h(t)\| \le M \left(e^{\omega(t-t_0)} \|u_{0k} - u_{0h}\| + \int_{t_0}^{t} e^{\omega(t-s)} \|f_k(s) - f_h(s)\| ds \right),$$

hence $\{u_k\}$ converges to u in $C(t_0,T;E)$. This implies $u(t) \in \overline{D}$ for each $t \in [t_0,T]$ and also (2.27). Finally (2.28) is a consequence of (2.27). ∎

Remark. By virtue of estimate (2.28), in order to prove the existence of an F-solution in $L^p(t_0,T;E)$ of (1.2) for each $u_0 \in \overline{D}$ and $f \in L^p(t_0,T;E)$ it will be sufficient to prove it for each (u_0, f) of a dense subset of $\overline{D} \times L^p(t_0,T;E)$.

3. Existence of F-solutions in L^p

To prove the existence of F-solutions in L^p for problem (1.2) we will first prove the existence of a strict solution u in the case $f \in C^3(t_0,T;E), f^k(t_0) = 0$, $k = 0,1,2,3$ and $u_0 = 0$. This will be obtained as the limit of the solutions v_n of its approximating problem. To prove the convergence of v_n to a strict solution we need uniform estimates for $v_n^{(k)}$ up to $k = 3$. Let us start with a lemma on perturbations of Volterra type of the approximating problem.

Lemma 3.1. *Let $\Gamma_n : C(t_0,T;E) \to C(t_0,T;E)$ be such that for each $u \in C(t_0,T;E)$ we have*

$$\|(\Gamma_n u)(t)\| \le C_5 \|u\|_{C(t_0,t;E)}, \; t \in]t_0,T], \tag{3.1}$$

where C_5 is independent of t, u and n. Let $w_n \in C(t_0,T;D)$ and set

$$B_n w_n + \Lambda w_n + \Gamma_n w_n = g_n. \tag{3.2}$$

Then there exist ν and C_6 depending on $C_5, M, \omega, T-t_0$ such that

$$\|(w_n(t)\| \le C_6 \|g_n\|_{C(t_0,t;E)}, \; t \in]t_0,T], \; n > \nu. \tag{3.3}$$

If $\Gamma_n : L^1(t_0,T;E) \to L^1(t_0,T;E)$ are such that for each $u \in L^1(t_0,T;E)$ we have

$$\|\Gamma_n\|_{L^1(t_0,t;E)} \le C_5 \|u\|_{L^1(t_0,t;E)}, \; t \in]t_0,T], \tag{3.4}$$

then given $v_n \in L^1(t_0,T;D)$ and setting g_n as in (3.2) we have

$$\|w_n\|_{L^1(t_0,t;E)} \le C_6 \|g_n\|_{L^1(t_0,T;E)}, \; t \in]t_0,T], n > \nu. \tag{3.5}$$

Proof. When $t_0 \leq \tau \leq t \leq T$ we get from (2.24) and (3.1)

$$\begin{aligned} \|w_n(\tau)\| &\leq C_2 C_5 \int_{t_0}^{\tau} \|w_n\|_{C(t_0,s;E)} ds \\ &+ C_2 \left(\int_{t_0}^{\tau} \|g_n(s)\| ds + \frac{1}{n} \|g_n\|_{C(t_0,\tau;E)} + \frac{C_5}{n} \|w_n\|_{C(t_0,\tau;E)} \right), \end{aligned}$$

and so setting

$$\begin{aligned} \varphi_n(t) &= \|w_n\|_{C(t_0,t;E)}, \quad a = 2C_2C_5 \\ \lambda_n(t) &= 2C_2[(T - t_0) + 1] \|g_n\|_{C(t_0,t;E)} \end{aligned}$$

and

$$\nu = \max\{2C_2C_5, \omega\}, \tag{3.6}$$

we deduce for $n > \nu$

$$\varphi_n(t) \leq a \int_{t_0}^{t} \varphi_n(s) ds + \lambda_n(t).$$

From Gronwall's inequality we get

$$\varphi_n(t) \leq \int_{t_0}^{t} \lambda_n(s) a e^{a(t-s)} ds + \lambda_n(t) \leq \lambda_n(t) e^{a(t-t_0)},$$

which is equivalent to (3.3) if we set

$$C_6 = 2C_2[(T - t_0) + 1] e^{2C_2C_5(T-t_0)}. \tag{3.7}$$

Let us prove the second part of the lemma. For $\tau \in [t_0, T]$ a.e. we have

$$\begin{aligned} \|w_n(\tau)\| &\leq C_2 \left(\int_{t_0}^{\tau} \|(\Gamma_n w_n)(s)\| ds + \int_{t_0}^{\tau} \|g_n(s)\| ds \right) \\ &+ \frac{C_2}{n} \left(\|(\Gamma_n w_n)(\tau)\| + \|g_n(\tau)\| \right), \end{aligned}$$

hence for $t \in]t_0, T]$

$$\begin{aligned} \int_{t_0}^{t} \|w_n(\tau)\| d\tau &\leq C_2 \left(\int_{t_0}^{t} d\tau \int_{t_0}^{\tau} \|(\Gamma_n w_n)(s)\| ds + \int_{t_0}^{t} d\tau \int_{t_0}^{\tau} \|g_n(s)\| ds \right) \\ &+ \frac{C_2}{n} \int_{t_0}^{t} \left(\|(\Gamma_n w_n)(\tau)\| + \|g_n(\tau)\| \right) d\tau], \end{aligned}$$

and so

$$\begin{aligned} \|w_n\|_{L^1(t_0,t;E)} &\leq C_2 \left(\int_{t_0}^{t} \|\Gamma_n w_n\|_{L^1(t_0,\tau;E)} d\tau + (t - t_0) \|g_n\|_{L^1(t_0,t;E)} \right) \\ &+ \frac{C_2}{n} \left(\|\Gamma_n w_n\|_{L^1(t_0,t;E)} + \|g_n\|_{L^1(t_0,t;E)} \right). \end{aligned}$$

Now if we set $\varphi_n(t) = \|w_n\|_{L^1(t_0,t;E)}$ the proof can be concluded as for the first part. ∎

Theorem 3.2. *Let $A \in C^3(t_0,T;\mathcal{L}(D;E))$ and $f \in C^3(t_0,T;E)$ be such that $f^{(k)}(t_0)=0$, $k=0,1,2,3$. For each $n>\omega$ there exists a unique $v_n \in C^3(t_0,T;D)$ being a solution of*

$$B_n v_n + \Lambda v_n + f = 0, \tag{3.8}$$

such that $v_n^{(k)}(t_0)=0$, $k=0,1,2,3$. Moreover, for each $k=1,2,3$ there exist ν_k and γ_k(depending on $M,\omega,T-t_0,C_0$ and on the seminorm $[A]_{C^k(t_0,T;\mathcal{L}(D;E))}$) such that for $n>\nu_k$ we have

$$\|v_n^{(k)}\|_{C^k(t_0,T;E)} \leq \gamma_k\|f\|_{C^k(t_0,T;E)}. \tag{3.9}$$

Proof. From Theorem 2.2 we deduce the existence of $v_n \in C^3(t_0,T;D)$, a solution of (3.8) and such that $v_n(t_0)=0$. Let us set for $t\in[t_0,T]$

$$\omega_1=\omega+1, \quad R(t)=(\omega_1-A(t))^{-1}. \tag{3.10}$$

In this proof we will consider the restriction of B to $C(t_0,T;E)$, so that (2.7) holds. Since $v_n \in D(B)$ we deduce from (2.8) and (3.8)

$$-n(nR(n,B)v_n')(t)-(\omega_1-A(t))v_n(t)+\omega_1 v_n(t)+f(t)=0 \tag{3.11}$$

and so

$$v_n(t)=-nR(t)(R(n,B)v_n')(t)+\omega_1 R(t)v_n(t)+R(t)f(t). \tag{3.12}$$

Differentiating (3.8) we have

$$(B_n v_n')(t)+A(t)v_n'(t)+A'(t)v_n(t)+f'(t)=0 \tag{3.13}$$

and so from (2.16) we deduce $v_n'(t_0)=0$. Setting

$$C(t)=A'(t)R(t) \tag{3.14}$$

and substituting v_n given by (3.12) in (3.13) we get

$$\begin{aligned}(B_n v_n')(t)+A(t)v_n'(t)-nC(t)(R(n,B)v_n')(t)+\\ \omega_1 C(t)\int_{t_0}^t v_n'(s)ds+C(t)f(t)+f'(t)=0.\end{aligned} \tag{3.15}$$

For brevity we will denote by $c=c(\|A'\|)$ a generic function of $\|A'\|_{C(t_0,T;\mathcal{L}(D,E))}$ and of $M,\omega,T-t_0$ and C_0, and analogously for $c=c(\|A'\|,\|A''\|)$, etc. If we set for each $v\in C(t_0,T;E)$ and $t\in[t_0,T]$

$$(\Gamma_n v)(t) \;=\; -nC(t)(R(n,B)v)(t)+\omega_1 C(t)\int_{t_0}^t v(s)ds \tag{3.16}$$

$$g(t) \;=\; C(t)f(t)+f'(t), \tag{3.17}$$

equation (3.15) can be written as

$$B_n v_n' + \Lambda v_n' + \Gamma_n v_n' + g = 0. \tag{3.18}$$

Now from (1.2) we get

$$\|C(t)\|_{\mathcal{L}(E)} \leq c(\|A'\|), \tag{3.19}$$

and so by virtue of (2.8)

$$\|(\Gamma_n v)(t)\| \le C(\|A'\|)\|v\|_{C(t_0,t;E)}, \quad t \in]t_0, T]. \tag{3.20}$$

Now we can apply Lemma 3.1 to equation (3.18) and deduce the existence of ν_1 and γ_1 depending on M, ω, $T - t_0$ and $\|A'\|_{C(t_0,T;\mathcal{L}(D;E))}$ such that (3.9) holds with $k = 1$. Let us estimate now v_n''. As $v_n' \in D(B)$ we can write (3.18) as

$$-n(R(n,B)v_n'')(t) - (\omega_1 - A(t))v_n'(t) + \omega_1 v_n'(t) + (\Gamma_n v_n')(t) + g(t) = 0. \tag{3.21}$$

Proceeding as above we obtain

$$v_n'(t) = -nR(t)(R(n,B)v_n'')(t) + \omega_1 R(t)v_n'(t) + R(t)(\Gamma_n v_n')(t) + R(t)g(t) = 0. \tag{3.22}$$

By differentiating (3.18) we get

$$(B_n v_n'')(t) + A(t)v_n''(t) + A'(t)v_n'(t) + (\Gamma_n v_n')'(t) + g'(t) = 0. \tag{3.23}$$

Now we have

$$C'(t) = A''(t)R(t) - C^2(t) \tag{3.24}$$

and

$$(\Gamma_n v)'(t) = -nC'(t)(R(n,B)v)(t) + (\Gamma_n v')(t) + \omega_1 C'(t)\int_{t_0}^t v(s)ds. \tag{3.25}$$

From (3.23)–(3.25) we deduce $v_n''(t_0) = 0$ (by virtue of (2.16)) and also

$$A'(t)v_n'(t) = (\Gamma_n v_n'')(t) + C(t)(\Gamma_n v_n')(t) + C(t)g(t). \tag{3.26}$$

Hence, substituting in (3.23) we get

$$(B_n v_n'')(t) + A(t)v_n''(t) + 2(\Gamma_n v_n'')(t) + (S_n v_n')(t) + h(t) = 0, \tag{3.27}$$

where we have set

$$(S_n v_n')(t) = C(t)(\Gamma_n v_n')(t) + C'(t)\left[(-nR(n,B)v_n')(t) + \omega_1 \int_{t_0}^t v_n'(s)ds\right], \tag{3.28}$$

$$h(t) = C(t)g(t) + g'(t). \tag{3.29}$$

From (3.19), (3.20) and (3.24) we deduce

$$\|(S_n v_n')(t)\| \le C(\|A'\|, \|A''\|)\|v_n'\|_{C(t_0,t;E)}, \tag{3.30}$$

$$\|h(t)\| \le C(\|A'\|, \|A''\|)\|f\|_{C^2(t_0,t;E)}. \tag{3.31}$$

Since we can apply Lemma 3.1 to equation (3.27), there exist ν_2 and γ_2 such that (3.9) with $k = 2$ is true. If we use the same procedure for v_n''' we obtain $v_n'''(t_0) = 0$ and (3.9) with $k = 3$ and suitable ν_3 and γ_3. ∎

Now we prove the existence of a strict solution in a particular case. This will enable us to prove the existence of the F-solution in $L^p(t_0, T; E)$.

Theorem 3.3. *Let $A \in C^3(t_0,T;\mathcal{L}(D;E))$ and $f \in C^3(t_0,T;E)$ be such that $f^{(k)}(t_0) = 0$, $k = 0,1,2,3$. Then problem (1.2) with $u_0 = 0$ has a unique strict solution in C. In addition, there exists C_7 depending on $M,\omega,T-t_0,C_0$ and $\|A'\|_{C(t_0,T;\mathcal{L}(D;E))}$ such that*

$$\|u\|_{C(t_0,T;D)} \le C_7\|f\|_{C^1(t_0,T;E)}. \tag{3.32}$$

Proof. If $\{v_n\}$ is the sequence defined in Theorem 3.2 we have for sufficiently large n and m

$$B_n(v_n - v_m) + \Lambda(v_n - v_m) + (B_n - B_m)v_m = 0 \tag{3.33}$$

and so from (2.25)

$$\|v_n - v_m\|_{C(t_0,T;E)} \le C_3\|(B_n - B_m)v_m\|_{C(t_0,T;E)}. \tag{3.34}$$

As $v_m \in D(B^2)$ we deduce from (3.9) with $k=2$ and (2.7)

$$\begin{aligned}\|(B_n - B_m)v_m\|_{C(t_0,T;E)} &= \|(m-n)R(n,B)R(m,B)B^2v_m\|_{C(t_0,T;E)} \\ &\le M^2\left|\frac{m-n}{(m-\omega)(n-\omega)}\right|\|v''_m\|_{C(t_0,T;E)} \\ &\le M^2\gamma_2\left|\frac{m}{m-\omega}-\frac{n}{n-\omega}\right|\|f\|_{C^2(t_0,T;E)}.\end{aligned} \tag{3.35}$$

Hence there exists

$$u = \lim_{n\to\infty} v_n \tag{3.36}$$

in $C(t_0,T;E)$, which also implies $u(t_0) = 0$. Now from (3.18) we obtain

$$B_n(v'_n - v'_m) + \Lambda(v'_n - v'_m) + \Gamma_n(v'_n - v'_m) + (B_n - B_m)v'_m + (\Gamma_n - \Gamma_m)v'_m = 0,$$

and so from (3.20) and Lemma 3.1 we deduce the existence of $c^\star$ such that for n and m sufficiently large we have

$$\|v'_n - v'_m\|_{C(t_0,T;E)} \le c^\star\left(\|(B_n - B_m)v'_m\|_{C(t_0,T;E)} + \|(\Gamma_n - \Gamma_m)v'_m\|_{C(t_0,T;E)}\right). \tag{3.37}$$

As $v'_m \in D(B^2)$ we deduce as above that

$$\|(B_n - B_m)v'_m\|_{C(t_0,T;E)} \le M^2\gamma_3\left|\frac{m}{m-\omega}-\frac{n}{n-\omega}\right|\|f\|_{C^3(t_0,T;E)}, \tag{3.38}$$

whereas from (3.16), (3.35) and (3.19) we have

$$\begin{aligned}\|(\Gamma_n - \Gamma_m)v'_m\|_{C(t_0,T;E)} &= \|C(t)\left[-nR(n,B)v'_m + mR(m,B)v'_m\right]\|_{C(t_0,T;E)} \\ &= \|C(t)(B_n - B_m)v_m\|_{C(t_0,T;E)} \\ &\le c\left|\frac{m}{m-\omega}-\frac{n}{n-\omega}\right|\|f\|_{C^2(t_0,T;E)},\end{aligned} \tag{3.39}$$

where c is independent on n and m. From (3.37)-(3.39) it follows that there exists $\lim_{n\to\infty} v'_n$ in $C(t_0,T;E)$; hence from (3.36) we get that $u \in C^1(t_0,T;E)$ and $u' = \lim_{n\to\infty} v'_n$ in $C(t_0,T;E)$. From (3.9) with $k=1$ we deduce

$$\|u'\|_{C(t_0,T;E)} \le \gamma_1\|f\|_{C^1(t_0,T;E)}. \tag{3.40}$$

Let us prove that u satisfies the equation (1.2). To this end let us observe that by virtue of (3.9) we have for sufficiently large n

$$\begin{aligned}\|B_n v_n - Bv_n\|_0 &= \|nR(n,B)Bv_n - Bv_n\|_0 \\ &= \|nR(n,B)Bv_n - (n-B)R(n,B)Bv_n\|_0 = \|BR(n,B)Bv_n\|_0 \\ &= \|R(n,B)B^2 v_n\|_0 \le \frac{M}{n-\omega}\|v_n''\|_0 \le \frac{M\gamma_2}{n-\omega}\|f\|_2,\end{aligned}$$

where $\|\cdot\|_0 = \|\cdot\|_{C(t_0,T;E)}$ and $\|\cdot\|_2 = \|\cdot\|_{C^2(t_0,T;E)}$.

Since

$$B_n v_n - Bu = B_n v_n - Bv_n - v_n' + u'$$

we deduce that

$$\lim_{n\to\infty} \|B_n v_n - Bu\|_{C(t_0,T;E)} = 0. \tag{3.41}$$

For each $t \in [t_0,T]$, $A(t) : D \subset E \to E$ is a closed operator, $v_n(t) \in D$ and $\lim_{n\to\infty} v_n(t) = u(t)$; from (3.8) and (3.41) we have

$$\lim_{n\to\infty} A(t)v_n(t) = \lim_{n\to\infty} [f(t) - (B_n v_n)(t)] = f(t) - u'(t),$$

which implies that $u(t) \in D$ and $Au(t) = f(t) - u'(t)$, i.e., u satisfies the equation in (1.2). From this fact and (3.40) we deduce (3.32). ■

We are now in position to prove the existence of the F-solutions in L^p. For this we need the following Hypothesis on A.

Hypothesis 3.4. *$A \in C(t_0,T;\mathcal{L}(D;E))$ and there exists $A_k \in C^3(t_0,T;\mathcal{L}(D;E))$ verifying Hypothesis 1.1–1.3 with C_0, ω, M independent of k and such that*

$$\lim_{k\to\infty} \|A - A_k\|_{C(t_0,T;\mathcal{L}(D;E))} = 0,$$

$$\sup_{k\in\mathbb{N}} \|A_k'\|_{C(t_0,T;\mathcal{L}(D;E))} < \infty.$$

Theorem 3.5. *Assume Hypothesis 3.4. Then, for any $f \in L^p(t_0,T;E)$ and $u_0 \in \overline{D}$, there exists a unique F-solution in L^p of problem (1.2).*

Proof. Let us first suppose that $f \in C^3(t_0,T;E)$ and $f^k(t_0) = u_0 = 0$, $k = 0,1,2,3$. By virtue of Theorem 3.3 there exists $u_k \in C^1(t_0,T;E) \cap C(t_0,T;D)$ as the solution of

$$\begin{cases} u_k'(t) = A_k(t)u_k(t) + f(t), \ t \in [t_0,T] \\ u_k(t_0) = 0. \end{cases}$$

Let us set for $t \in [t_0,T]$

$$f_k(t) = u_k'(t) - A(t)u_k(t) = [A_k(t) - A(t)]u_k(t) + f(t).$$

From (3.32) and Hypothesis 3.4 we deduce that

$$\|u_k\|_{C(t_0,T;D)} \le c\|f\|_{C^1(t_0,T;E)},$$

with c independent of k and so $\lim_{k\to\infty} \|f - f_k\|_{C(t_0,T;E)} = 0$; by virtue of (2.27) there exists $u = \lim_{k\to\infty} u_k$ in $C(t_0,T;E)$. This proves that problem (1.2) has an F-solution in L^p. By density (see Remark 2.6) the same conclusion holds when $f \in L^p(t_0,T;E)$ and $u_0 = 0$; if $u_0 \in D$ we can substitute f with $f - A(\cdot)u_0$ and then by density prove the general result. ■

4. Existence of strict solutions in L^p

In this section we will need the following hypothesis.

Hypothesis 4.1. *Let $A \in C^1(t_0, T; \mathcal{L}(D; E))$ verify Hypothesis 1.1–1.3. Moreover, for each $k \in \mathbb{N}$ there exists $A_k \in C^4(t_0, T; \mathcal{L}(D; E))$ verifying Hypothesis 1.1–1.3 with C_0, ω, M independent of k and such that*

$$\lim_{k\to\infty} \|A - A_k\|_{C^1(t_0,T;\mathcal{L}(D;E))} = 0.$$

Theorem 4.2. *Suppose that A satisfies Hypothesis 4.1. Then, for each $f \in W^{1,p}(t_0, T; E)$ and $u_0 \in D$ such that*

$$u_1 =: A(t_0)u_0 + f(t_0) \in \overline{D}, \tag{4.1}$$

there exists $u \in C^1(t_0, T; E) \cap C(t_0, T; D)$ as solution of

$$\begin{cases} u'(t) = A(t)u(t) + f(t), \ t \in [t_0, T] \\ u(t_0) = u_0. \end{cases} \tag{4.2}$$

In addition, u' is an F-solution in L^p of the same problem with $f(t)$ replaced by $A'(t)u(t) + f'(t)$ and u_0 by u_1.

Proof. We can suppose that $\omega < 0$ by substituting $A(t)$ with $A(t) - (\omega + 1)I$ and $f(t)$ with $f(t)e^{-(\omega+1)(t-t_0)}$. By virtue of Theorem 3.5 there exists an F-solution u in L^p of problem (4.2). From Theorem 2.4 we deduce

$$\lim_{n\to\infty} \|v_n - u\|_{L^p(t_0,T;E)} = 0, \tag{4.3}$$

where $v_n \in W^{1,p}(t_0, T; D)$ are solutions of the approximating problem (2.14), i.e.,

$$\begin{aligned} 0 &= B_n(v_n - u_0) + \Lambda v_n + f \\ &= B_n(v_n - v_n(t_0)) + B_n(v_n(t_0) - u_0) + \Lambda v_n + f. \end{aligned} \tag{4.4}$$

From (2.16) we deduce

$$v_n(t_0) - u_0 = R(n, A(t_0))u_1, \tag{4.5}$$

and so from (4.4) and (2.6) we get for $t \in [t_0, T]$

$$-(nR(n, B)v_n')(t) - ne^{-n(t-t_0)}R(n, A(t_0))u_1 + A(t)v_n(t) + f(t) = 0,$$

from which we deduce

$$\begin{aligned} v_n(t) &= A^{-1}(t)(nR(n, B)v_n')(t) \\ &\quad + ne^{-n(t-t_0)}A^{-1}(t)R(n, A(t_0))u_1 - A^{-1}(t)f(t). \end{aligned} \tag{4.6}$$

From (4.4) we have by differentiation

$$(B_n(v_n - u_0))'(t) + A(t)v_n'(t) + A'(t)v_n(t) + f'(t) = 0. \tag{4.7}$$

By substituting

$$(B_n(v_n - u_0))'(t) = (B_n(v_n' - u_1))(t) + e^{-n(t-t_0)} n A(t_0) R(n, A(t_0)) u_1$$

and (4.6) in (4.7) after setting

$$\begin{aligned} C(t) &= A'(t)A^{-1}(t), \quad h_0(t) = f'(t) - C(t)f(t), \\ h_n(t) &= n e^{-n(t-t_0)}(A(t_0) + C(t))R(n, A(t_0))u_1, \end{aligned} \tag{4.8}$$

we obtain

$$(B_n(v_n' - u_1))(t) + A(t)v_n'(t) + C(t)(nR(n, B)v_n')(t) + h_n(t) + h_0(t) = 0. \tag{4.9}$$

From (1.2) and (1.3) and for

$$c^\star = \sup_{t_0 \le t \le T} \|C(t)\|_{\mathcal{L}(E)} \tag{4.10}$$

we have

$$c^\star = c_0 \left(1 - \frac{M}{\omega}\right) \|A'\|_{C(t_0, T; \mathcal{L}(D, E))}. \tag{4.11}$$

Hence setting for $t \in [t_0, T]$

$$\widehat{A}(t) = A(t) + C(t),$$

we conclude (see Proposition 3.5 of [8]) that $\widehat{A}(t) : D \subset E \to E$ verifies Hypotheses (1.1)–(1.3) and (1.3) with c_0 and ω substituted by $c_0(1 + c^\star)$ and $\omega + Mc^\star$. From Hypothesis (4.1) it can be seen that $\widehat{A}$ also satisfies Hypothesis 3.4 with A_k substituted by $A_k + A_k' A_k^{-1}$. As $h \in L^p(t_0, T; E)$ and $u_1 \in \overline{D}$ we deduce from Theorem 3.5 the existence of z, the F-solution in L^p of problem

$$\begin{cases} z'(t) = [A(t) + C(t)]z(t) + h_0(t) \\ z(t_0) = u_1. \end{cases} \tag{4.12}$$

Hence if z_n are the solutions of its approximating problem

$$B_n(z_n - u_1)(t) + A(t)z_n(t) + C(t)z_n(t) + h_0(t) = 0, \tag{4.13}$$

we have (by virtue of Theorem 2.4)

$$\lim_{n\to\infty} \|z - z_n\|_{L^p(t_0, T; E)} = 0. \tag{4.14}$$

Setting

$$(\Gamma_n v)(t) = C(t)(nR(n, B)v)(t) \tag{4.15}$$

$$\begin{aligned} g_n(t) \quad = \quad & C(t)[nR(n, B)(z_n - z)](t) \\ & + C(t)[(nR(n, B)z)(t) - z_n(t)] + h_n(t), \end{aligned} \tag{4.16}$$

we obtain from (4.9) and (4.13)

$$B_n(v_n' - z_n) + \Lambda(v_n' - z_n) + \Gamma_n(v_n' - z_n) + g_n = 0. \tag{4.17}$$

Now from (4.10) we deduce that for each $v \in L^1(t_0, T; E)$ and $t \in [t_0, T]$ we have

$$\|\Gamma_n v\|_{L^1(t_0,t;E)} \leq c^\star \|v\|_{L^1(t_0,t;E)},$$

and so by applying Lemma 3.1 to (4.17) we get for large n

$$\|v_n' - z_n\|_{L^1(t_0,T;E)} \leq c\|g_n\|_{L^1(t_0,T;E)}, \tag{4.18}$$

where c is independent on n. To estimate $\|g_n\|_{L^1(t_0,T;E)}$ let us observe that

$$\|C(\cdot)nR(n,B)(z_n - z)\|_{L^1(t_0,T;E)} \leq c^\star \|z_n - z\|_{L^1(t_0,T;E)}$$

$$\|C(\cdot)[nR(n,B)z - z_n]\|_{L^1(t_0,T;E)}$$

$$\leq c^\star \left\{\|nR(n,B)z - z\|_{L^1(t_0,T;E)} + \|z_n - z\|_{L^1(t_0,T;E)}\right\},$$

$$\|h_n\|_{L^1(t_0,T;E)} =$$

$$\int_{t_0}^T ne^{-n(t-t_0)}\| - u_1 + nR(n, A(t_0))u_1 + C(t)R(n, A(t_0)u_1)\|dt,$$

and so from (4.14) we deduce that $\lim_{n\to\infty} \|g_n\|_{L^1(t_0,T;E)} = 0$. Hence by virtue of (4.18) and (4.14) we obtain

$$\lim_{n\to\infty} \|v_n' - z\|_{L^1(t_0,T;E)} = 0. \tag{4.19}$$

This (together with (4.3)) implies that $u \in W^{1,1}(t_0, T; E)$ and $u' = z$ a.e. But $z \in C(t_0, T; E)$ is an F-solution in L^p (see Theorem 2.5) and so $u \in C^1(t_0, T; E)$; as u is an F-solution of (4.2) we have $u(t_0) = u_0$.

Now let us show that u satisfies the equation in (4.2). Since $u \in D(B)$ we have

$$B_n v_n - Bu = nR(n,B)(u' - v_n') + B_n R(n, A(t_0))u_1 + (B_n - B)u$$

and so

$$\lim_{n\to\infty} \|B_n v_n - Bu\|_{L^1(t_0,T;E)} = 0.$$

By virtue of (4.4) this implies

$$\begin{aligned} &\lim_{n\to\infty} \|\Lambda v_n + Bu + f\|_{L^1(t_0,T;E)} \\ &= \lim_{n\to\infty} \|Bu - B_n(v_n - u_0)\|_{L^1(t_0,T;E)} = 0. \end{aligned} \tag{4.20}$$

Since $\Lambda : D(\Lambda) \subset L^1(t_0, T; E) \to L^1(t_0, T; E)$ is a closed operator and $v_n \in L^1(t_0, T; D) = D(\Lambda)$, (4.3) and (4.20) imply that $u \in L^1(t_0, T; D)$ and $\Lambda u = -Bu - f$. Since $-Bu = u'$ and f are continuous we deduce $A(t)u(t) = u'(t) - f(t)$ for each $t \in [t_0, T]$. The last part of the theorem is a consequence of the fact that $u' = z$ is a F-solution in L^p of (4.12). ■

5. Existence of F-solutions in C

The regularity theorem of §4 enables us to prove the existence of an F-solution in C.

Theorem 5.1. *Let A satisfy the assumptions of Theorem 4.2. Then for each $f \in C(t_0, T; E)$ and $u_0 \in \overline{D}$ there exists a unique F- solution in C of problem* (1.2).

Proof. Setting $\omega_1 = \omega + 1$ and for $k > \omega$

$$u_{0k} = kR(k, A(t_0))[u_0 - R(\omega_1, A(t_0))f(t_0)] + R(\omega_1, A(t_0))f(t_0)$$

we have $u_{0k} \in \overline{D}$ and (since $u_0 \in \overline{D}$)

$$\lim_{k \to \infty} u_{0k} = u_0. \tag{5.1}$$

Moreover

$$\begin{aligned} A(t_0)u_{0k} + f(t_0) &= (k^2 R(k, A(t_0)) - k)[u_0 - R(\omega_1, A(t_0))f(t_0)] \\ &\quad + \omega_1 R(\omega_1, A(t_0))f(t_0), \end{aligned} \tag{5.2}$$

hence

$$A(t_0)u_{0k} + f(t_0) \in \overline{D}.$$

Let $f_k \in W^{1,p}(t_0, T; E)$ be such that $f_k(t_0) = f(t_0)$ and

$$\lim_{k \to \infty} \|f - f_k\|_{C(t_0, T; E)} = 0. \tag{5.3}$$

From Theorem 4.2 we deduce the existence of $u_k \in C^1(t_0, T; E) \cap C(t_0, T; D)$ as solution of

$$\begin{cases} u_k'(t) = A(t)u_k(t) + f_k(t), \ t \in [t_0, T] \\ u_k(t_0) = u_{0k}. \end{cases} \tag{5.4}$$

From the estimate (2.27) we deduce by virtue of (5.1) and (5.3) that there exists $u \in C(t_0, T; E)$ such that

$$\lim_{k \to \infty} \|u - u_k\|_{C(t_0, T; E)} = 0,$$

and so u is an F-solution in C of problem (1.2). ■

6. An application

In this section we want to show an application of the preceding theory to the following first order partial differential problem:

$$u_t(t, x) + a(t, x)u_x(t, x) = f(t, x), \ t \in I, x \in J \ \text{ a.e.} \tag{6.1}$$

$$u(t_0, x) = u_0(x), \qquad x \in J \ \text{ a.e.} \tag{6.2}$$

$$u(t, 0) = u(t, l), \qquad t \in I, \tag{6.3}$$

where $I = [t_0, T]$, $J = [0, l]$. We will suppose that $x \to a(t, x)$ is discontinuous so that the method of characteristics cannot be used. We will consider only the strict solutions of our problem, more precisely we will prove the following result.

Theorem 6.1. *Let $a \in C^1(I; L^\infty(J))$ and $\gamma > 0$ be such that for each $t \in I$ we have*

$$\gamma^{-1} < a(t,x) < \gamma, \quad x \in J \ a.e. \tag{6.4}$$

and let $f \in W^{1,1}(I; L^\infty(J))$ and

$$u_0 \in \ Lip_\#(J) = \{u : [0,l] \to \mathbb{R} \mid \sup_{0 \le x < y \le l} \frac{|u(x) - u(y)|}{|x-y|} < \infty; u(0) = u(l)\}$$

verify the compatibility condition

$$a(t_0,\cdot)u_0'(\cdot) + f(t_0,\cdot) \in C_\#(J). \tag{6.5}$$

Then there is a unique solution u of problem (6.1)-(6.3) *such that $u, u_t \in C(I \times J)$ and $u_x(t,x)$ exists for $t \in I$ and $x \in J$ a.e.. In addition we have for each $t \in I$*

$$\sup_{x\in J} |u(t,\cdot)| \le \sup_{x\in J} |u_0(x)| + \int_{t_0}^t ess \ sup_{x\in J} |f(s,x)| ds. \tag{6.6}$$

Proof. We will denote by E the Banach space $L^\infty(J)$ with norm $\|u\|_\infty =$ ess $\sup_{x\in J}|u(x)|$ and by D the Banach space $\mathrm{Lip}_\#(J)$ with norm $\|u\|_\infty + \|u'\|_\infty$. Let $a : J \to \mathbb{R}_+^*$ be such that $a, a^{-1} \in L^\infty(J)$ and define $A : D \subset E \to E$ by setting $(Au)(x) = -a(x)u'(x), u \in D, x \in J$ a.e.. Given $\lambda > 0$ and $f \in E$ the only solution $u \in D$ of $\lambda u - Au = f$ is

$$u(x) = ce^{-\lambda b(x)} + \int_0^x e^{-\lambda[b(x)-b(y)]} a^{-1}(y) f(y) dy, \quad x \in J,$$

where $b(x) = \int_0^x a^{-1}(y)dy$ and

$$c = (e^{\lambda b(l)} - 1)^{-1} \int_0^l e^{\lambda b(y)} a^{-1}(y) f(y) dy.$$

As $e^{\lambda b(y)} a^{-1}(y) = \frac{1}{\lambda}\frac{d}{dy} e^{\lambda b(y)} > 0$, $y \in J$ a.e. we deduce for each $x \in J$

$$\begin{aligned} |u(x)|e^{\lambda b(x)} \ &\le \ (e^{\lambda b(l)} - 1)^{-1} \textstyle\int_0^l \frac{1}{\lambda}\frac{d}{dy} e^{\lambda b(y)} |f(y)| dy \\ &\quad + \textstyle\int_0^x \frac{1}{\lambda}\frac{d}{dy} e^{\lambda b(y)} |f(y)| dy \le e^{\lambda b(x)} \frac{\|f\|_\infty}{\lambda}, \end{aligned}$$

and so $\rho(A) \supset]0, +\infty[$ and $\|\lambda R(\lambda, A)\| \le 1$, for $\lambda > 0$.
Now let us define for each $t \in I$, $A(t) : D \to E$ as

$$(A(t)u)(x) = -a(t,x)u'(x).$$

From (6.4) we deduce the estimate in Hypothesis (1.2) with $c_0 = \max\{1, \gamma\}$ and from the preceding result we obtain that Hypothesis (1.3) holds with $\omega = 0$ and $M = 1$. We can now prove that A, f and u_0 satisfy the assumptions of Theorem 4.1. Since $a \in C^1(I; E)$ we get $A \in C^1(I; \mathcal{L}(D, E))$ and the existence of $\{a_k\} \subset C^\infty(I; E)$ such that $\lim_{k\to\infty} \|a - a_k\|_{C^1(I;E)} = 0$. Hence setting $a_k(t)(x) = a_k(t,x)$ there exists $\gamma_1 > 0$ such that for each $t \in I$ and $k \in \mathbb{N}$ we have $\gamma_1^{-1} < a_k(t,x) < \gamma_1$ for $x \in J$, a.e. Setting for $t \in I$ and $k \in \mathbb{N}$, $A_k(t)u = -a_k(t)u'$, Hypothesis 4.1 is satisfied with the corresponding constants $c_0 = \max\{1, \gamma_1\}$, $\omega = 0$ and $M = 1$. As $\overline{D} = C_\#(J)$, the conclusion follows from Theorem 4.2 and estimate (2.27). ∎

In the preceding theorem the stability condition in Hypothesis 1.3 has been easily verified because $M = 1$. This is not the case in the following example which can be solved when a is more regular, i.e.,

$$a(t,x) > 0 \text{ for } (t,x) \in Q = I \times J \tag{6.7}$$

$$a \in C^1(I;\ \mathrm{Lip}_{\#}(J)) \tag{6.8}$$

($\mathrm{Lip}_{\#}(J)$ is endowed with the norm $\|u\|_\infty + \|u'\|_\infty$). In this case there exist $\gamma, N, M > 0$ such that for each $t, s \in I$ and $x \in J$

$$\begin{array}{l} \gamma^{-1} < a(t,x) < \gamma \\ \sup_{t\in I}\left\{\|a(t,\cdot)\|_{\mathrm{Lip}_{\#}(J)}, \|a^{-1}(t,\cdot)\|_{\mathrm{Lip}_{\#}(J)}\right\} = L < \infty \\ |a(t,x) - a(s,x)| \le L|t-s|. \end{array} \tag{6.9}$$

Let us replace E and D with

$$E_1 = \ \mathrm{Lip}_{\#}(J),\ \ D_1 = \{u \in E_1 \mid u' \in E_1\}, \tag{6.10}$$

and define for each $t \in I$, $A(t) : D_1 \to E_1$ as in Theorem 6.1. Then Hypothesis 1.2 holds with $c_0 = \max\{1, N\}$ and $\rho(A(t)) \supset]0, +\infty[$.

To prove Hypothesis 1.3 let us define $B : D \to E$ as

$$Bu = u - u'.$$

For $t_0 \le s \le t \le T$, we deduce from the first part of the proof of Theorem 6.1 the following estimates

$$\begin{array}{rcl} \|B^{-1}\|_{\mathcal{L}(E)} & \le & 1 \\ \|(I - A(t))B^{-1}\|_{\mathcal{L}(E)}, \|B(I - A(t))^{-1}\|_{\mathcal{L}(E)} & \le & 2\gamma + 1 \\ \|(I - A(t))(I - A(s))^{-1}\|_{\mathcal{L}(E)} & \le & e^{c(t-s)}, \end{array} \tag{6.11}$$

where $c = 3L$.

Now for $u \in E_1, \lambda > 0$ and $t_0 \le t_n \le \cdots \le t_1 \le T$ we have

$$\begin{array}{rcl} \|R(\lambda; t_1, \cdots, t_n)u\|_{E_1} & = & \|R(\lambda, t_1, \cdots, t_n)u\|_E + \|(I - B)R(\lambda, t_1, \cdots, t_n)u\|_E \\ & \le & \dfrac{2}{\lambda^n}\|u\|_E + \|BR(\lambda, t_1, \cdots, t_n)u\|_E. \end{array}$$

As by virtue of (6.11) we obtain

$$\begin{array}{rcl} l\|BR(\lambda, t_1, \cdots, t_n)u\|_E & \le & \|B(I - A(t_1))^{-1}\|_{\mathcal{L}(E)}\|R(\lambda, A(t_1))\|_{\mathcal{L}(E)} \cdot \\ & & \|(I - A(t_1))(I - A(t_2))^{-1}\|_{\mathcal{L}(E)}\|R(\lambda, A(t_2))\|_{\mathcal{L}(E)} \cdot \\ & & \|(I - A(t_{n-1}))(I - A(t_n))^{-1}\|_{\mathcal{L}(E)}\|R(\lambda, A(t_n))\|_{\mathcal{L}(E)} \cdot \\ & & \|(I - A(t_n))B^{-1}\|_{\mathcal{L}(E)}\|u - u'\|_E \\ & \le & (2\gamma + 1)^2 e^{c(t_1 - t_n)} \dfrac{1}{\lambda^n}\|u\|_{E_1}. \end{array}$$

Estimate in Hypothesis 1.3 is verified with $M = 2 + (2\gamma + 1)^2 e^{c(T - t_0)}$.

In this case we could also prove that Hypothesis 4.1 holds and so we could apply Theorem 4.2. For the sake of brevity we omit the explicite application to problem (6.1)-(6.3).

References

[1] Arendt W., *Resolvent positive operators and integrated semigroups*, Proc. London Math. Soc. **54** (1987), 321-349.

[2] Daleckii, J. L. and M. G. Krein, "Stability of solutions of differential equations in Banach space," Am. Math. Soc., 1974.

[3] Da Prato, G. and E. Sinestrari, *On the Phillips and Tanabe regularity theorems*, Semesterbericht Funktionalanalysis, Tübingen, Sommersemester 1985, 117-124.

[4] Da Prato, G. and E. Sinestrari, *Differential operators with non dense domain*, Ann. Sc. Norm. Sup. Pisa **14** (1987), 285-344.

[5] Kato, T., "Perturbation theory of linear operators", Springer-Verlag, New-York, 1966.

[6] Kato, T., *Linear evolution equations of "hyperbolic type" II*, J. Math. Soc. Japan **25** (1973), 648-666.

[7] Kellermann, H. and M. Hieber, *Integrated semigroups*, J.Funct. Analysis **84** (1989), 160-180.

[8] Pazy, A., "Semigroups of Linear Operators and Applications to Partial Differential Equations," Springer-Verlag, New-York, 1983.

[9] Thieme, H. R., *Semiflows generated by Lipschitz perturbations of non-densely defined operators*, Diff. Integral Equations, **3** (1990), 1035-1066.

Scuola Normale di Pisa
56126 Pisa, Italy

Dipartimento di Matematica
Universitá di Roma I
Piazzale Aldo Moro 2
00185 Roma, Italy

Received October 6, 1990
and in final form December 12, 1991

Semigroup Forum Vol. 45 (1992) 322–331

RESEARCH ARTICLE

Finite images and elementary equivalence of completely regular inverse semigroups

Francis Oger

Communicated by K. Keimel

The definitions and results of model theory which are used here, in particular the notions of language, elementary equivalence and elementary extension, are given in [1]. The reader is referred to [9] concerning group theory, and to [3] concerning semigroup theory. We draw the reader's attention to the fact that our definition of a diagram of groups technically differs from some categorical definitions.

The finite images of a group G are the finite groups H such that there exists a surjective homomorphism from G to H. Two finitely generated abelian groups which have the same finite images are necessarily isomorphic. On the other hand, there are examples of finitely generated abelian-by-finite groups which have the same finite images without being isomorphic (several references are given in [4, Corollary 5.6]).

In [6], we show that two finitely generated abelian-by-finite groups G, H are elementarily equivalent if and only if they have the same finite images. It follows that two finitely generated abelian-by-finite groups can be elementarily equivalent without being isomorphic.

As far back as 1970, B. I. Zil'ber gave in [10] an example of two nonisomorphic finitely generated commutative semigroups which are elementarily equivalent. In the present paper, inspired by Zil'ber's example, we generalize the properties of finite images and the result of [6] by considering semigroups, or diagrams of groups, instead of groups.

In the first section, we prove that, if two diagrams of groups G, H are finite unions of finitely generated abelian-by-finite groups, then G and H are elementarily equivalent if and only if they have the same finite images. We also obtain an algebraic characterization of elementary equivalence for diagrams of groups which are infinite unions of finitely generated abelian-by-finite groups.

In the second section, using an interpretation of completely regular inverse semigroups in terms of diagrams of groups, we show that, if two completely regular inverse semigroups G, H are finite unions of finitely generated abelian-by-finite groups, then G and H are elementarily equivalent if and only if they have the same finite images. As the semigroups which are considered by Zil'ber in [10] are inverse and completely regular, we obtain Zil'ber's example as a consequence of our results. Moreover, we give some examples of nonisomorphic finitely generated commutative torsion-free completely regular inverse semigroups which are elementarily equivalent.

In the second section, we also prove that, for each integer $n \geq 1$ and for any subgroups S, T of $\mathbb{Z}^n$, if $(\mathbb{Z}^n, S)$ and $(\mathbb{Z}^n, T)$ are elementarily equivalent, then they are isomorphic. On the other hand, we give some examples of subgroups S_1, S_2, T_1, T_2 of $\mathbb{Z}^2$ such that $(\mathbb{Z}^2, S_1, S_2)$ and $(\mathbb{Z}^2, T_1, T_2)$ are elementarily equivalent, but not isomorphic.

In the third section, we consider diagrams of groups which consist of one group and one or several endomorphisms of this group; in particular, we consider modules over not necessarily commutative rings. For each integer $n \geq 1$, and for any endomorphisms g, h of $\mathbb{Z}^n$, with matrices A, B in the canonical basis of $\mathbb{Z}^n$, the diagrams of groups $(\mathbb{Z}^n, +, g)$ and $(\mathbb{Z}^n, +, h)$ are isomorphic if and only if A and B are conjugate as matrices with coefficients in $\mathbb{Z}$; we prove that $(\mathbb{Z}^n, +, g)$ and $(\mathbb{Z}^n, +, h)$ are elementarily equivalent if and only if, for each integer $k \geq 1$, the images of A and B modulo k are conjugate as matrices with coefficients in $\mathbb{Z}/k\mathbb{Z}$. In [7] and [8], V. N. Remeslennikov gave some examples of matrices A, B which are not conjugate as matrices with coefficients $\mathbb{Z}$, while their images modulo k are conjugate as matrices with coefficients in $\mathbb{Z}/k\mathbb{Z}$, for each integer $k \geq 1$.

1. Diagrams of groups

Definition and first properties of diagrams of groups.

We consider a *type of diagram* $D = (I, J, c, d)$ which is made of two sets I, J and two maps c, d from J to I. We call ***diagram of type D in the category of groups*** or, more briefly, *D-diagram of groups*, a system $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ which consists of:

- a family of pairwise disjoint groups $(G_i)_{i\in I}$, whose set theoretic union is the universe of the structure;
- a family of homomorphisms $g_j : G_{c(j)} \to G_{d(j)}$.

The language of a D-diagram of groups $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ is obtained by considering a binary relational symbol for each g_j and a ternary relational symbol for the law of each G_i.

A *homomorphism* from a D-diagram of groups $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ to a D-diagram of groups $H = ((H_i)_{i\in I}, (h_j)_{j\in J})$ is the union $f = \cup_{i\in I} f_i$ of a family of homomorphisms $f_i : G_i \to H_i$ which satisfy $h_j \circ f_{c(j)} = f_{d(j)} \circ g_j$ for each $j \in J$.

A *subdiagram* $H = ((H_i)_{i\in I}, (h_j)_{j\in J})$ of a D-diagram of groups $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ is obtained by taking subgroups H_i of the groups G_i such that $g_j(H_{c(j)}) \subset H_{d(j)}$ for each $j \in J$ and considering, for each $j \in J$, the restriction h_j of g_j to $H_{c(j)}$. We say that H is *normal* in G if the subgroups H_i are normal in the groups G_i. In that case, we denote by G/H the D-diagram of groups $((G_i/H_i)_{i\in I}, (k_j)_{j\in J})$ where, for each $j \in J$, k_j is the homomorphism from $G_{c(j)}/H_{c(j)}$ to $G_{d(j)}/H_{d(j)}$ which is induced by g_j.

Finite images and profinite completion.

In this subsection, we consider a type of diagram $D = (I, J, c, d)$ with I *finite*, and we generalize to D-diagrams of polycyclic-by-finite groups the notion of profinite completion, which is defined in [9, §10, B] for polycyclic-by-finite groups.

The *finite images* of a D-diagram of groups G are the finite D-diagrams H such that there exists a surjective homomorphism from G to H.

For each group M and each integer $n \geq 1$, we denote by M^n the subgroup of M which is generated by $\{x^n \mid x \in M\}$. Similarly, for each D-diagram of groups $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ and each integer $n \geq 1$, we denote by G^n the union of the subgroups G_i^n, which is a normal subdiagram of G.

In [9, §10, B], D. Segal observes that, if M is a polycyclic-by-finite group, then the quotients M/M^n for $n \geq 1$ are finite, and the following projective systems are equivalent:

a) the set of all finite quotients M/N;

b) $\{M/M^n \mid n \geq 1\}$;

c) $\{M/M^{n!} \mid n \geq 1\}$.

He defines the profinite completion $\widehat{M}$ of M as the projective limit of any of these systems.

Similarly, if $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ is a D-diagram of polycyclic-by-finite groups, then, for any normal subdiagrams $H \subset K$ of G, the canonical surjection from G to G/H induces a surjective homomorphism from G/K to G/H. Moreover, the quotients G/G^n for $n \geq 1$ are finite since I is finite, and any normal subdiagram H of G such that G/H is finite contains G^n for some integer $n \geq 1$. So, the following projective systems are equivalent:

a) the set of all finite quotients G/H;

b) $\{G/G^n \mid n \geq 1\}$;

c) $\{G/G^{n!} \mid n \geq 1\}$.

We define the *profinite completion* $\widehat{G}$ of G as the projective limit of any of these systems. We have $\widehat{G} = \left((\widehat{G}_i)_{i\in I}, (\hat{g}_j)_{j\in J}\right)$, where, for each $i \in I$, $\widehat{G}_i$ is the profinite completion of G_i and, for each $j \in J$, $\hat{g}_j$ is the unique continuous homomorphism from $\widehat{G}_{c(j)}$ to $\widehat{G}_{d(j)}$ which extends g_j (see [9, §10, Prop. 3, p. 222] for the definition of $\hat{g}_j$).

By [9, §10, Prop. 4, p. 225], the following properties are equivalent for two polycyclic-by-finite groups M, N:

1) M and N have the same finite images;

2) $M/M^n \cong N/N^n$ for each integer $n \geq 1$;

3) $\widehat{M} \cong \widehat{N}$.

The key argument in the proof is that, for each integer $n \geq 1$, the set of all isomorphisms from M/M^n to N/N^n is finite. We can show in a similar way that, if G and H are D-diagrams of polycyclic-by-finite groups, then the following properties are equivalent:

1) G and H have the same finite images;

2) $G/G^n \cong H/H^n$ for each integer $n \geq 1$;

3) $\widehat{G} \cong \widehat{H}$.

Finite images and elementary equivalence.

The following result generalizes [6, Theorem 1]:

Theorem 1.1. *Let $D = (I, J, c, d)$ be a type of diagram with I finite. Then, any D-diagram of finitely generated abelian-by-finite groups is an elementary submodel of its profinite completion.*

Notations. For each group M, for each ultrafilter U over a set I, for each $x \in M$ and for each $a \in \mathbb{Z}^U$, we denote by x^a the element of M^U which is represented by $\left(x^{a(i)}\right)_{i\in I}$, where $(a(i))_{i\in I}$ is a representative of a in $\mathbb{Z}^I$. The element x^a does not depend on the choice of the representative $(a(i))_{i\in I}$.

For each polycyclic-by-finite group M, for each $x \in M$ and for each $a \in \widehat{\mathbb{Z}}$, if $(a(n))_{n\in\mathbb{N}}$ is a sequence of elements of $\mathbb{Z}$ which converges to a in

$\widehat{\mathbb{Z}}$, then, $(x^{a(n)})_{n\in\mathbb{N}}$ converges in $\widehat{G}$ to an element that we denote by x^a. This element does not depend on the choice of sequence $(a(n))_{n\in\mathbb{N}}$.

Proof of Theorem 1.1. Let $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ be a D-diagram of finitely generated abelian-by-finite groups and let $\widehat{G} = \left((\widehat{G}_i)_{i\in I}, (\hat{g}_j)_{j\in J}\right)$ be its profinite completion. We want to show that G is an elementary submodel of $\widehat{G}$, which means that G and $\widehat{G}$ satisfy the same sentences with parameters in G. By [1, Theorem 4.1.9], it suffices to prove that there exists an ultrafilter U and an isomorphism $f : G^U \to \widehat{G}^U$ which fixes the elements of G.

As I is finite, [6, Lemma 1.1] implies the existence of an integer $n \geq 1$ such that the groups G_i^n are abelian and torsion-free. According to the proof of [6, Theorem 1], there exist an ultrafilter U, an isomorphism of groups $h : \mathbb{Z}^U \to \widehat{\mathbb{Z}}^U$ and some isomorphisms $f_i : G_i^U \to \widehat{G}_i^U$ which satisfy $f_i(x) = x$ for each $x \in G_i$ and $f_i(x^a) = x^{h(a)}$ for each $x \in G_i^n$ and each $a \in \mathbb{Z}^U$.

The map $f = \cup_{i\in I} f_i$ is a bijection from $G^U = ((G_i^U)_{i\in I}, (g_j^U)_{j\in J})$ to $\widehat{G}^U = \left((\widehat{G}_i^U)_{i\in I}, (\hat{g}_j^U)_{j\in J}\right)$ which fixes the elements of G. In order to show that f is an isomorphism, it remains to be proved that the isomorphisms f_i satisfy $\hat{g}_j^U \circ f_{c(j)} = f_{d(j)} \circ g_j^U$ for each $j \in J$.

We have $\left(\hat{g}_j^U \circ f_{c(j)}\right)(x) = \hat{g}_j^U(x) = g_j(x)$ and $\left(f_{d(j)} \circ g_j^U\right)(x) = f_{d(j)}(g_j(x)) = g_j(x)$ for each $x \in G_{c(j)}$. We also have $\left(\hat{g}_j^U \circ f_{c(j)}\right)(y^a) = \hat{g}_j^U(y^{h(a)}) = g_j(y)^{h(a)}$ and $\left(f_{d(j)} \circ g_j^U\right)(y^a) = f_{d(j)}(g_j(y)^a) = g_j(y)^{h(a)}$ for each $y \in G_{c(j)}^n$ and each $a \in \mathbb{Z}^U$; the last equality is true because $g_j(y)$ belongs to $G_{d(j)}^n$. So, we have $\left(\hat{g}_j^U \circ f_{c(j)}\right)(x) = \left(f_{d(j)} \circ g_j^U\right)(x)$ for $x \in G_{c(j)}$ and for $x = y^a$ with $y \in G_{c(j)}^n$ and $a \in \mathbb{Z}^U$. This implies $\hat{g}_j^U \circ f_{c(j)} = f_{d(j)} \circ g_j^U$ since $G_{c(j)}^U$ is generated by $G_{c(j)}$ and the elements y^a for $y \in G_{c(j)}^n$ and $a \in \mathbb{Z}^U$. ■

The following result generalizes [6, Corollary]. It is a consequence of Theorem 1.1:

Corollary 1.2. *Let $D = (I, J, c, d)$ be a type of diagram, with I finite, and let G, H be D-diagrams of finitely generated abelian-by-finite groups. Then, the following properties are equivalent:*

a) *G and H are elementarily equivalent;*
b) *G and H have the same finite images;*
c) *$G/G^n \cong H/H^n$ for each integer $n \geq 1$;*
d) *$\widehat{G} \cong \widehat{H}$.*

The corollary below gives a criterion of elementary equivalence which is valid for any type of diagram $D = (I, J, c, d)$. We say that a type of diagram $D^* = (I^*, J^*, c^*, d^*)$ is a *restriction* of D if we have $I^* \subset I$, $J^* \subset J$, $c(J^*) \subset I^*$, $d(J^*) \subset I^*$, $c^* = c \restriction J^*$ and $d^* = d \restriction J^*$. The *restriction* a D-diagram of groups $G = ((G_i)_{i\in I}, (g_j)_{j\in J})$ to D^* is the D^*-diagram of groups $G^* = ((G_i)_{i\in I^*}, (g_j)_{j\in J^*})$.

Corollary 1.3. *Let $D = (I, J, c, d)$ be a type of diagram and let G, H be two D-diagrams of finitely generated abelian-by-finite groups. Then, G and H are elementarily equivalent if and only if. for each restriction $D^* = (I^*, J^*, c^*, d^*)$ of D associated to finite subsets I^*, J^* of I, J, the restrictions of G and H to D^* have the same finite images.*

Proof. Each first-order sentence is built up from a finite number of symbols. So, G and H are elementarily equivalent if and only if, for each restriction $D^* = (I^*, J^*, c^*, d^*)$ of D associated with finite subsets I^*, J^* of I, J, the restrictions of G and H to D^* are elementarily equivalent. Now, Corollary 1.3 follows from Corollary 1.2 applied to the "finite" restrictions of G and H. ■

2. Completely regular inverse semigroups

The definitions and results of semigroup theory which are used here are given in [3]. A *semigroup* is a set equipped with an associative law. We interpret this law with a binary functional symbol. A semigroup G is *inverse* if, for each $x \in G$, there exists a unique $x' \in G$ such that $xx'x = x$ and $x'xx' = x'$. An inverse semigroup G is *completely regular* if we have $xx' = x'x$ for each $x \in G$.

If we consider:
- a partially ordered set $(I, \leq)$ such that $\inf(i, j)$ exists for any $i, j \in I$,
- a family $(G_i)_{i \in I}$ of pairwise disjoint groups,
- a family of homomorphisms of groups $g_{j,i} : G_j \to G_i$ defined for $i \leq j$ and such that $g_{i,i} = id_{G_i}$ for each $i \in I$ and $g_{k,j} \circ g_{j,i} = g_{k,i}$ for $i \leq j \leq k$,

then we define a structure of completely regular inverse semigroup on $\cup_{i \in I} G_i$ by writing $xy = g_{i,k}(x) g_{j,k}(y)$ for any $i, j, k \in I$ such that $k = \inf(i, j)$, for each $x \in G_i$ and for each $y \in G_j$.

Conversely, if G is a completely regular inverse semigroup, then there exist:
- a partially ordered set $(I, \leq)$ such that $\inf(i, j)$ exists for any $i, j \in I$;
- a partition $(G_i)_{i \in I}$ of G in subsemigroups such that, for each $i \in I$, G_i equipped with the law of G is a group;
- a family of homomorphisms of groups $g_{j,i} : G_j \to G_i$ defined for $i \leq j$ and such that $g_{i,i} = id_{G_i}$ for each $i \in I$ and $g_{k,j} \circ g_{j,i} = g_{k,i}$ for $i \leq j \leq k$;

all three of them defined in such a way that $xy = g_{i,k}(x) g_{j,k}(y)$ for any $i, j, k \in I$ such that $k = \inf(i, j)$, for each $x \in G_i$ and for each $y \in G_j$.

The semigroup G is completely determined by the diagram of groups $G^* = ((G_i)_{i \in I}, (g_{j,i})_{i<j})$, where $i < j$ means $i \leq j$ and $i \neq j$. The diagram of groups G^* is itself determined by G since the groups G_i are the maximal groups contained in G. We say that G^* is *associated* with G.

If G and H are completely regular inverse semigroups, and if the diagrams of groups $((G_i)_{i \in K}, (g_{j,i})_{i<j})$ and $((H_i)_{i \in L}, (h_{j,i})_{i<j})$ are associated with G and H respectively, then G and H are isomorphic if and only if there exists an isomorphism σ from $(K, \leq)$ to $(L, \leq)$ such that the diagrams of groups $((G_i)_{i \in K}, (g_{j,i})_{i<j})$ and $((H_{\sigma(i)})_{i \in K}, (h_{\sigma(j),\sigma(i)})_{i<j})$ are isomorphic.

The following proposition shows the connection between the elementary equivalence of two completely regular inverse semigroups which are unions of finitely many groups, and the elementary equivalence of the associated diagrams of groups:

Proposition 2.1. *Let G and H be completely regular inverse semigroups. Let $((G_i)_{i \in K}, (g_{j,i})_{i<j})$ and $((H_i)_{i \in L}, (h_{j,i})_{i<j})$ be diagrams of groups associated with G and H. Let us suppose that K and L are finite. Then, G and H are elementarily equivalent if and only if there exists an isomorphism σ from $(K, \leq)$ to $(L, \leq)$ such that the diagrams of groups $((G_i)_{i \in K}, (g_{j,i})_{i<j})$ and $((H_{\sigma(i)})_{i \in K}, (h_{\sigma(j),\sigma(i)})_{i<j})$ are elementarily equivalent.*

Proof. By [1, Theorem 6.1.15], two structures M, N are elementarily equivalent if and only if there exists an ultrafilter U such that M^U and N^U are isomorphic. For each ultrafilter U, G^U and H^U are respectively associated with $((G_i)_{i\in K},(g_{j,i})_{i<j})^U = ((G_i^U)_{i\in K},(g_{j,i}^U)_{i<j})$ and $((H_i)_{i\in L},(h_{j,i})_{i<j})^U = ((H_i^U)_{i\in L},(h_{j,i}^U)_{i<j})$. So, the following properties are equivalent:

1) G and H are elementarily equivalent;
2) There exists an ultrafilter U such that G^U and H^U are isomorphic;
3) There exist an ultrafilter U and an isomorphism σ from $(K,\leq)$ to $(L,\leq)$ such that $((G_i^U)_{i\in K},(g_{j,i}^U)_{i<j})$ and $((H_{\sigma(i)}^U)_{i\in K},(h_{\sigma(j),\sigma(i)}^U)_{i<j})$ are isomorphic;
4) There exists an isomorphism σ from $(K,\leq)$ to $(L,\leq)$ such that $((G_i)_{i\in K},(g_{j,i})_{i<j})$ and $((H_{\sigma(i)})_{i\in K},(h_{\sigma(j),\sigma(i)})_{i<j})$ are elementarily equivalent. ■

Now, we only consider *completely regular inverse semigroups which are disjoint finite unions of polycyclic-by-finite groups*. The *finite images* of a semigroup G are the semigroups H such that there exists a surjective homomorphism from G to H. The semigroups G/G^n and the *profinite completion* $\widehat{G}$ are defined from the diagrams of groups M/M^n and the profinite completion $\widehat{M}$, where M is a diagram of groups associated with G. Any finite image of G is an image of G/G^n for some integer $n\geq 1$. So, the following properties are equivalent for two semigroups G, H:

1) G and H have the same finite images;
2) $G/G^n \cong H/H^n$ for each $n\geq 1$;
3) $\widehat{G}\cong\widehat{H}$.

The two following results are easy consequences of Theorem 1.1 and Corollary 1.2:

Theorem 2.2. *If a completely regular inverse semigroup is a disjoint finite union of finitely generated abelian-by-finite groups, then it is an elementary submodel of its profinite completion.* ■

Corollary 2.3. *Let G, H be completely regular inverse semigroups. Let us suppose that G and H are disjoint finite unions of finitely generated abelian-by-finite groups. Then, the following properties are equivalent:*

1) *G and H are elementarily equivalent;*
2) *G and H have the same finite images;*
3) *$G/G^n \cong H/H^n$ for each integer $n\geq 1$;*
4) *$\widehat{G}\cong\widehat{H}$.* ■

The hypotheses of Theorem 2.2 and Corollary 2.3 are satisfied by finitely generated commutative completely regular inverse semigroups. Corollary 2.4 and Corollary 2.7 provide some examples of such semigroups which satisfy 1), 2), 3), 4) without being isomorphic.

Corollary 2.4. *Let p be a prime number, let a_1, a_2, b_1, b_2 be nonzero elements of $\mathbb{Z}/p\mathbb{Z}$, and let G_0, G_1, G_2, H_0, H_1, H_2 be pairwise disjoint cyclic groups written with additive notation, respectively generated by u_0, u_1, u_2, v_0, v_1, v_2 and respectively isomorphic to $\mathbb{Z}/p\mathbb{Z}$, $\mathbb{Z}$, $\mathbb{Z}$, $\mathbb{Z}/p\mathbb{Z}$, $\mathbb{Z}$, $\mathbb{Z}$. For each $i\in\{1,2\}$, let us consider the homomorphisms $g_i : G_i \to G_0$ and $h_i : H_i \to H_0$ which satisfy $g_i(u_i) = a_iu_0$ and $h_i(v_i) = b_iv_0$. Then, the semigroups G, H associated with the diagrams (G_0,G_1,G_2,g_1,g_2) and (H_0,H_1,H_2,h_1,h_2) are elementarily equivalent. They are isomorphic if and only if $a_1b_2 = \pm a_2b_1$ or $a_1b_1 = \pm a_2b_2$.*

Proof. For each integer $n \geq 1$, there exist two integers r_1, r_2 which are prime to n and such that $a_1 = r_1 b_1$ and $a_2 = r_2 b_2$; the homomorphism $f_n : G \to H$ defined by $f_n(u_0) = v_0$, $f_n(u_1) = r_1 v_1$ and $f_n(u_2) = r_2 v_2$ induces an isomorphism from G/nG to H/nH.

The semigroups G, H are isomorphic if and only if $(G_0, G_1, G_2, g_1, g_2)$ is isomorphic to $(H_0, H_1, H_2, h_1, h_2)$ or to $(H_0, H_2, H_1, h_2, h_1)$. We are going to prove that the first property is true if and only if $a_1 b_2 = \pm a_2 b_1$. We can show in a similar way that the second one is true if and only if $a_1 b_1 = \pm a_2 b_2$.

Any isomorphism f from $(G_0, G_1, G_2, g_1, g_2)$ to $(H_0, H_1, H_2, h_1, h_2)$ is obtained by writing $f(u_0) = dv_0$ with d a non zero element of $\mathbb{Z}/p\mathbb{Z}$, $f(u_1) = \varepsilon_1 v_1$ with $\varepsilon_1 = \pm 1$ and $f(u_2) = \varepsilon_2 v_2$ with $\varepsilon_2 = \pm 1$. If such an isomorphism exists, then we have $(h_1 \circ f)(u_1) = h_1(\varepsilon_1 v_1) = \varepsilon_1 b_1 v_0 = (f \circ g_1)(u_1) = f(a_1 u_0) = a_1 d v_0$ and $(h_2 \circ f)(u_2) = h_2(\varepsilon_2 v_2) = \varepsilon_2 b_2 v_0 = (f \circ g_2)(u_2) = f(a_2 u_0) = a_2 d v_0$, which implies $\varepsilon_1 b_1 = a_1 d$ and $\varepsilon_2 b_2 = a_2 d$, whence $a_1 a_2 d = \varepsilon_1 a_2 b_1 = \varepsilon_2 a_1 b_2$ and $a_2 b_1 = \pm a_1 b_2$.

Conversely, if $a_1 b_2 = \varepsilon a_2 b_1$ with $\varepsilon = \pm 1$, then we define an isomorphism f from $(G_0, G_1, G_2, g_1, g_2)$ to $(H_0, H_1, H_2, h_1, h_2)$ by writing $f(u_0) = b_1 a_1^{-1} v_0$, $f(u_1) = v_1$ and $f(u_2) = \varepsilon v_2$. ∎

Example 2.5. It follows from Corollary 2.4 that G and H are elementarily equivalent and non-isomorphic if $p = 5$, $a_1 = 1$, $a_2 = 2$, $b_1 = 1$ and $b_2 = 4$. This example was obtained by Zil'ber in [10].

Now, we consider structures which consist of a finitely generated torsion-free abelian group and two subgroups. According to Proposition 2.6 below, two such structures can be elementarily equivalent and non-isomorphic. By Corollary 2.7, there exist non-isomorphic finitely generated commutative torsion-free completely regular inverse semigroups which are elementarily equivalent. For each group M and each $x \in M$, we denote by $\langle x \rangle$ the subgroup of M which is generated by x.

Proposition 2.6. *Let $r \geq 2$ be an integer and let a, b be integers which are prime to r. For each integer p, let us write $M_p = (\mathbb{Z} \times \mathbb{Z}, \langle (1,0) \rangle, \langle (p,r) \rangle)$. Then, M_a and M_b are elementarily equivalent; they are isomorphic if and only if $b \equiv \pm a \pmod{r}$.*

Proof. For each $p \in \mathbb{Z}$, we consider the diagram of groups $D_p = (\mathbb{Z}, \mathbb{Z}, \mathbb{Z} \times \mathbb{Z}, f, f_p)$, where f is the homomorphism from the first copy of $\mathbb{Z}$ to $\mathbb{Z} \times \mathbb{Z}$ which satisfies $f(1) = (1,0)$, and f_p is the homomorphism from the second copy of $\mathbb{Z}$ to $\mathbb{Z} \times \mathbb{Z}$ which satisfies $f_p(1) = (p,r)$. In order to prove that M_a and M_b are elementarily equivalent, it suffices to show that D_a and D_b are elementarily equivalent, or, by Corollary 1.2, that D_a/nD_a and D_b/nD_b are isomorphic for each integer $n \geq 1$.

For each integer $n \geq 1$, there exist an integer d which is prime to n and an integer k such that $b = ad + kr$. We consider the endomorphisms α, β, γ of $\mathbb{Z}$, $\mathbb{Z}$, $\mathbb{Z} \times \mathbb{Z}$ which satisfy $\alpha(1) = d$, $\beta(1) = 1$, $\gamma(1,0) = (d,0)$ and $\gamma(0,1) = (k,1)$. We have $\gamma(f(1)) = \gamma(1,0) = (d,0) = f(d) = f(\alpha(1))$ and $\gamma(f_a(1)) = \gamma(d,r) = (ad + kr, r) = (b,r) = f_b(1) = f_b(\beta(1))$. Consequently, (α, β, γ) is an homomorphism from D_a to D_b. It follows that (α, β, γ) induces an isomorphism from D_a/nD_a to D_b/nD_b, since α, β, γ induce automorphisms of $\mathbb{Z}$, $\mathbb{Z}$, $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$.

If $\alpha : M_a \to M_b$ is an isomorphism, then we have $\alpha(1,0) = (\varepsilon_1, 0)$ and $\alpha(a,r) = (\varepsilon_2 b, \varepsilon_2 r)$ with $\varepsilon_1 = \pm 1$ and $\varepsilon_2 = \pm 1$. Moreover, we have $b \equiv \varepsilon_1 \varepsilon_2 a \pmod{r}$ since $(\varepsilon_2 b - \varepsilon_1 a, \varepsilon_2 r) = \alpha(a,r) - a\,\alpha(1,0) = r\,\alpha(0,1)$ belongs to $r(\mathbb{Z}\times\mathbb{Z})$. Conversely, if $b=\varepsilon a+kr$ with $\varepsilon=\pm 1$ and $k\in\mathbb{Z}$, then we obtain an isomorphism $\alpha: M_a \to M_b$ by writing $\alpha(1,0)=(\varepsilon,0)$ and $\alpha(0,1)=(k,1)$. ■

Corollary 2.7. *The completely regular inverse semigroups S_a, S_b associated with the diagrams of groups D_a, D_b defined above are elementarily equivalent; they are isomorphic if and only if $b\equiv \pm a \pmod r$ or $ab\equiv \pm 1 \pmod r$.*

Proof. The elementary equivalence of S_a and S_b follows from the elementary equivalence of D_a and D_b.

The semigroups S_a and S_b are isomorphic if and only if $(\mathbb{Z}\times\mathbb{Z},\langle(1,0)\rangle,\langle(a,r)\rangle)$ is isomorphic to $(\mathbb{Z}\times\mathbb{Z},\langle(1,0)\rangle,\langle(b,r)\rangle)$ or to $(\mathbb{Z}\times\mathbb{Z},\langle(b,r)\rangle,\langle(1,0)\rangle)$. According to Proposition 2.6, the first case is realized if and only if $b\equiv\pm a \pmod r$.

In order to treat the second case, we consider two integers $d,m\in\mathbb{Z}$ such that $bd+rm=1$. We obtain an isomorphism α from $(\mathbb{Z}\times\mathbb{Z},\langle(1,0)\rangle,\langle(d,r)\rangle)$ to $(\mathbb{Z}\times\mathbb{Z},\langle(b,r)\rangle,\langle(1,0)\rangle)$ by writing $\alpha(1,0)=(b,r)$ and $\alpha(0,1)=(m,-d)$. So, the second case is realized if and only if $(\mathbb{Z}\times\mathbb{Z},\langle(1,0)\rangle,\langle(a,r)\rangle)$ is isomorphic to $(\mathbb{Z}\times\mathbb{Z},\langle(1,0)\rangle,\langle(d,r)\rangle)$. By Proposition 2.6, this property is true if and only if $d\equiv\pm a \pmod r$, and therefore if and only if $ab\equiv\pm 1 \pmod r$. ■

In contrast with Proposition 2.6, it follows from the proposition below that, for each integer $n\geq 1$ and for any subgroups S, T of $\mathbb{Z}^n$, $(\mathbb{Z}^n,S)$ and $(\mathbb{Z}^n,T)$ are elementarily equivalent if and only if they are isomorphic.

Proposition 2.8. *For any integers $m,n\geq 0$ and for any homomorphisms $g,h:\mathbb{Z}^m\to\mathbb{Z}^n$, the diagrams of groups $G=(\mathbb{Z}^m,\mathbb{Z}^n,g)$ and $H=(\mathbb{Z}^m,\mathbb{Z}^n,h)$ are elementarily equivalent if and only if they are isomorphic.*

Proof. For each homomorphism $g:\mathbb{Z}^m\to\mathbb{Z}^n$, there exist a unique integer $r\leq \inf(m,n)$, a unique integer $s\leq r$ and a unique sequence $(t_1,\dots,t_s)$ such that:

1) $\mathbb{Z}^m/\ker g\cong g(\mathbb{Z}^m)\cong\mathbb{Z}^r$;
2) $\mathbb{Z}^n/g(\mathbb{Z}^m)\cong\mathbb{Z}^{n-r}\times(\mathbb{Z}/t_1\mathbb{Z})\times\cdots\times(\mathbb{Z}/t_s\mathbb{Z})$;
3) $t_i\geq 2$ for $1\leq i\leq s$;
4) t_i divisible by t_{i+1} for $1\leq i\leq s-1$.

Moreover, there exist a basis $(u_1,\dots,u_m)$ of $\mathbb{Z}^m$ and a basis $(v_1,\dots,v_n)$ of $\mathbb{Z}^n$ such that $g(u_i)=t_iv_i$ for $1\leq i\leq s$, $g(u_i)=v_i$ for $s+1\leq i\leq r$ and $g(u_i)=0$ for $r+1\leq i\leq m$. So, the invariant $(r,t_1,\dots,t_s)$ determines $G=(\mathbb{Z}^m,\mathbb{Z}^n,g)$ up to isomorphism. This invariant is completely determined by the diagrams of groups G/kG for $k\geq 1$. ■

3. Endomorphisms of groups; modules

Endomorphisms of groups.

For each group G and for each endomorphism g of G, we consider the diagram of groups (G,g). Here, the language consists of one binary relational symbol for the endomorphism and one ternary relational symbol for the law of the group. As as matter of fact, we obtain the same notion of elementary equivalence if we interpret the endomorphism with a unary functional symbol and the law of the group with a binary functional symbol. The following result is a consequence of Corollary 1.2:

Corollary 3.1. *Let $n\geq 1$ be an integer, and let g, h be endomorphisms of $\mathbb{Z}^n$, with matrices A, B in the canonical basis of $\mathbb{Z}^n$. Then, $(\mathbb{Z}^n,+,g)$ and*

$(\mathbb{Z}^n, +, h)$ are elementarily equivalent if and only if, for each integer $k \geq 1$, the images of A and B modulo k are conjugate as matrices with coefficients in $\mathbb{Z}/k\mathbb{Z}$. ■

Example 3.2. In [7, p. 74], V. N. Remeslennikov proves that the matrices

$$A = \begin{pmatrix} 2 & 2 \\ -3 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 1 \\ -6 & 3 \end{pmatrix}$$

are not conjugate as matrices with coefficients in $\mathbb{Z}$ and that, for each integer $k \geq 1$, the images of A and B modulo k are conjugate as matrices with coefficients in $\mathbb{Z}/k\mathbb{Z}$. Consequently, the diagrams of groups $G = (\mathbb{Z}^2, +, g)$ and $H = (\mathbb{Z}^2, +, h)$ associated with A and B are elementarily equivalent, but not isomorphic.

In [7], Remeslennikov constructs from A and B two torsion-free nilpotent groups of class 4 with 2 generators which have the same finite images without being isomorphic. It follows from [5, Theorem 3.1] that these groups are not elementarily equivalent.

Example 3.3. For each integer $n \geq 2$, there exist two automorphisms g, h of $\mathbb{Z}^n$ such that $(\mathbb{Z}^n, +, g)$ and $(\mathbb{Z}^n, +, h)$ are elementarily equivalent, but not isomorphic. In fact, Remeslennikov proves in [8, §7] that there exist two $n \times n$ matrices A, B, invertible as matrices with coefficients in $\mathbb{Z}$, which are not conjugate as matrices with coefficients in $\mathbb{Z}$ and whose images modulo k are conjugate as matrices with coefficients in $\mathbb{Z}/k\mathbb{Z}$, for each integer $k \geq 1$. Concerning the case $n = 2$, Remeslennikov proves that these conditions are satisfied by the matrices

$$A = \begin{pmatrix} 21 & -8 \\ -8 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 2 \\ 12 & 23 \end{pmatrix}.$$

Remark 3.4. In contrast with Examples 3.2 and 3.3, it follows from Proposition 2.8 that, for any integers $m, n \geq 0$ and for any homomorphisms $g, h : \mathbb{Z}^m \to \mathbb{Z}^n$, the diagrams of groups $G = (\mathbb{Z}^m, \mathbb{Z}^n, g)$ and $H = (\mathbb{Z}^m, \mathbb{Z}^n, h)$ are elementarily equivalent if and only if they are isomorphic.

Modules.

Let R be a unitary ring. The language of a left (respectively right) R-module M is usually defined by considering a binary functional symbol for the addition of M and, for each $a \in R$, a unary functional symbol for the map $M \to M : x \to ax$ (respectively $x \to xa$). Anyhow, we obtain the same notion of elementary equivalence between R-modules if we interpret the addition with a ternary relational symbol and, for each $a \in R$, the map $x \to ax$ (respectively $x \to xa$) with a binary relational symbol. So, the following result is a consequence of Corollary 1.2:

Corollary 3.5. *Let R be a unitary ring and let M, N be two left (respectively right) R-modules whose additive structures are finitely generated abelian groups. Then, the following properties are equivalent: a) M and N are elementarily equivalent; b) M and N have the same finite images; c) For each integer $k \geq 1$, the R-modules M/kM and N/kN are isomorphic.*

Remark 3.6. Let $n \geq 1$ be an integer, let A, B be $n \times n$ matrices with coefficients in $\mathbb{Z}$ and let M, N be the structures of $\mathbb{Z}[X]$-modules which are defined on $\mathbb{Z}^n$ respectively by $Xx = Ax$ and $Xx = Bx$ for each $x \in \mathbb{Z}^n$. Then, M and N are isomorphic if and only if A and B are conjugate as matrices with coefficients in $\mathbb{Z}$. By Corollary 3.5, M and N are elementarily equivalent if and only if, for each integer $k \geq 1$, the images of A and B modulo k are conjugate as matrices with coefficients in $\mathbb{Z}/k\mathbb{Z}$.

Remark 3.7. If R is a Dedekind ring, it is possible to deduce Corollary 3.5 from [2, Corollary 5.3]. In particular, if R is the ring of integers of a extension of finite degree of $\mathbb{Q}$, it follows from Corollary 3.5 or [2, Corollary 5.3] that all ideals of R except (0) are elementarily equivalent as R-modules. Of course, if R is non principal, there exist such ideals which are not isomorphic.

References

[1] Chang, C. C. and H. J. Keisler, "Model Theory", Studies in Logic **73**, North-Holland, Amsterdam, 1973.

[2] Eklof, P. C. and E. R. Fisher, *The elementary theory of abelian groups*, Ann. Math. Logic **4** (1972), 115–171.

[3] Ljapin, E. S., "Semigroups," Translations of Math. Monographs **3**, Amer. Math. Soc., Providence, R. I., 1974.

[4] Oger, F., *Equivalence élémentaire entre groupes finis-par-abéliens de type fini*, Comment. Math. Helv. **57** (1982), 469–480.

[5] Oger, F., *Elementary equivalence and isomorphism of finitely generated nilpotent groups*, Communications in Algebra **12** (1984), 1899–1915.

[6] Oger, F., *Elementary equivalence and profinite completions: a characterization of finitely generated abelian-by-finite groups*, Proc. Amer. Math. Soc. **103** (1988), 1041–1048.

[7] Remeslennikov, V. N., *Conjugacy of subgroups in nilpotent groups*, Algebra I Logika **6** (1967), 61–76 (in Russian).

[8] Remeslennikov, V. N., *Groups that are residually finite with respect to conjugacy*, Siberian Math. Journal **12** (1971), 783–792.

[9] Segal, D., "Polycyclic Groups," Cambridge Tracts in Math. 82, Cambridge University Press, Cambridge, 1983.

[10] Zil'ber, B. I., *On the isomorphism and elementary equivalence of commutative semigroups*, Algebra and Logic **9** (1970), 400-402.

U.A. 753,
Département de Mathématiques
Université Paris 7
2 Place Jussieu
75 251 Paris cédex 05
France

Received January 23, 1991
and in final form July 1, 1991

Semigroup Forum Vol. 45 (1992) 332–341

RESEARCH ARTICLE

The Semigroup Generated by Certain Operators On the Congruence Lattice of A Clifford Semigroup

Mario Petrich

Communicated by N. R. Reilly

1. Introduction and Summary

Clifford semigroups, also known as semilattices of groups, form a class of semigroups nearest to both groups and semilattices (commutative idempotent semigroups). Much has been written about Clifford semigroups mostly on account of their transparent structure best described by an early theorem of Clifford, see ([3], II.2). In spite of their simplicity, they still provide a source of interesting problems.

For any regular semigroup S, there are two relations on the congruence lattice $\mathcal{C}(S)$ of S of great importance. For any $\rho \in \mathcal{C}(S)$, the kernel of ρ is the set of elements of S ρ-related to idempotents, the trace of ρ is the restriction of ρ to the set of idempotents of S. We denote by ρK and ρk (respectively ρT and ρt) the greatest and the least congruences on S with the same kernel (respectively trace) as ρ. We are generally interested in the set $\Gamma = \{K, k, T, t\}$ of operators on $\mathcal{C}(S)$, and, by iteration, in the semigroup Γ^+ or monoid Γ^* generated by Γ. If we fix ρ in $\mathcal{C}(S)$ and consider Γ^* acting on ρ as above, we obtain a network of congruences $\rho, \rho K, \rho k, \ldots$ ordered by inclusion.

Networks of congruences were first studied, for inverse semigroups, by Petrich and Reilly [6]. For the lattice of fully invariant congruences on the free completely regular semigroup of countably infinite rank, networks of congruences were investigated by Pastijn and Trotter [2] and Petrich and Reilly [7]. The latter amounts to analogues of congruence networks for the lattice of varieties of completely regular semigroups. In all these cases, only certain subsets of the entire network $\rho, \rho K, \rho k, \ldots$ are considered. The entire network for a congruence on a Rees matrix semigroup is determined in Petrich [5], as well as the monoid generated by it and the sublattice of $\mathcal{C}(S)$ generated by this monoid. The trend is obvious: the smaller the class of semigroups we consider, and the more explicitly described its structure and its congruences, the better we can describe the network, the monoid and the lattice evoked above.

For a Clifford semigroup S, we represent a congruence ρ on S by means of a normal pair (N, ξ), where N is the kernel of ρ and ξ is the congruence on the underlying semilattice Y of S induced by the trace of ρ. Using this representation, we are able to determine the defining relations for the generators Γ as Γ^+ acts on all Clifford semigroups. The success of this study is due primarily to the relatively simple description of congruences on a Clifford semigroup as indicated above. It is highly probable that our method could be applied to a similar analysis of congruences on more general classes of semigroups.

In Section 2, we state briefly most of the needed notation, terminology and preliminary results; the rest is relegated to the literature. A sequence of lemmas in Section 3 leads to the first main result of the paper which describes the semigroup generated by Γ with relations valid in all networks on Clifford semigroups. Examples in Section 4 illustrate certain aspects related to this study; the main example is used in the proof of the second principal result of this paper.

2. Congruences and Operators

For any semigroup S, we denote by $E(S)$ its set of idempotents and by $\mathcal{C}(S)$ its congruence lattice. For any set A, we use ε, ω and ι to denote the equality, the universal relation and the identity mapping on A, respectively, with A affixed as a subscript if necessary to avoid confusion. For a function $\varphi : A \to B$, we denote by $\overline{\varphi}$ the equivalence on A induced by φ. The free semigroup generated by A is denoted by A^+, the monoid by A^*.

Whenever possible, we follow the notation and terminology of [3].

By definition, a Clifford semigroup S is a completely regular inverse semigroup. That is, it is a union of its (maximal) subgroups and its idempotents commute. This implies that the set $E(S)$ of idempotents is in the center of S, that is idempotents commute with all elements of S. The Clifford structure theorem for them says that they are precisely (strong) semilattices of groups. This semigroup is described by the following construction.

Let Y be a semilattice. For each $\alpha \in Y$, let G_α be a group such that $G_\alpha \cap G_\beta = \emptyset$ if $\alpha \neq \beta$. For any $\alpha, \beta \in Y$ such that $\alpha \geq \beta$, let $\varphi_{\alpha,\beta} : G_\alpha \to G_\beta$ be a homomorphism and assume that

$$\varphi_{\alpha,\alpha} = \iota_{G_\alpha} ,$$
$$\varphi_{\alpha,\beta}\varphi_{\beta,\gamma} = \varphi_{\alpha,\gamma} \text{ if } \alpha > \beta > \gamma .$$

On $S = \bigcup_{\alpha \in Y} G_\alpha$ define a multiplication by: for $a \in G_\alpha$, $b \in G_\beta$,

$$a * b = (a\varphi_{\alpha,\alpha\beta})(b\varphi_{\beta,\alpha\beta}) .$$

Then S is a Clifford semigroup, to be denoted by $[Y; G_\alpha, \varphi_{\alpha,\beta}]$. Conversely, every Clifford semigroup can be so constructed.

In view of the above, when studying Clifford semigroups, we may restrict our attention to those of the form $S = [Y; G_\alpha, \varphi_{\alpha,\beta}]$. Since they form a class of inverse semigroups, the construction and basic properties of congruences on inverse semigroups can be readily specialized to the case at hand. In particular, we can represent congruences on S by congruence pairs, which on $[Y; G_\alpha, \varphi_{\alpha,\beta}]$ take on the following simple form.

For each $\alpha \in Y$, let N_α be a normal subgroup of G_α such that

$$N_\alpha\varphi_{\alpha,\beta} \subseteq N_\beta \text{ whenever } \alpha > \beta . \tag{1}$$

Let ξ be a congruence on Y such that

$$N_\beta\varphi_{\alpha,\beta}^{-1} \subseteq N_\alpha \text{ whenever } \alpha > \beta \text{ and } \alpha\,\xi\,\beta . \tag{2}$$

Define a relation ρ on S by: for $a \in G_\alpha$, $b \in G_\beta$,

$$a\,\rho\,b \text{ if } ab^{-1} \in N_{\alpha\beta},\ \alpha\,\xi\,\beta .$$

Then ρ is a congruence on $S = [Y; G_\alpha, \varphi_{\alpha,\rho}]$, and conversely every congruence on S can be so constructed for unique N_α and ξ.

For $\rho \in \mathcal{C}(S)$, we define the *kernel* of ρ and the *trace* of ρ, by

$$\ker\rho = \{a \in S \mid a\,\rho\, e \ \text{ for some } \ e \in E(S)\}, \quad \operatorname{tr}\rho = \rho|_{E(S)} \ ,$$

respectively. To each $\rho \in \mathcal{C}(S)$, we associate further four congruences on S by:

ρK, ρk – the greatest and the least congruences on S with the same kernel as ρ, respectively,

ρT, ρt – the greatest and the least congruences on S with the same trace as ρ, respectively,

$$\Gamma = \{K, k, T, t\} \ .$$

We thus arrive at four operators K, k, T and t on $\mathcal{C}(S)$. In order to express their values on a congruence ρ on S represented as above, it will be convenient to introduce some notation and terminology.

With N_α's satisfying condition (1), let $N = \bigcup\limits_{\alpha\in Y} N_\alpha$. Then N is a subsemigroup of S which we term *normal.* If both conditions (1) and (2) are satisfied, then we call the pair (N,ξ) *normal.* Therefore every congruence is obtained from a normal pair (N,ξ) by the formula

$$a\,\rho\, b \Leftrightarrow ab^{-1} \in N, \ \ \alpha\,\xi\,\beta \qquad (a \in G_\alpha, b \in G_\beta) \ ,$$

and we write $\rho \sim (N,\xi)$. In fact $N = \ker\rho$ and ξ is induced on Y by the trace of ρ. Hence a normal pair is essentially a congruence pair. We can transfer the above operators K, k, T and t from $\mathcal{C}(S)$ to the isomorphic copy $\mathcal{NP}(S)$ of normal pairs for S in an obvious way, where $\mathcal{NP}(S)$ is ordered componentwise.

For a normal subsemigroup N of S, let $N\kappa$ be the relation on Y given by

$$\alpha\, N\kappa\, \beta \ \text{ if for every } \ \gamma \in Y, \ \ N_{\alpha\beta\gamma}\varphi^{-1}_{\alpha\gamma,\alpha\beta\gamma} \subseteq N_{\alpha\gamma}$$
$$\text{and } \ N_{\alpha\beta\gamma}\varphi^{-1}_{\beta\gamma,\alpha\beta\gamma}N_{\beta\gamma} \ .$$

For a congruence ξ on Y, let $\xi\theta$ be the subset of S such that for every $\alpha \in Y$,

$$\xi\theta \cap G_\alpha = \cup\{\ker\overline{\varphi_{\alpha,\beta}} \mid \alpha \geq \beta, \ \ \alpha\,\xi\,\beta\} \ .$$

We can derive easily the lattice operations in $\mathcal{NP}(S)$ from ([4], Theorem 4.4) to obtain

$$(N,\xi) \wedge (N',\xi') = (N \cap N', \xi \cap \xi') \ , \tag{3}$$
$$(N,\xi) \vee (N',\xi') = (M, \xi \vee \xi') \ , \tag{4}$$

where

$$M = \{a \in S \mid a \in G_\alpha, \ \ \alpha \geq \beta, \ \ \alpha\,\xi \vee \xi'\,\beta, \ \ a\varphi_{\alpha,\beta} \in N_\beta N'_\beta \ \text{ for some } \ \beta \in Y\}.$$

For a congruence ρ on a Clifford semigroup S, or equivalently, for its corresponding normal pair (N,ξ), we call the set $\{\rho, \rho K, \rho k, \rho T, \rho t, \rho KT, \ldots\}$ partially ordered by inclusion the *network of* ρ, equivalently of (N,ξ).

As usual, for a Clifford semigroup S, we denote by σ, τ and μ the least group, the greatest idempotent pure and the greatest idempotent separating congruences on S (note $\mu = \mathcal{H}$), respectively.

Throughout we fix a Clifford semigroup $S = [Y; G_\alpha, \varphi_{\alpha,\beta}]$ with the semilattice of idempotents E.

3. The semigroup Γ^+/Σ

Six lemmas lead to the main result of this section which provides a system of representatives Ω for the congruence on Γ^+ generated by certain relations Σ thereby implicitly providing an isomorphic copy of Γ^+/Σ.

Lemma 1. *Let $\xi \in \mathcal{C}(Y)$ and N be a normal subsemigroup of S. Then (N,ξ) is a normal pair if and only if $\xi \subseteq N\kappa$.*

Proof. Necessity. Let $\alpha\,\xi\,\beta$ and $\gamma \in Y$. Then $\alpha\gamma\,\xi\,\beta\gamma$ and $\alpha\gamma \geq \alpha\beta\gamma$, $\alpha\gamma\,\xi\,\alpha\beta\gamma$, $\beta\gamma \geq \alpha\beta\gamma$, $\beta\gamma\,\xi\,\alpha\beta\gamma$ and thus, by definition of a normal pair,

$$N_{\alpha\beta\gamma}\varphi^{-1}_{\alpha\gamma,\alpha\beta\gamma} \subseteq N_{\alpha\gamma}, \quad N_{\alpha\beta\gamma}\varphi^{-1}_{\beta\gamma,\alpha\beta\gamma} \subseteq N_{\beta\gamma}$$

so that $\alpha\, N\kappa\, \beta$.

Sufficiency. Let $\alpha,\beta \in Y$ be such that $\alpha \geq \beta$ and $\alpha\,\xi\,\beta$. Hence $\alpha\, N\kappa\,\beta$ and thus $N_\beta\varphi^{-1}_{\alpha,\beta} \subseteq N_\alpha$ where we have set $\gamma = \alpha$. But then (N,ξ) is a normal pair.

Lemma 2. *For $(N,\xi) \in \mathcal{NP}(S)$, we have*

$$(N,\xi)K = (N,N\kappa), \qquad (N,\xi)k = (N,\varepsilon)\ ,$$
$$(N,\xi)T = (S,\xi) \quad , \qquad (N,\xi)t = (\xi\theta,\xi)\ .$$

Proof. Straightforward. ∎

It follows from the above lemma that the operators k and T are order preserving. That this holds for t will follow from the first part of the next lemma. Example 1 in Section 4 will show that this property fails for K.

Lemma 3. *Let $\xi,\xi' \in \mathcal{C}(Y)$.*
 (i) *If $\xi \subseteq \xi'$, then $\xi\theta \subseteq \xi'\theta$.*
 (ii) $\xi \subseteq \xi\theta\kappa$.
 (iii) $\xi\theta = \xi\theta\kappa\theta$.
 (iv) $\varepsilon\theta = E$.

Proof. (i) This follows directly from the definition of θ.
 (ii) Straightforward.
 (iii) We have

$$(\xi\theta,\xi)tK = (\xi\theta,\xi)K = (\xi\theta,\xi\theta\kappa)\ ,$$
$$(\xi\theta,\xi)tKt = (\xi\theta,\xi\theta\kappa)t = (\xi\theta\kappa\theta,\xi\theta\kappa)\ ,$$

and hence $\xi\theta\kappa\theta \subseteq \xi\theta$. By part (ii), we have $\xi \subseteq \xi\theta\kappa$ and thus by part (i), $\xi\theta \subseteq \xi\theta\kappa\theta$ and equality prevails.
 (iv) Obvious. ∎

Lemma 4. *Let N be a normal subsemigroup of S.*
 (i) $N\kappa\theta \subseteq N$.
 (ii) $S\kappa = \omega$.

Proof. Straightforward. ∎

The content of Lemma 4 is an analogue of a part of Lemma 3. Examples in Section 4 will show that the analogues of the remaining assertions of Lemma 3 fail.

Lemma 5. *Let* $\rho \in \mathcal{C}(S)$.

(i) $\rho TK = \omega \sim (S, \omega)$.
(ii) $\rho TKt = \sigma \sim (\omega\theta, \omega)$.
(iii) $\rho kT = \rho Tk = \mu \sim (S, \varepsilon)$.
(iv) $\rho TKtk = \sigma \wedge \mu \sim (\omega\theta, \varepsilon)$.
(v) $\rho kt = \varepsilon \sim (E, \varepsilon)$.
(vi) $\rho ktK = \tau \sim (E, E\kappa)$.
(vii) $\rho ktKT = \tau \vee \mu \sim (S, E\kappa)$.
(viii) $\rho tK = \rho tKt \sim (\xi\theta, \xi\theta\kappa)$.

Proof. Indeed,

(i) $(N, \xi)TK = (S, \xi)K = (S, S\kappa) = (S, \omega)$ by Lemma 4(ii),
(v) $(N, \xi)kt = (N, \varepsilon)t = (\varepsilon\theta, \varepsilon) = (E, \varepsilon)$ by Lemma 3(iv),
(viii) $(N, \xi)tKt = (\xi\theta, \xi)Kt = (\xi\theta, \xi\theta\kappa)t = (\xi\theta\kappa\theta, \xi\theta\kappa)$
$= (\xi\theta, \xi\theta\kappa) = (N, \xi)tK$ by Lemma 3(iii).

The proof of the remaining items is even simpler and is omitted. ∎

In view of Lemma 5(i)–(vii), we may drop ρ in those expressions and use the following notation:

$$\begin{aligned} &\varepsilon = kt, \quad &\tau = ktK, \quad &\tau \vee \mu = ktKt, \quad \mu = kT\ , \\ &\omega = TK, \quad &\sigma = TKt, \quad &\sigma \wedge \mu = TKtk\ . \end{aligned}$$

Note that the second line is obtained from the first, except for μ, by the transformation $K \leftrightarrow t$, $T \leftrightarrow k$. Finally let

$$\Delta = \{\varepsilon, \sigma, \mu, \tau, \sigma \wedge \mu, \tau \vee \mu, \omega\}\ .$$

Lemma 6. *Operators* Γ *satisfy the following relations* Σ:

(i) $K^2 = kK = K, \quad k^2 = Kk = k,$
$t^2 = Tt = t, \quad T^2 = tT = T.$
(ii) $KTK = TKT = TK,$
$tkt = ktk = kt.$
(iii) $tKt = tK,$
(iv) $kT = Tk.$

Proof. (i) This follows directly from the definition of the operators Γ.
(ii) For any $\rho \in \mathcal{C}(S)$, we have $\rho KT = \omega$ and $\rho kt = \varepsilon$ from which the claims follow immediately.
(iii) This follows from Lemma 3(iii) and its proof.
(iv) This follows from Lemma 5(iii). ∎

We are now ready for the first principal result of the paper.

Theorem 1. *The set*

$$\begin{aligned} \Omega = \{K, KT, Kt, KtK,\ &Ktk, KtKT,\ k, \\ t,\ tk,\ tK,\ &tKT,\ T\} \cup \Delta \end{aligned}$$

is a system of representatives for the congruence on Γ^+ *generated by the relations* Σ. *The set* Ω *is given the product of representatives. The elements of* Δ *act as right zeros of* Ω.

Proof. The multiplication of elements of Δ by elements of Γ follows from Lemma 5 in a straightforward manner. This also shows that $\Delta\Gamma = \Delta$.

For each w in Γ^+ we consider four words wK, wk, wT and wt. If the last letter of w is K or k, because of relations $K^2 = kK = K$ and $k^2 = Kk = k$,

wK and wk have representatives of the same length as w, so it suffices to consider the words wT and wt. Symmetrically for the case when the last letter of w is either T or t. In the remaining cases we search among the relations Σ whether the resulting word can be replaced by a shorter one and replace Tk by kT. If in this process we reach an element of Δ, we may stop because of the first paragraph of the proof. We now apply this procedure successively to K, k, T and t, thereby obtaining the words in Ω. That no two elements of Ω are related by our congruence will follow from Example 1 in the next section.

Routine analysis of the multiplication table of Ω leads to the following result.

Proposition 1. *The $\mathcal{D}$-structure of Ω has the form of Diagram 1 with the partial order of $\mathcal{D}$-classes and the egg-box picture of each $\mathcal{D}$-class (horizontal rows are the $\mathcal{R}$-classes, vertical rows the $\mathcal{L}$-classes). The $\mathcal{D}$-class $\{KT, Kt\}$ is irregular. The principal factor of tK is isomorphic to the Rees matrix semigroup $\mathcal{M}^0(\{1,2\}, e, \{1,2,3\}; P)$, where*

$$P = \begin{bmatrix} e & e \\ e & e \\ e & 0 \end{bmatrix}.$$

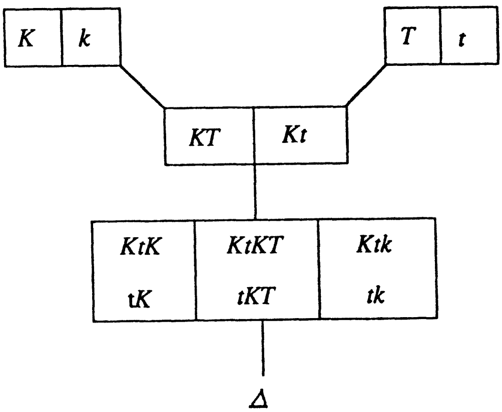


DIAGRAM 1

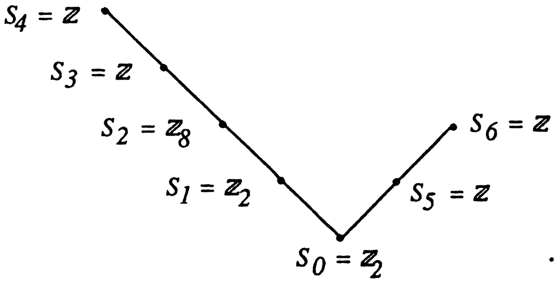
4. Examples

The first example provides an instance of a Clifford semigroup S for which the semigroup generated by the operators K, k, T and t on $\mathcal{C}(S)$ is isomorphic to the semigroup Γ^+/Σ constructed in the preceding section. In addition, it illustrates some of the processes hidden in the general theory. We may go one step further by constructing the partially ordered set underlying the semigroup generated by the operators, that is the congruence network, and the lattice generated by this partially ordered set. The second and third examples will be used to show that certain normal subsemigroups and congruences in the projections of the partially ordered set $\{\rho, \rho K, \dots\}$ are incomparable in general.

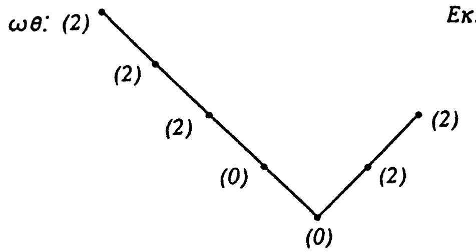
We represent Clifford semigroups in this section by means of a diagram of the underlying semilattice with a group for each vertex. These groups are the additive groups of integers $\mathbb{Z}$ and some of its homomorphic images $\mathbb{Z}_n$. All connecting homomorphisms are the natural homomorphisms $\mathbb{Z} \to \mathbb{Z}$, $\mathbb{Z} \to \mathbb{Z}_m$, $\mathbb{Z}_m \to \mathbb{Z}_n$.

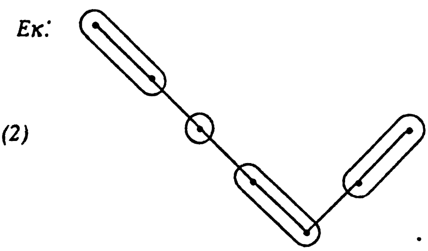
Normal subsemigroups are represented by cyclic subgroups associated with vertices and congruences on the semilattice by grouping some of the vertices. This is certainly not the most concise way of representing these quantities but is definitely the most adequate for visualizing and checking.

Example 1. Let S be the Clifford semigroup given by

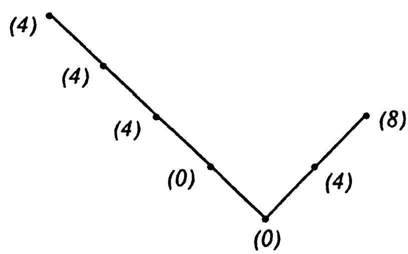


Then

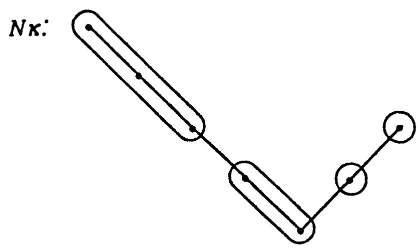




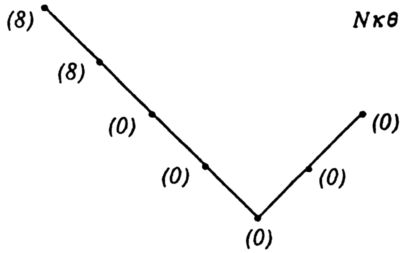
Let N be defined by



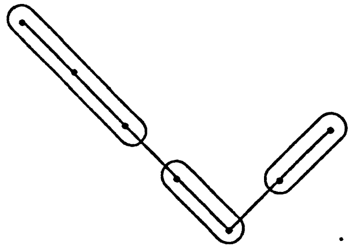
Then



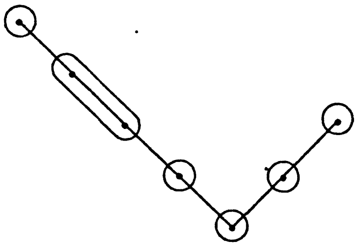
$N\kappa\theta$:



$N\kappa\theta\kappa$:

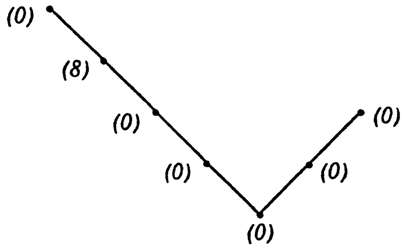


Let ξ be defined by

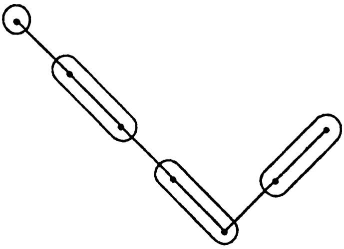


Then

$\xi\theta$:



$\xi\theta\kappa$:



Note that $\xi \subseteq N\kappa$ which by Lemma 1 gives that (N,ξ) is a normal pair for S.

In order to depict the partially ordered set $\{\rho, \rho K, \ldots\}$, we consider the list of elements Ω in Theorem 1. We replace each element of Ω^1 by the corresponding normal pair and first form the partially ordered sets of the first and of the second projections of the elements of Ω^1. The elements ρK, ρk, ρT and ρt are given the normal pair represenation in Lemma 2; the representation of elements of Δ can be found in Lemma 5. The remaining elements of Ω^1 are:

$$\begin{array}{rlrll}
\rho KT \sim (S, N\kappa), & & \rho tk \sim (\xi\theta, \varepsilon), & & \\
\rho Kt \sim (N\kappa\theta, N\kappa), & & \rho tK \sim (\xi\theta, \xi\theta\kappa), & & (5) \\
\rho KtK \sim (N\kappa\theta, N\kappa\theta\kappa), & & & & (6) \\
\rho Ktk \sim (N\kappa\theta, \varepsilon). & & \rho tKT \sim (S, \xi\theta\kappa), & & \\
\rho KtKT \sim (S, N\kappa\theta\kappa). & & & &
\end{array}$$

After drawing the partially ordered set of the first and the second projections, we plot the pairs corresponding to the elements of Ω^1.

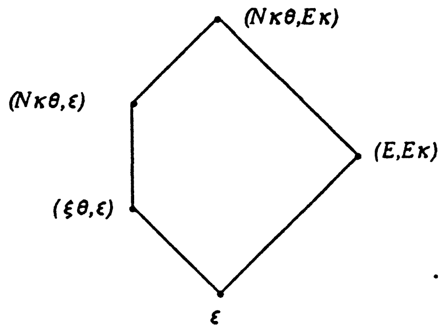
Observe that the sets of the first and the second projections of the elements Ω^1 are represented in this example by distinct elements, that is to say there is no collapsing. It follows that there is no collapsing in the set Ω^1 in our example. From the results of the preceding section and this discussion, we derive the second main result of the paper as follows.

Theorem 2. *Let S be a Clifford semigroup. The semigroup $\Omega(S)$ generated by the operators K, k, T and t on $C(S)$ is a homomorphic image of Ω. For the semigroup S in Example 1 above, we have $\Omega(S) \cong \Omega$.*

Continuing with Example 1, we may attempt to construct the lattice L generated by this congruence network. The meet of two normal pairs is by components, as we noted in (3), whereas the formula for their join is more complicated, see (4). That the join is not by components in our example shows:

$$(\xi\theta, \varepsilon) \vee (E, E\kappa) = (N\kappa\theta, E\kappa)$$

giving the join $\rho tk \vee \tau$, with $N\kappa\theta \supset \xi\theta \supset E$. Using this, we easily check that the following diagram depicts a sublattice of L.



Therefore L is not modular. Note that the above lattice is contained in two trace classes of L

We denote a congruence on Y by listing its nontrivial classes in parentheses, for example (01)(23) stands for the congruence with classes

$$\{0,1\}, \{2,3\}, \{4\}, \{5\}, \{6\} \ .$$

Of course, we have $E \subset N$. In our example, $E\kappa = (01)(34)(56)$ and $N\kappa = (01)(234)$ which are incomparable under inclusion. It follows that the function κ is not monotone. Furthermore, here we have $N\kappa\theta\kappa = (01)(234)(56)$ so that $N\kappa \neq N\kappa\theta\kappa$. It is instructive to contrast these two phenomena with the properties of θ, see Lemma 3(i) and (iii).

We continue with the same notation and in addition represent normal subsemigroups by $(N_0, N_1, \ldots, N_k)$.

Example 2. Let S be the Clifford semigroup given by

$$S_3 = \mathbb{Z}$$
$$S_2 = \mathbb{Z}$$
$$S_1 = \mathbb{Z}_8$$
$$S_0 = \mathbb{Z}_2$$

with natural epimorphisms and let

$$N = ((1),(4),(4),(4)), \quad \xi = (12)$$

in terms of cyclic subgroups and nontrivial congruence classes. Then

$$\omega\theta = ((0),(2),(2),(2)), \quad E\kappa = (23), \quad N\kappa = (123)$$

so that $\xi \subseteq N\kappa$ and by Lemma 1, (N,ξ) is a normal pair for S.

Example 3. Let S be the Clifford semigroup given by

$$\begin{array}{l} S_2 = \mathbb{Z} \\ S_1 = \mathbb{Z} \\ S_0 = \mathbb{Z}_2 \end{array}$$

with natural epimorphisms and let $N = ((0),(2),(4))$. Then

$$\begin{aligned} &N\kappa = (01), \quad N\kappa\theta = ((0),(2),(0)), \quad N\kappa\theta\kappa = (01) \\ &E\kappa = (12), \quad \varepsilon\theta\kappa = (12) \end{aligned}$$

with (N,ε) a normal pair for S.

References

[1] Pastijn, F. and M. Petrich, *Congruences on regular semigroups*, Trans. Amer. Math. Soc. **295** (1986), 607–633.

[2] Pastijn, F. and P. G. Trotter, *Lattices of completely regular semigroup varieties*, Pacific J. Math. **119** (1985), 191–214.

[3] Petrich, M., "Inverse semigroups", Wiley, New York, 1984.

[4] Petrich, M., *Congruences on strong semilattices of regular simple semigroups*, Semigroup Forum **37** (1988), 167–199.

[5] Petrich, M., "Congruence networks for completely simple semigroups", J. Austral. Math. Soc. (to appear).

[6] Petrich, M. and N. R. Reilly, *A network of congruences on an inverse semigroup*, Trans. Amer. Math. Soc. **270** (1982), 309–325.

[7] Petrich, M. and N. R. Reilly, *Semigroups generated by certain operators on varieties of completely regular semigroups*, Pacific J. Math. **132** (1988), 151–175.

Department of Mathematics
Simon Fraser University
Burnaby, British Columbia
Canada, V5A 1S6

Received February 26, 1991
and in final form May 17, 1991

Semigroup Forum Vol. 45 (1992) 342–363

RESEARCH ARTICLE

Integral Transforms of Vector Measures on Semigroups with Applications to Spectral Operators

Werner J. Ricker

Communicated by R. Nagel

Many moment problems, which include problems of characterization of classical integral transforms (e.g. Fourier, Laplace, Stieltjes, Poisson), fit into the following scheme. Given are locally compact Hausdorff spaces Ω and Γ (with Γ usually a topological semigroup), an $\mathbb{R}$ or $\mathbb{C}$-valued kernel K on $\Omega \times \Gamma$, a locally convex Hausdorff space X (briefly, lcs) and a function $f : \Gamma \to X$. The problem is whether there exists a regular vector measure $m : \mathcal{B}(\Omega) \to X$ (where $\mathcal{B}(\Omega)$ are the Borel sets on Ω) such that

$$f(\gamma) = \int_\Omega K(w,\gamma)dm(w), \quad \gamma \in \Gamma, \tag{1}$$

and, possibly, how to construct such a measure. For example, the choice $\Omega = [0,\infty)$, $\Gamma = (0,\infty)$ and $K(w,\gamma) = \exp(-w\gamma)$, for $(w,\gamma) \in \Omega \times \Gamma$, leads to the problem of characterizing certain Laplace transforms. If $\Omega = \Gamma = \mathbb{R}$ and $K(w,\gamma) = \exp(iw\gamma)$, for $(w,\gamma) \in \Omega\times\Gamma$, we have the problem of Fourier-Stieltjes transforms.

An important application of such characterizations arises in the theory of scalar-type spectral operators (briefly, scalar operators). Given a bounded operator T on a Banach space Y, say, we choose for X the lcs $L(Y)$ consisting of all continuous linear operators of Y into itself (equipped with the strong operator topology). Under appropriate conditions on T it is often possible to construct a suitable X-valued function $f_T : \Gamma \to X$, where Γ is usually some "natural" semigroup containing $\sigma(T)$, such that T is a scalar operator iff f_T is the K-transform (for an appropriate kernel K) of some spectral measure $m : \mathcal{B}(\Omega) \to X$, that is, f_T is of the form (1). Of course, to ensure that m is a spectral measure, as distinct from general operator-valued measures, it is necessary that the kernel $K : \Omega \times \Gamma \to \mathbb{C}$ satisfy certain algebraic criteria with respect to the semigroup Γ.

Many particular examples of the above set-up are well known; some of these will be listed in the sequel. The aim of this paper is to formulate an abstract framework which unifies the particular examples scattered about in the literature and which provides a systematic approach to the problem which they collectively encompass. It is hoped that the suggested framework is versatile enough to accommodate most of the important cases which occur in practice, but which is not too general to lose contact with "real" examples.

1. FORMULATION AND CLASSICAL EXAMPLES

Let Ω and Γ be locally compact Hausdorff spaces which are σ-compact. The space of all regular $\mathbb{C}$-valued measures on $\mathcal{B}(\Omega)$ is denoted by $\mathcal{M}(\Omega)$; it is the dual Banach space to $C_0(\Omega)$, the space of continuous functions on Ω which vanish at infinity (equipped with the supremum norm). The space $\mathcal{M}(\Gamma)$ is defined similarly.

Let $K : \Omega \times \Gamma \to \mathbb{C}$ be a function which is measurable with respect to the σ-algebra generated by all product sets $E \times F$ with $E \in \mathcal{B}(\Omega)$, $F \in \mathcal{B}(\Gamma)$, and such that

(i) for $\gamma \in \Gamma$, the function $K(\cdot,\gamma) : w \mapsto K(w,\gamma)$, for $w \in \Omega$, is bounded and

(ii) for $w \in \Omega$, the function $K(w,\cdot) : \gamma \mapsto K(w,\gamma)$, for $\gamma \in \Gamma$, is continuous.

If X is a lcs, then a set function $m : \mathcal{B}(\Omega) \to X$ which is σ- additive is called an X-valued measure or a vector measure. Regularity of m means that the complex measure

$$\langle m, x' \rangle : E \mapsto \langle m(E),\ x' \rangle, \quad E \in \mathcal{B}(\Omega),$$

is regular, for each $x' \in X'$ (the continuous dual space of X). In this case we define the K-transform $\widehat{m} : \Gamma \to X$, of m, by the formula

$$\widehat{m}(\gamma) = \int_\Omega K(w,\gamma)dm(w), \quad \gamma \in \Gamma. \tag{2}$$

It follows from property (i) that this integral is defined, [7; II §3, Lemma 1]. Similarly, if $n : \mathcal{B}(\Gamma) \to X$ is a regular vector measure, we define $\widetilde{n}$ by

$$\widetilde{n}(w) = \int_\Gamma K(w,\gamma)dn(\gamma), \tag{3}$$

for every $w \in \Omega$ for which this integral exists. The X-valued function $\widetilde{n}$, when defined on all of Ω, is called the inverse K-transform of n.

Let $\widehat{\mathcal{M}}(\Gamma) = \{\widehat{\mu};\ \mu \in \mathcal{M}(\Omega)\}$ be the family of K-transforms of all scalar measures on Ω. It is assumed that

(iii) the K-transform map $\mu \mapsto \widehat{\mu}$, $\mu \in \mathcal{M}(\Omega)$, is injective.

In the sequel, it will be assumed that a vector space $\mathcal{F}$ of functions on Γ is given such that

(iv) $\mathcal{F} \subset C(\Gamma)$ and, for every $w \in \Omega$, the function $K(w,\cdot)$ belongs to $\mathcal{F}$.

Furthermore, a Radon measure $\lambda \geq 0$ on $\mathcal{B}(\Omega)$, a directed index set $\mathcal{I}$ and, for every $i \in \mathcal{I}$, a linear map $L_i : \mathcal{F} \to \mathbb{C}^\Omega$ are assumed given such that

(v) a function $f : \Gamma \to \mathbb{C}$ belongs to $\widehat{\mathcal{M}}(\Gamma)$ if and only if $f \in \mathcal{F}$, $L_i(f) \in L^1(\lambda)$, for every $i \in \mathcal{I}$, and $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} < \infty$;

(vi) if $f \in \mathcal{F}$ and $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} < \infty$, then

$$\lim_{i\in\mathcal{I}} \int_\Omega K(w,\gamma)L_i(f)(w)d\lambda(w) = f(\gamma), \quad \gamma \in \Gamma. \tag{4}$$

It is also assumed that, for every $i \in \mathcal{I}$, a product measurable function $H_i : \Omega \times \Omega \to \mathbb{C}$ is given such that

(vii) for every $w \in \Omega$, the function $\xi \mapsto H_i(w,\xi)$, $\xi \in \Omega$, is bounded and measurable; for every $\xi \in \Omega$, the function $w \mapsto H_i(w,\xi)$, $w \in \Omega$, is λ-integrable and

$$\sup\{\| H_i(\cdot,\xi) \|_1;\ i \in \mathcal{I},\ \xi \in \Omega\} < \infty; \tag{5}$$

for every set $E \in \mathcal{B}(\Omega)$, the function

$$\xi \mapsto \int_E H_i(w,\xi)d\lambda(w), \quad \xi \in \Omega,$$

is measurable and, for every measure $\mu \in \mathcal{M}(\Omega)$,

$$L_i(\widehat{\mu})(w) = \int_\Omega H_i(w,\xi)d\mu(\xi), \quad w \in \Omega. \tag{6}$$

Finally, a vector subspace $\mathcal{S}(\Gamma)$ of $\mathcal{M}(\Gamma)$ is given such that

(viii) $\mathcal{F} \subset L^1(\nu)$, for every $\nu \in \mathcal{S}(\Gamma)$, and the space $\mathcal{S}(\Gamma)$ separates points of $\mathcal{F}$;

(ix) if $\nu \in \mathcal{S}(\Gamma)$, then the function

$$w \mapsto \int_\Gamma |K(w,\gamma)|d|\nu|(\gamma), \quad w \in \Omega,$$

is bounded and measurable and, for every $E \in \mathcal{B}(\Gamma)$, the function

$$w \mapsto \int_E K(w,\gamma)d\nu(\gamma), \quad w \in \Omega,$$

is also measurable;

(x) the vector space $\widetilde{\mathcal{S}}(\Gamma) = \{\widetilde{\nu};\ \nu \in \mathcal{S}(\Gamma)\}$ is dense in $C_0(\Omega)$, and

(xi) if $f \in \mathcal{F}$ and $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} < \infty$, then

$$\lim_{i\in\mathcal{I}} \int_\Gamma \left(\int_\Omega K(w,\gamma)L_i(f)(w)d\lambda(w)\right) d\nu(\gamma) = \int_\Gamma f(\gamma)d\nu(\gamma), \tag{7}$$

for every $\nu \in \mathcal{S}(\Gamma)$.

Remark 1.(a). The assumption (v) implies that $\widehat{\mathcal{M}}(\Gamma) \subset \mathcal{F}$. However, it is not assumed that $\widehat{\mathcal{M}}(\Gamma) = \mathcal{F}$.

(b) For every $\nu \in \mathcal{S}(\Gamma)$, the inverse K-transform $\widetilde{\nu}$ exists because $\{K(w,\cdot);\ w \in \Omega\} \subset \mathcal{F} \subset L^1(\nu)$.

(c) In a majority of cases, the identity (7) follows from (vi) and from the Lebesgue Dominated Convergence theorem (if $\mathcal{I}$ is cofinal with $\mathbb{N}$) or from the fact that (4) holds uniformly on compact subsets of Γ. At any rate, in most of the classical situations – a selection of which will be presented later in the section – all the assumptions made are satisfied; some hold trivially, others follow from well known theorems. Firstly however, we state some simple, but useful consequences of the above assumptions. ∎

Proposition 1.1. *If* $f \in \mathcal{F}$ *and if* $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} < \infty$, *then*

$$\lim_{i\in\mathcal{I}} \int_\Omega \widetilde{\nu}(w)L_i(f)(w)d\lambda(w) = \int_\Gamma f(\gamma)d\nu(\gamma), \tag{8}$$

for every $\nu \in \mathcal{S}(\Gamma)$.

Proof. Let $M > 0$ satisfy $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} \le M$. If $\nu \in \mathcal{S}(\Gamma)$ and $\alpha_\nu > 0$ is such that

$$\sup_{w\in\Omega} \int_\Gamma |K(w,\gamma)|.d|\nu|(\gamma) \le \alpha_\nu,$$

then it follows that

$$\int_\Omega \int_\Gamma |K(w,\gamma)L_i(f)(w)|d\lambda(w)d|\nu|(\gamma) \le \alpha_\nu M,$$

for each $i \in \mathcal{I}$. The result then follows from the definition of $\widetilde{\nu}$, the Fubini theorem and the identity (7). ∎

Proposition 1.2. *If $f \in \mathcal{F}$ and if* $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} < \infty$, *then*

$$\lim_{i \in \mathcal{I}} \int_\Omega \psi(w) L_i(f)(w) d\lambda(w)$$

exists, for every $\psi \in C_0(\Omega)$.

Proof. Let $\psi \in C_0(\Omega)$ and $\epsilon > 0$ be given. Since $\widetilde{\mathcal{S}}(\Gamma)$ is dense in $C_0(\Omega)$, there is $\varphi \in \widetilde{\mathcal{S}}(\Gamma)$ such that $\| \varphi - \psi \|_\infty < \epsilon/3M$, where M satisfies $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} \le M$. It follows from Proposition 1.1 that there exist $i_0 \in \mathcal{I}$ and $j_0 \in \mathcal{I}$ such that

$$\begin{aligned} |\int_\Omega \psi(L_i(f) - L_j(f))d\lambda| &\le |\int_\Omega (\psi - \varphi)L_i(f)d\lambda| + |\int_\Omega \varphi(L_i(f) - L_j(f))d\lambda| + \\ &+ |\int_\Omega (\varphi - \psi)L_j(f)d\lambda| < 2M \| \varphi - \psi \|_\infty + \epsilon/3 < \epsilon, \end{aligned}$$

for all $i \ge i_0$ and $j \ge j_0$. ∎

Proposition 1.3. *If, for every $\gamma \in \Gamma$, the function $K(\cdot, \gamma)$ vanishes at infinity, then the set of measures on Γ with finite support can be taken for $\mathcal{S}(\Gamma)$.*

Proof. The conditions (viii) and (ix) are trivially satisfied. If $\nu = \sum_{n=1}^N \alpha_n \epsilon_{\gamma_n}$, where $\alpha_n \in \mathbb{C}$ and $\gamma_n \in \Gamma$, $1 \le n \le N$, are arbitrary and ϵ_{γ_n} is the Dirac point mass at γ_n, then

$$\widetilde{\nu}(w) = \sum_{n=1}^N \alpha_n K(w, \gamma_n), \quad w \in \Omega.$$

It is then clear from the hypothesis of the proposition that $\widetilde{\mathcal{S}}(\Gamma) \subset C_0(\Omega)$. Since $\widetilde{\mathcal{S}}(\Gamma)$ is the linear span in $C_0(\Omega)$ of $\{K(\cdot, \gamma);\ \gamma \in \Gamma\}$, it follows from the Hahn-Banach theorem and (iii) that $\widetilde{\mathcal{S}}(\Gamma)$ is dense in $C_0(\Omega)$. Hence, (x) is satisfied.

Finally, if $f \in \mathcal{F}$ satisfies $\sup\{\| L_i(f) \|_1;\ i \in \mathcal{I}\} < \infty$, then

$$\int_\Gamma \left(\int_\Omega K(w, \gamma) L_i(f)(w) d\lambda(w) \right) d\nu(\gamma) = \sum_{n=1}^N \alpha_n \int_\Omega K(w, \gamma_n) L_i(f)(w) d\lambda(w),$$

for each $i \in \mathcal{I}$. The identity (7) then follows from (4). ∎

In the remainder of this section examples of classical kernels are given which will be used in the sequel.

Example 1.1 (Fourier-Stieltjes transform). Let Ω be a σ-compact locally compact Abelian group, Γ its dual group and λ be (a choice of) Haar measure on Ω. Define a kernel $K : \Omega \times \Gamma \to \mathbb{C}$ by

$$K(w, \gamma) = \langle -w, \gamma \rangle, \quad (w, \gamma) \in \Omega \times \Gamma.$$

Let $\mathcal{I}$ be an index set directed by a relation $\le$ and, for every $i \in \mathcal{I}$, let u_i and w_i be functions on Ω and Γ respectively, with the properties:

(1) $u_i \ge 0$, $\int_\Omega u_i d\lambda = 1$.

(2) $w_i \in C_{00}(\Gamma)$ and $w_i(\Gamma) \subseteq [0, 1]$.

(3) For every neighbourhood V of 0 in Ω and $\epsilon > 0$, there is $i_0 \in \mathcal{I}$ such that, for every i with $i_0 \leq i$, $\int_V u_i d\lambda > 1 - \epsilon$ and $w \notin V$ implies $u_i(w) < \epsilon$.

(4) $\lim_i w_i(\gamma) = 1$ uniformly with respect to γ on compact subsets for Γ.

(5) $w_i(\gamma) = \int_\Omega u_i(w) K(w,\gamma) d\lambda(w)$, $\gamma \in \Gamma$ and $u_i(w) = \int_\Gamma w_i(\gamma)\langle w,\gamma\rangle d\gamma$, $w \in \Omega$, where $d\gamma$ is a choice of Haar measure on Γ.

The existence of such a kernel was proved in [1] and [16] generalizing the Fejér means on the unit circle, say.

Let $\mathcal{F}$ be the space of all bounded, continuous functions on Γ. Define linear maps $L_i : \mathcal{F} \to \mathbb{C}^\Omega$, $i \in \mathcal{I}$, by

$$L_i(f)(w) = \int_\Gamma f(\gamma) w_i(\gamma)\langle w,\gamma\rangle d\gamma, \quad w \in \Omega,$$

for each $f \in \mathcal{F}$. It is well known that the requirement (v) is satisfied (see [16] or [20]) and (vi) follows from the Inversion Theorem for Fourier transforms and the property (4).

For $i \in \mathcal{I}$, define $H_i : \Omega \times \Omega \to \mathbb{C}$ by

$$H_i(w,\xi) = u_i(w - \xi), \quad (w,\xi) \in \Omega \times \Omega.$$

Then it is easily verified that (vii) is valid.

Finally, let $\mathcal{S}(\Gamma) = L^1(d\gamma)$, considered as a subspace of $\mathcal{M}(\Gamma)$. The Riemann-Lebesgue lemma guarantees that $\widetilde{\mathcal{S}}(\Gamma) \subset C_0(\Omega)$ and the properties (viii) - (xi) are easily verified or are well known facts. ∎

Example 1.2 (Laplace transform). Let $\Omega = [0,\infty)$ and $\Gamma = (0,\infty)$. Define a kernel $K : \Omega \times \Gamma \to \mathbb{C}$ by

$$K(w,\gamma) = \exp(-w\gamma), \quad (w,\gamma) \in \Omega \times \Gamma.$$

Let λ denote the sum of Lebesgue measure on Ω and the Dirac point mass at zero.

Let $\mathcal{F}$ be the subspace of $C^\infty(\Gamma)$ consisting of those functions f for which $f(\infty) = \lim_{t\to\infty} f(t)$ exists. Define the Widder inversion operators $L_k : \mathcal{F} \to \mathbb{C}^\Omega$, $k \in \mathbb{N}$, by

$$L_k(f)(w) = (-1)^k (k!)^{-1} (k/w)^{k+1} f^{(k)}(k/w), \quad w > 0,$$

and

$$L_k(f)(0) = f(\infty),$$

for each $f \in \mathcal{F}$. It is well known that the requirement (v) is satisfied (see VII Theorem 12a of [19]) and (vi) follows from the Inversion Theorem of D. V. Widder, [19; VII Theorem 11b].

For $k \in \mathbb{N}$, define $H_k : \Omega \times \Omega \to \mathbb{C}$ by

$$H_k(w,\xi) = (k!)^{-1}(k/w)^{k+1}\xi^k \exp(-k\xi/w), \quad w > 0, \xi \geq 0,$$

and

$$H_k(0,\xi) = \chi_{\{0\}}(\xi), \quad \xi \geq 0.$$

Using the identity $(k!)^{-1}\int_0^\infty s^k \exp(-s) ds = 1$, $k = 0, 1, \ldots$, it is easily verified that (vii) holds (see also VII, §7 of [19]).

Since $\{K(\cdot,\gamma);\ \gamma \in \Gamma\} \subset C_0(\Omega)$, it follows from Proposition 1.3. that the set of measures on Γ with finite support can be taken for $\mathcal{S}(\Gamma)$. ∎

Example 1.3 (Stieltjes transform). Let $\Omega = [0,\infty)$ and $\Gamma = (0,\infty)$. Define a kernel $K : \Omega \times \Gamma \to \mathbb{C}$ by

$$K(w,\gamma) = (w+\gamma)^{-1}, \quad (w,\gamma) \in \Omega \times \Gamma.$$

Let λ denote the sum of Lebesgue measure on Ω and the Dirac point mass at zero.

Let $\mathcal{F}$ be the subspace of $C^\infty(\Gamma)$ consisting of those functions f for which $A_f = \lim_{t\to 0+} tf(t)$ exists. Define the Widder inversion operators $L_k : \mathcal{F} \to \mathbb{C}^\Omega$, $k \in \mathbb{N}$, by

$$L_k(f)(w) = (k!(k-2)!)^{-1}(-w)^{k-1}(w^k f(w))^{(2k-1)}, \quad w > 0, \ k = 2,3,\ldots,$$
$$L_1(f)(w) = (wf(w))^{(1)}, \quad w > 0,$$

and

$$L_k(f)(0) = A_f, \quad k \in \mathbb{N},$$

for each $f \in \mathcal{F}$. It is well known that the requirement (v) is satisfied (see [4] or VIII, §16 of [19]) and (vi) follows from the Inversion Theorem of D. V. Widder, [19; VIII Theorem 15].

For $k \in \mathbb{N}$, define $H_k : \Omega \times \Omega \to \mathbb{C}$ by

$$H_k(w,\xi) = c_k w^{k-1}\xi^k(w+\xi)^{-2k}, \quad w > 0, \ \xi \geq 0,$$

and

$$H_k(0,\xi) = \chi_{\{0\}}(\xi), \quad \xi \geq 0,$$

where $c_1 = 1$, $c_k = B(k-1,k+1)$, $k = 2,3,\ldots$, and B is the Beta function. Using the identity $\int_0^\infty w^{k-1}\xi^k(w+\xi)^{-2k}dw = B(k,k)$ for each $\xi > 0$, it is easily verified that (vii) holds (see [4] for example).

Since $\{K(\cdot,\gamma);\ \gamma \in \Gamma\} \subset C_0(\Omega)$, it follows from Proposition 1.3 that the set of measures on Γ with finite support can be taken for $\mathcal{S}(\Gamma)$. ■

Example 1.4 (Hausdorff moments). Let $\Omega = [0,1]$ and $\Gamma = \{0,1,2,\ldots\}$. Define a kernel $K : \Omega \times \Gamma \to \mathbb{C}$ by

$$K(w,\gamma) = w^\gamma, \quad (w,\gamma) \in \Omega \times \Gamma.$$

Let λ denote the sum of Lebesgue measure on Ω and the Dirac point mass at 1.

Let $\mathcal{F}$ be the subspace of $\ell^\infty(\Gamma)$ consisting of those functions f such that $f(\infty) = \lim_{n\to\infty} f(n)$ exists. For $f \in \ell^\infty(\Gamma)$, define operators $\Delta^k f$ and $\lambda_f(k,m)$ by

$$\Delta^k f(n) = \sum_{m=0}^{k} (-1)^m \binom{k}{m} f(n+k-m), \quad k,n = 0,1,2,\ldots,$$

and

$$\lambda_f(k,m) = \binom{k}{m}(-1)^{k-m}\Delta^{k-m}f(m), \quad k = 0,1,2,\ldots; 0 \leq m \leq k.$$

Then define linear maps $L_k : \mathcal{F} \to \mathbb{C}^\Omega$, $k = 0,1,2,\cdots$, by

$$L_k(f)(w) = (k+1)\lambda_f(k,[kw]), \quad 0 \leq w < 1,$$

and

$$L_k(f)(1) = f(k),$$

for each $f \in \mathcal{F}$, where [ ] denotes integer part. The requirement (v) is known, [19; III Theorem 2b], and (vi) follows from a well known inversion theorem, [19; III Theorem 3].

Define $H_k : \Omega \times \Omega \to \mathbb{C}$, $k = 0, 1, 2, \ldots$, by

$$H_k(w, \xi) = (k+1)\binom{k}{[kw]}\xi^{[kw]}(1-\xi)^{k-[kw]}, \quad 0 \le w < 1, \ \xi \in \Omega,$$

and

$$H_k(1, \xi) = \xi^k, \quad \xi \in \Omega.$$

Using the identity $\int_0^1 \binom{k}{[kw]}\xi^{[kw]}(1-\xi)^{k-[kw]}dw = k^{-1}(1-\xi^k)$, $k \in \mathbb{N}$, it can be verified that (vii) holds.

Since Ω is compact, $C_0(\Omega) = C(\Omega)$. Hence, $\{K(\cdot, \gamma);\ \gamma \in \Gamma\} \subset C_0(\Omega)$, and it follows from Proposition 1.3 that the set of measures on Γ with finite support can be taken for $\mathcal{S}(\Gamma)$. ■

Example 1.5 (Poisson integrals). Let $\Omega = (-\pi, \pi]$ and $\Gamma = \{z \in \mathbb{C}; |z| < 1\}$. Let λ be normalized Lebesgue measure on Ω. Define a kernel $K : \Omega \times \Gamma \to \mathbb{C}$ by

$$K(w, \gamma) = Re((\exp(iw) + \gamma)/(\exp(iw) - \gamma)), \quad (w, \gamma) \in \Omega \times \Gamma.$$

If $\gamma = r\exp(i\xi)$ with $0 \le r < 1$, $\xi \in (-\pi, \pi]$, then in more familiar notation we have $K(w, r\exp(i\xi)) = P_r(\xi - w)$ where

$$P_R(\theta) = (1 - R^2)/(1 + R^2 - 2R\cos(\theta)), \quad R \in [0, 1), \ \theta \in (-\pi, \pi],$$

is the Poisson kernel; it is an approximate identity for $L^1(\lambda)$.

Let $\mathcal{F}$ be the linear space of all $\mathbb{C}$-valued harmonic functions in Γ. Let $\mathcal{I}$ denote the interval [0,1) directed by the usual order induced from $\mathbb{R}$ and, for $r \in \mathcal{I}$, let $L_r : \mathcal{F} \to \mathbb{C}^\Omega$ be defined by

$$L_r(f)(w) = f(r\exp(iw)), \quad w \in \Omega,$$

for each $f \in \mathcal{F}$. For each $r \in \mathcal{I}$, let $H_r : \Omega \times \Omega \to \mathbb{C}$ be defined by

$$H_r(\theta, \varphi) = (2\pi)^{-1}P_r(\theta - \varphi), \quad (\theta, \varphi) \in \Omega \times \Omega.$$

Since we identify Ω with the circle group and consider only 2π-periodic functions on Ω, the space $\mathcal{S}(\Gamma)$ can be taken as the finitely supported measures on Γ (c.f. Proposition 1.3), in which case properties (viii) - (xi) hold (see also Remark 1 (c)). Properties (i) - (vii) hold from well known properties of Poisson integrals of complex measures (see [15], for example). ■

Example 1.6 (Orthogonal systems: discrete case). Let $\Omega = [a, b]$ be a compact interval and $\Gamma = \{0, 1, 2 \ldots\}$. Let $\{\psi_\gamma; \gamma \in \Gamma\}$ be an orthonormal system (in $L^2(\Omega)$) of continuous functions on Ω satisfying the following conditions, [8]:

(I) not all the ψ_γ vanish at a or at b,

(II) the linear span of $\{\psi_\gamma; \gamma \in \Gamma\}$ is dense in $C_0(\Omega) = C(\Omega)$,

(III) there exists a regular summation method $\{\lambda_{nN}; 0 \le n \le N-1,\ N = 1,2,\ldots\}$ such that

$$\sup\left\{\int_a^b \left|\sum_{n=0}^{N-1} \lambda_{nN}\psi_n(s)\overline{\psi_n(t)}\right| ds; t \in \Omega,\ N = 1,2\ldots\right\} < \infty.$$

If all the ψ_γ vanish at a then we put $\Omega = (a,b]$ and if all the ψ_γ vanish at b then we put $\Omega = [a,b)$.

Define $K : \Omega \times \Gamma \to \mathbb{C}$ by

$$K(w,\gamma) = \psi_\gamma(w), \quad (w,\gamma) \in \Omega \times \Gamma.$$

Let λ denote Lebesgue measure on Ω.

Let $\mathcal{F} = \mathbb{C}^\Gamma$ and $\mathcal{I} = \{1,2,\ldots\}$. Define linear maps $L_k : \mathcal{F} \to \mathbb{C}^\Omega$, $k \in \mathcal{I}$, by

$$L_k(f)(w) = \sum_{n=0}^{k-1} \lambda_{kn} f(n)\overline{\psi_n(w)}, \quad w \in \Omega,$$

for each $f \in \mathcal{F}$. The L_k are the maps σ_k given in [8] with ψ_n replaced by $\overline{\psi_n}$.

Elements of $\widehat{\mathcal{M}}(\Gamma)$ are characterized by the maps L_k, $k \in \mathcal{I}$, see [8; Theorem 1]. Furthermore, the density of the span of $\{\psi_\gamma; \gamma \in \Gamma\}$ in $C_0(\Omega)$ implies that the transform map $\widehat{\cdot} : \mathcal{M}(\Omega) \to C(\Gamma)$ is injective. If $\sup\{\| L_k(f) \|_1 ; k \in \mathcal{I}\}$ is finite, then for each $\gamma \in \Gamma$,

$$\lim_{k\to\infty} \int_\Omega K(w,\gamma)L_k(f)(w)d\lambda(w) = \lim_{k\to\infty} \int_a^b \psi_\gamma(w)L_k(f)(w)dw =$$
$$= \lim_{k\to\infty} \lambda_{k\gamma} f(\gamma) = f(\gamma);$$

see Section 2 of [8].

Define $H_k : \Omega \times \Omega \to \mathbb{C}$, $k \in \mathcal{I}$, by

$$H_k(w,\xi) = \sum_{n=0}^{k-1} \lambda_{kn}\overline{\psi_n(w)}\psi_n(\xi), \quad (w,\xi) \in \Omega \times \Omega,$$

($H_k(w,\xi) = K_k(w,\xi)$ in the notation of [8]).

Observe that $H_k(w,\cdot) = \sum_{n=0}^{k-1} \lambda_{kn}\overline{\psi_n(w)}\psi_n(\cdot)$ is bounded and measurable for each $k \in \mathcal{I}$, $w \in \Omega$. Also, if $\xi \in \Omega$ is fixed, then

$$\| H_k(\cdot,\xi) \|_1 = \int_a^b \left|\sum_{n=0}^{k-2} \lambda_{kn}\overline{\psi_n(w)}\psi_n(\xi)\right| dw$$

Hence, by Assumption (III) on the family $\{\psi_\gamma;\ \gamma \in \Gamma\}$ we have $H_k(\cdot,\xi) \in L^1(\lambda)$ and

$$\sup\{\| H_k(\cdot,\xi) \|_1; \xi \in \Omega,\ k \in \mathcal{I}\} < \infty.$$

If $\mu \in \mathcal{M}(\Omega)$, then for $w \in \Omega$ and $k \in \mathcal{I}$ we have

$$\int_\Omega H_k(w,\xi)d\mu(\xi) = \int_\Omega \sum_{n=0}^{k-1} \lambda_{kn}\overline{\psi_n(w)}\psi_n(\xi)d\mu(\xi) =$$
$$= \sum_{n=0}^{k-1} \lambda_{kn}\overline{\psi_n(w)} \int_\Omega \psi_n(\xi)d\mu(\xi) =$$
$$= \sum_{n=0}^{k-1} \lambda_{kn}\overline{\psi_n(w)}\widehat{\mu}(n) = L_k(\widehat{\mu})(w).$$

Finally, since $\{K(\cdot,\gamma);\ \gamma \in \Gamma\} = \{\psi_\gamma;\ \gamma \in \Gamma\}$ is a subset of $C_0(\Omega)$, Proposition 1.3 implies that the finitely supported measures on Γ can be taken for $\mathcal{S}(\Gamma)$. ■

Example 1.7 (Orthogonal systems: continuous case). Let $\Omega = (A, B)$ and $\Gamma = (C, D)$ be intervals in $\mathbb{R}$ (possibly unbounded). Let $K : \Omega \times \Gamma \to \mathbb{C}$ and $H : \Omega \times \Gamma \to \mathbb{C}$ be bounded, continuous functions. Let $\{\lambda_k;\ k = 1, 2, \ldots\}$ be a sequence of $\mathbb{R}$-valued, continuous functions in $L^1(\Gamma)$ (Lebesgue measure on Γ) and λ be Lebesgue measure on Ω.

Let $\mathcal{F} = \{f \in C(\Gamma); f \text{ is bounded }\}$ and $\mathcal{I} = \{1, 2, \ldots\}$. Define linear maps $L_k : \mathcal{F} \to \mathbb{C}^\Omega$, $k \in \mathcal{I}$, by

$$L_k(f)(w) = \int_\Gamma \lambda_k(\gamma) f(\gamma) H(w,\gamma)\, d\gamma, \quad w \in \Omega,$$

for each $f \in \mathcal{F}$; these are the maps σ_N of [9].

For $\phi \in L^1(\Gamma)$, considered as a member of $\mathcal{M}(\Gamma)$, define $\widehat{\phi} : \Omega \to \mathbb{C}$ by (2) with $dm = \phi(\cdot)d\gamma$. Also define $H_k : \Omega \times \Omega \to \mathbb{C}$, $k \in \mathcal{I}$, by

$$H_k(w,\xi) = \int_\Gamma \lambda_k(s) K(\xi, s) H(w, s)\, ds, \quad (w,\xi) \in \Omega \times \Omega.$$

The following conditions are *assumed* to hold; see Section 1 of [9].

[I] If $f \in \mathcal{F}$ is such that $L_k(f) \in L^1(\lambda)$, $k \in \mathcal{I}$, then

$$f(\gamma) = \lim_{k\to\infty} \int_\Omega K(w,\gamma) L_k(f)(w)\, d\lambda(w), \quad \gamma \in \Gamma.$$

[II] The set $\{\widehat{\phi};\ \phi \in L^1(\Gamma)\}$ is dense in $C_0(\Omega)$.

[III] $\sup\{\| H_k(\cdot,\xi) \|_1;\ k \in \mathcal{I},\ \xi \in \Omega\} < \infty$.

It should be pointed out that Assumption IV of [9] is not needed to characterize elements of $\widehat{\mathcal{M}}(\Gamma)$. It is only used to ensure that certain functions are transforms of L^1-functions (rather than measures), a problem which we do not consider in this article. Anyway, under Assumption (I)–(III) the maps L_k, $k \in \mathcal{I}$, do characterize $\widehat{\mathcal{M}}(\Gamma)$ and guarantee that the K-transform $\widehat{\cdot} :$ $\mathcal{M}(\Omega) \to C(\Gamma)$ is injective, [9; Theorem 1].

For $w \in \Omega$ and $k \in \mathcal{I}$ fixed, the conditions on K, H and the inclusion $\{\lambda_k;\ k \in \mathcal{I}\} \subset L^1(\Gamma)$ imply that each function $H_k(w,\cdot)$ is bounded and measurable. If $\mu \in \mathcal{M}(\Omega)$, then an application of Fubini's theorem yields

$$\begin{aligned}\int_\Omega H_k(w,\xi)\, d\mu(\xi) &= \int_\Omega \int_\Gamma \lambda_k(s) K(\xi,s) H(w,s)\, ds\, d\mu(\xi) = \\ &= \int_\Gamma \lambda_k(s) H(w,s) \left(\int_\Omega K(\xi,s)\, d\mu(\xi) \right) ds = \\ &= \int_\Gamma \lambda_k(s) H(w,s) \widehat{\mu}(s)\, ds = L_k(\widehat{\mu})(w),\end{aligned}$$

for each $w \in \Omega$.

Let $\mathcal{S}(\Gamma) = \{\nu = \phi d\gamma;\ \phi \in L^1(\Gamma)\}$. Then certainly $\mathcal{F} \subset L^1(\nu)$ for all $\nu \in \mathcal{S}(\Gamma)$. If $f \in \mathcal{F}$ and

$$\int_\Gamma f(\gamma)\phi(\gamma)\, d\gamma = 0, \quad \phi \in L^1(\Gamma),$$

then f is the zero element of $(L^1(\Gamma))' = L^\infty(\Gamma)$, that is, $f = 0$ a.e. The continuity of f implies that $f \equiv 0$. Hence, $\mathcal{S}(\Gamma)$ seperates the points of $\mathcal{F}$.

If $\nu = \phi d\gamma \in \mathcal{S}(\Gamma)$ and $M > 0$ satisfies $|K(w,\gamma)| \le M$, for all $w \in \Omega$, $\gamma \in \Gamma$, then

$$w \mapsto \int_\Gamma |K(w,\gamma)| d|\nu|(\gamma), \quad w \in \Omega,$$

has $\| \cdot \|_\infty$-norm not greater than $M \| \phi \|_1$ and hence, is a bounded measurable function. The linear space $\{\widehat{\nu};\ \nu \in \mathcal{S}(\Gamma)\}$ is dense in $C_0(\Omega)$ by Assumption II.

If $f \in \mathcal{F}$ satisfies $\sup_{k \in \mathcal{I}} \| L_k(f) \|_1 = R < \infty$, then the sequence $g_k(\gamma) = \int_\Omega K(w,\gamma) L_k(f)(w) d\lambda(w)$, $\gamma \in \Gamma$, converges pointwise to f on Γ (by Assumption I). Furthermore, if $\nu = \phi d\gamma$ is a member of $\mathcal{S}(\Gamma)$, then

$$|g_k(\gamma)| \le MR, \quad \gamma \in \Gamma.$$

Since the constant function $\gamma \mapsto MR$, $\gamma \in \Gamma$, is ν-integrable it follows by Dominated Convergence that

$$\lim_{k\to\infty} \int_\Gamma \int_\Omega K(w,\gamma) L_k(f)(w) d\lambda(w) = \int_\Gamma f(\gamma) d\nu(\gamma).$$

So, the conditions (i) – (xi) are satisfied. ∎

2. VECTOR-VALUED TRANSFORMS

In this section we consider the problem of when a given *vector*-valued function is the K-transform of some regular *vector* measure. The notation and set up throughout is as in Section 1.

Let X be a quasi-complete lcs. It is assumed that there is a linear space, $\mathcal{F}(X)$, of functions $f : \Gamma \to X$ such that

(xii) for every $f \in \mathcal{F}(X)$, the functions $\langle f(\cdot), x' \rangle$ belong to $\mathcal{F}$, for each $x' \in X'$, and

(xiii) for every $f \in \mathcal{F}(X)$, the linear functional

$$x' \mapsto L_i(\langle f(\cdot), x' \rangle)(w), \quad x' \in X', \tag{9}$$

is $\sigma(X',X)$-continuous, for each $w \in \Omega$ and $i \in \mathcal{I}$, where $\sigma(X',X)$ denotes the weak * topology on X'.

The unique element of X representing the functional (9) is denoted by $L_i(f)(w)$, for each $w \in \Omega$ and $i \in \mathcal{I}$. Hence, for every $i \in \mathcal{I}$, there is a linear map $L_i : \mathcal{F}(X) \to X^\Omega$ given by

$$L_i(f) : w \mapsto L_i(f)(w), \quad w \in \Omega,$$

for each $f \in \mathcal{F}(X)$. It follows from (xii) and the inclusion $\mathcal{F} \subset C(\Gamma)$, that each member of $\mathcal{F}(X)$ is weakly continuous.

Remark 2. A space $\mathcal{F}(X)$ with properties (xii) and (xiii) certainly exists for each of the examples given in Section 1. For Fourier-Stieltjes transforms, the space of all bounded, weakly continuous, X-valued functions on Γ can be taken for $\mathcal{F}(X)$. The space of all X-valued functions f on Γ, having weak derivatives of all orders (in the sense of Definition 3.2.3 in [2]) and such that $f(\infty) = \lim_{t\to\infty} f(t)$ exists weakly in X, can be taken for $\mathcal{F}(X)$ in the case of Laplace transforms. For Stieltjes transforms, the space of all X-valued functions f on Γ having weak derivatives of all orders and such that $A_f = \lim_{t\to 0+} tf(t)$ exists weakly in X, can be taken for $\mathcal{F}(X)$. For Hausdorff moments the space of all bounded, X-valued functions f on Γ for which $f(\infty) = \lim_{n\to\infty} f(n)$ exists weakly in X, can be taken as the space $\mathcal{F}(X)$. For Poisson integrals, $\mathcal{F}(X)$ is the space of X-valued harmonic functions on the open unit disc in $\mathbb{C}$. In Example 1.6 we can take $\mathcal{F}(X) = X^\Gamma$ and for Example 1.7 the space of bounded, weakly continuous functions on Γ can be taken for $\mathcal{F}(X)$. ∎

Lemma 2.1. *If $\mu : \mathcal{B}(\Omega) \to X$ is a regular vector measure, then the function $\widehat{\mu}$ belongs to $\mathcal{F}(X)$ and*

$$L_i(\widehat{\mu})(w) = \int_\Omega H_i(w,\xi)d\mu(\xi), \quad w \in \Omega, \tag{10}$$

for each $i \in \mathcal{I}$.

Proof. If $x' \in X'$, then $\langle \mu, x'\rangle \in \mathcal{M}(\Omega)$ and $\langle \mu, x'\rangle\widehat{\ }(\gamma) = \langle \widehat{\mu}(\gamma),\ x'\rangle$, $\gamma \in \Gamma$. Since $\langle \mu, x'\rangle\widehat{\ } \in \widehat{\mathcal{M}}(\Gamma) \subset \mathcal{F}$, for each $x' \in X'$, it follows that $\langle \widehat{\mu}(\cdot),\ x'\rangle$ belongs to $\mathcal{F}$ for each $x' \in X'$. Furthermore, for each $w \in \Omega$ and $i \in \mathcal{I}$, it follows from (6) that

$$L_i(\langle \widehat{\mu}(\cdot), x'\rangle)(w) = \langle \int_\Omega H_i(w,\xi)d\mu(\xi),\ x'\rangle, \quad x' \in X',$$

which shows that the linear functional

$$x' \mapsto L_i(\langle \widehat{\mu}(\cdot),\ x'\rangle)(w), \quad x' \in X',$$

is $\sigma(X', X)$-continuous and represented by the element $\int_\Omega H_i(w,\xi)d\mu(\xi)$ of X. Accordingly, $\widehat{\mu} \in \mathcal{F}(X)$ and (10) is valid. ∎

Let Σ be a σ-algebra of subsets of a set Λ and μ be a σ-additive measure on Σ, either $\mathbb{C}$-valued or non-negative extended real-valued. Let X be a lcs. A weakly measurable function $f : \Lambda \to X$ is called Pettis integrable with respect to μ if the $\mathbb{C}$-valued function

$$\langle f, x'\rangle : \xi \mapsto \langle f(\xi),\ x'\rangle, \quad \xi \in \Lambda,$$

is μ-integrable for each $x' \in X'$, and if, for every $E \in \Sigma$, there is an element $(f\mu)(E)$ of X such that

$$\langle (f\mu)(E),\ x'\rangle = \int_E \langle f,\ x'\rangle d\mu, \quad x' \in X'.$$

The Orlicz-Pettis lemma implies that the indefinite integral of f with respect to μ, that is, the map $f\mu : E \mapsto (f\mu)(E)$, $E \in \Sigma$, is an X-valued measure. The element $(f\mu)(\Lambda)$ is denoted simply by $\mu(f)$.

So, after this slight digression, let $f \in \mathcal{F}(X)$ be a function such that the function $L_i(f)$ is Pettis λ-integrable, for each $i \in \mathcal{I}$. Then, for every $i \in \mathcal{I}$, the linear map $\Phi_i(f) : C_0(\Omega) \to X$ is defined by

$$\Phi_i(f)(\psi) = \int_\Omega \psi(w) L_i(f)(w)d\lambda(w), \quad \psi \in C_0(\Omega). \tag{11}$$

The collection of linear maps (11), $i \in \mathcal{I}$, is said to be weakly equicompact if the subset,

$$\{\Phi_i(f)(\psi);\ \psi \in C_0(\Omega),\ \| \psi \|_\infty \le 1,\ i \in \mathcal{I}\}, \tag{12}$$

of X, is relatively weakly compact; [6].

Proposition 2.2. *A function $f \in \mathcal{F}(X)$ is the K-transform of a regular, X-valued vector measure on $\mathcal{B}(\Omega)$, if and only if, $L_i(f)$ is Pettis λ-integrable for each $i \in \mathcal{I}$, and the family of maps (11), $i \in \mathcal{I}$, is weakly equicompact. If such a measure exists, then it is unique.*

Proof. Suppose that $f = \widehat{\mu}$ for some regular measure $\mu : \mathcal{B}(\Omega) \to X$. Fix $i \in \mathcal{I}$. If $\psi \in C_{00}(\Omega)$, the subspace of $C_0(\Omega)$ whose elements have compact support, then it follows from (vii) of Section 1 that

$$\xi \mapsto \int_\Omega \psi(w) H_i(w,\xi)\, d\lambda(w), \quad \xi \in \Omega, \tag{13}$$

is bounded and measurable and hence, is μ-integrable. If x^i_ψ denotes the element of X given by

$$\int_\Omega \left(\int_\Omega \psi(w) H_i(w,\xi)\, d\lambda(w) \right) d\mu(\xi), \tag{14}$$

then it follows that

$$\langle x^i_\psi, x' \rangle = \int_\Omega \int_\Omega H_i(w,\xi)\, d(\psi\lambda)(w)\, d\langle \mu(\xi),\ x' \rangle, \quad x' \in X'.$$

If $\alpha = \sup\{\| H_i(\cdot,\xi) \|_1;\ i \in \mathcal{I},\ \xi \in \Omega\}$, then

$$\int_\Omega \int_\Omega |H_i(w,\xi)|\, d|\psi\lambda|(w)\, d|\langle \mu, x' \rangle|(\xi) \le \alpha \| \psi \|_\infty |\langle \mu, x' \rangle|(\Omega),$$

for each $x' \in X'$. Lemma 2.1 and the Fubini theorem imply that

$$\langle x^i_\psi, x' \rangle = \int_\Omega \psi(w) \langle L_i(f)(w),\ x' \rangle\, d\lambda(w), \quad x' \in X'.$$

Let $W_i = \{x^i_\psi;\ \psi \in C_{00}(\Omega),\ \| \psi \|_\infty \le 1\}$. Since (13) is measurable and bounded by $\alpha \| \psi \|_\infty$, it follows from (14) that $W_i \subset \alpha \overline{bco}\mu(\mathcal{B}(\Omega))$. Hence, W_i is relatively weakly compact, [7; IV Theorem 6.1]. Lemma 3 of [6] then implies that $L_i(f)$ is Pettis λ-integrable and

$$x^i_\psi = \int_\Omega \psi(w) L_i(f)(w)\, d\lambda(w), \quad \psi \in C_{00}(\Omega).$$

In fact, it has been shown that $x^i_\psi \in \alpha \overline{bco}\mu(\mathcal{B}(\Omega))$, for all $i \in \mathcal{I}$ and all $\psi \in C_{00}(\Omega)$ satisfying $\| \psi \|_\infty \le 1$, where *bco* denotes balanced convex hull and the bar denotes closure. Since $C_0(\Omega)$ is the uniform closure of $C_{00}(\Omega)$ it follows that

$$\int_\Omega \psi(w) L_i(f)(w)\, d\lambda(w) \in \alpha \overline{bco}\mu(\mathcal{B}(\Omega)),$$

for every $i \in \mathcal{I}$ and every $\psi \in C_0(\Omega)$ satisfying $\| \psi \|_\infty \le 1$. Hence, the maps (11), $i \in \mathcal{I}$, are weakly equicompact.

Conversely, suppose that $L_i(f)$ is Pettis λ-integrable for each $i \in \mathcal{I}$ and the maps (11), $i \in \mathcal{I}$, are weakly equicompact. The boundedness of (12) implies the existence of $M_{x'} > 0$, for each $x' \in X'$, such that

$$|\langle \Phi_i(f)(\psi), x' \rangle| \le M_{x'}, \quad i \in \mathcal{I},\ \| \psi \|_\infty \le 1.$$

Accordingly, if $x' \in X'$, then

$$\| L_i(\langle f(\cdot), x'\rangle) \|_1 = \sup_{\|\psi\|_\infty \le 1} \left| \int_\Omega \psi L_i(\langle f(\cdot), x'\rangle) d\lambda \right| = \sup_{\|\psi\|_\infty \le 1} |\langle \Phi_i(f)(\psi), x'\rangle| \le M_{x'},$$

for each $i \in \mathcal{I}$. It follows from Proposition 1.2 that for each $\psi \in C_0(\Omega)$,

$$\lim_i \langle \Phi_i(f)(\psi), x'\rangle = \lim_i \int_\Omega \psi L_i(\langle f(\cdot), x'\rangle) d\lambda,$$

exists for each $x' \in X'$. Hence, for each $\psi \in C_0(\Omega)$, it follows that $\{\Phi_i(f)(\psi);\ i \in \mathcal{I}\}$ is a weak Cauchy net contained in the weakly compact set

$$\| \psi \|_\infty \overline{bco}\{\Phi_i(f)(\varphi);\ i \in \mathcal{I}, \| \varphi \|_\infty \le 1\},$$

and is therefore, weakly convergent. Define

$$\Phi(f)(\psi) = (\text{weak}) \lim_{i \in \mathcal{I}} \Phi_i(f)(\psi), \quad \psi \in C_0(\Omega).$$

The so defined linear map $\Phi(f) : C_0(\Omega) \to X$ is weakly compact and it follows from Proposition 1 of [6] that there is a unique regular measure $\mu : \mathcal{B}(\Omega) \to X$ such that

$$\Phi(f)(\psi) = \int_\Omega \psi d\mu, \quad \psi \in C_0(\Omega).$$

For each $x' \in X'$ and $\nu \in \mathcal{S}(\Gamma)$, Proposition 1.1 implies that

$$\langle \int_\Omega \widetilde{\nu} d\mu, x'\rangle = \lim_i \langle \Phi_i(f)(\widetilde{\nu}), x'\rangle = \int_\Gamma \langle f, x'\rangle d\nu. \tag{15}$$

It follows from the Fubini theorem, (ix) of Section 1 and the definition of $\widetilde{\nu}$ that

$$\langle \int_\Omega \widetilde{\nu} d\mu, x'\rangle = \int_\Gamma \langle \widehat{\mu}, x'\rangle d\nu, \quad x' \in X'.$$

Then (15) implies that for each $x' \in X'$,

$$\int_\Gamma \langle f, x'\rangle d\nu = \int_\Gamma \langle \widehat{\mu}, x'\rangle d\nu,$$

for every $\nu \in \mathcal{S}(\Gamma)$. Since $\mathcal{S}(\Gamma)$ separates points of $\mathcal{F}$ it follows that $f = \widehat{\mu}$. ■

Remark 3. If we apply Proposition 2.2 to the particular case of Examples 1.1–1.7 of Section 1, with the spaces $\mathcal{F}(X)$ as suggested in Remark 2, then we recover several well known results. Indeed, for the case of Fourier-Stieltjes transforms (Example 1.1) we refer to [6] and for Laplace transforms (Example 1.2) we refer to [12] and [17]. The vector versions of Example 1.3 (Stieltjes transforms) and Example 1.5 (Poisson integrals) can be found in [13] and [14], respectively. The results of [18] are vector-valued analogues of the Hausdorff moment problem (Example 1.4) which can easily be recast into the setting of Proposition 2.2. The vector versions of Examples 1.6 and 1.7 can be found in [8] and [9], respectively.■

3. OPERATOR-VALUED TRANSFORMS

In this section we use the results of Section 2 to characterize operator-valued transforms. Let X be a quasi-complete lcs and $P : \mathcal{B}(\Omega) \to L(X)$ be a regular operator-valued measure. If the space $L(X)$ is sequentially complete or if P is equicontinuous, then bounded measurable functions are P-integrable. Accordingly, the K-transform $\widehat{P} : \Gamma \to L(X)$, of P, can be defined by

$$\widehat{P}(\gamma) = \int_\Omega K(w,\gamma)dP(w), \quad \gamma \in \Gamma.$$

In particular, for each $x \in X$ the X-valued measure $P(\cdot)(x) : E \mapsto P(E)(x)$, $E \in \mathcal{B}(\Omega)$, satisfies

$$\widehat{P}(\gamma)(x) = \int_\Omega K(w,\gamma)dP(w)(x) = (P(\cdot)(x))\widehat{\ }(\gamma), \quad \gamma \in \Gamma.$$

Proposition 3.1. *Let the lcs X be quasi-complete and barrelled. If the kernel $K : \Omega \times \Gamma \to \mathbb{C}$ is bounded, then a map $U : \Gamma \to L(X)$ is the K-transform of an equicontinuous, regular operator-valued measure, if and only if, for each $x \in X$,*
- *(i) $U(\cdot)(x) \in \mathcal{F}(X)$,*
- *(ii) $L_i(U(\cdot)(x))$ is Pettis λ-integrable for each $i \in \mathcal{I}$, and*
- *(iii) the family of maps $\{\Phi_i(U(\cdot)(x));\ i \in \mathcal{I}\}$, as given by (11), is weakly equicompact.*

If such a measure exists, then it is unique.

Proof. If $U = \widehat{P}$ for some equicontinuous, regular, operator-valued measure P, then $U(\cdot)(x) = (P(\cdot)(x))\widehat{\ }$ for each $x \in X$. Accordingly, (i) – (iii) follow from Proposition 2.2.

Conversely, suppose that (i) – (iii) hold. Let $x \in X$. By Proposition 2.2 there exists a (unique) regular measure $m_x : \mathcal{B}(\Omega) \to X$ such that $U(\cdot)(x) = \widehat{m}_x$. For each $E \in \mathcal{B}(\Omega)$, define an operator $P(E) : X \to X$ by

$$P(E)(x) = m_x(E), \quad x \in X. \tag{16}$$

The linearity of $P(E)$ follows from the uniqueness of K-transforms. To complete the proof it suffices to show that $P(E)$ is continuous.

Fix $x \in X$. It follows from (ix) of Section 1 that if $F \in \mathcal{B}(\Gamma)$, then the function $w \mapsto \int_F K(w,\gamma)d\nu(\gamma)$, $w \in \Omega$, is bounded and measurable for each $\nu \in \mathcal{S}(\Gamma)$. Hence, it is m_x-integrable. Fix $\nu \in \mathcal{S}(\Gamma)$. For each $F \in \mathcal{B}(\Gamma)$ define an element $z_F^x(\nu)$ of X by

$$z_F^x(\nu) = \int_\Omega \left(\int_F K(w,\gamma)d\nu(\gamma)\right) dm_x(w). \tag{17}$$

Then it follows from the Fubini theorem and the identity (16), that

$$\langle z_F^x(\nu), x'\rangle = \int_F \langle U(\gamma)(x), x'\rangle d\nu(\gamma),$$

for each $x' \in X'$. This shows that $U(\cdot)(x)$ is Pettis ν-integrable and $\int_F U(\cdot)(x)d\nu = z_F^x(\nu)$, $F \in \mathcal{B}(\Gamma)$. In particular, substituting $F = \Gamma$ and using the identities (16) and (17) gives

$$\int_\Gamma U(\cdot)(x)d\nu = \int_\Omega \tilde{\nu}dP(\cdot)(x), \quad \nu \in \mathcal{S}(\Gamma), \tag{18}$$

for each $x \in X$.

Define linear operators $T_\nu : X \to X$, $\nu \in \mathcal{S}(\Gamma)$, by

$$T_\nu(x) = \int_\Gamma U(\cdot)(x)d\nu, \quad x \in X. \tag{19}$$

Let β be a positive constant such that $|K(w,\gamma)| \le \beta$, $(w,\gamma) \in \Omega \times \Gamma$. Then it follows from the identity $U(\cdot)(x) = \widehat{m}_x$, $x \in X$, that for each $x \in X$, the map $U(\cdot)(x)$ takes its values in the weakly compact set $\beta\overline{bcom}_x(\mathcal{B}(\Omega))$. Since X is barrelled, $\{U(\gamma);\ \gamma \in \Gamma\}$ is an equicontinuous part of $L(X)$.

A typical seminorm on X has the form

$$q_N(x) = \sup\{|\langle x, x'\rangle|;\ x' \in N\}, \quad x \in X,$$

where N is an equicontinuous subset of X'. If $M = \{U(\gamma)'(x');\ \gamma \in \Gamma,\ x' \in N\}$, then M is an equicontinuous subset of X' and it follows easily that

$$q_N(T_\nu(x)) \le |\nu|(\Gamma)q_M(x), \quad x \in X,$$

for each $\nu \in \mathcal{S}(\Gamma)$. This shows that T_ν is a continuous operator for each $\nu \in \mathcal{S}(\Gamma)$. Furthermore, it follows from (16) and (18) that, for each $x \in X$,

$$\{T_\nu(x);\ \nu \in \mathcal{S}(\Gamma),\ \|\widetilde{\nu}\|_\infty \le 1\} \subset \overline{bcom}_x(\mathcal{B}(\Omega)).$$

Since X is barrelled, the family $\{T_\nu;\ \nu \in \mathcal{S}(\Gamma),\ \|\widetilde{\nu}\|_\infty \le 1\}$ is an equicontinuous part of $L(X)$.

Let $E \in \mathcal{B}(\Omega)$. If p is continuous semi-norm on X, then the equicontinuity of $\{T_\nu; \nu \in \mathcal{S}(\Gamma),\ \|\widetilde{\nu}\|_\infty \le 1\}$ implies the existence of continuous semi-norms $q_1, \ldots, q_k$ on X and a constant $\alpha > 0$ such that

$$p(T_\nu(x)) \le \alpha \|\widetilde{\nu}\|_\infty \max_{1 \le j \le k} q_j(x),$$

for each $x \in X$ and $\nu \in \mathcal{S}(\Gamma)$. That is,

$$p\left(\int_\Omega \widetilde{\nu}dm_x\right) \le \alpha \|\widetilde{\nu}\|_\infty \max_{1 \le j \le k} q_j(x), \quad x \in X,\ \nu \in \mathcal{S}(\Gamma).$$

Since $\widetilde{\mathcal{S}}(\Gamma)$ is dense in $C_0(\Omega)$ it follows that

$$p\left(\int_\Omega \psi dm_x\right) \le \alpha \max_{1 \le j \le k} q_j(x), \quad x \in X,\ \psi \in C_0(\Omega),\ \|\psi\|_\infty \le 1,$$

and hence, that

$$p(P(E)(x)) = p(m_x(E)) \le \alpha \max_{1 \le j \le k} q_j(x), \quad x \in X,\ E \in \mathcal{B}(\Omega).$$

This shows that $P(E)$ is continuous for each $E \in \mathcal{B}(\Omega)$ and that $\{P(E);\ E \in \mathcal{B}(\Omega)\}$ is an equicontinuous part of $L(X)$. ■

Remark 4. The kernels of Examples 1.1, 1.2, 1.4 and 1.7 are bounded on $\Omega \times \Gamma$ and hence, Proposition 3.1 is applicable to them. However, Examples 1.3, 1.5 and 1.6 show that there are kernels of interest which are not bounded. An examination of the previous proof shows that the only role played by the boundedness of the kernel was in guaranteeing that each of the linear operators T_ν, $\nu \in \mathcal{S}(\Gamma)$, is continuous. The following result shows that in some cases this conclusion is valid even for unbounded kernels. ■

Proposition 3.2. *Let the lcs X be quasi-complete and barrelled. If, for every $\gamma \in \Gamma$, the function $K(\cdot,\gamma)$ vanishes at infinity and $\mathcal{S}(\Gamma)$ is taken to be the set of measures on Γ with finite support, then a map $U : \Gamma \to L(X)$ is the K-transform of an equicontinuous, regular, operator-valued measure, if and only if, for each $x \in X$, the conditions (i) – (iii) of Proposition 3.1 are valid.*

Proof. As remarked, it suffices to show that each of the operators (19) is continuous. If $\nu = \sum_{j=1}^{m} \alpha_j \epsilon_{\gamma_j}$, where $\alpha_j \in \mathbb{C}$ and $\gamma_j \in \Gamma$ are arbitrary, then it follows from (18) that

$$T_\nu(x) = \int_\Gamma U(\cdot)(x)d\nu = \sum_{j=1}^{m} \alpha_j U(\gamma_j)(x), \quad x \in X.$$

That is, $T_\nu = \sum_{j=1}^{m} \alpha_j U(\gamma_j)$, and hence, T_ν is continuous. ■

4. MULTIPLICATIVE KERNELS

Up to now, Γ has just been a locally compact, σ-compact Hausdorff space. Throughout this section we assume that Γ is also a commutative semigroup with a continuous semigroup operation. If the kernel K also satisfies certain algebraic properties, then the operator-valued measures P constructed in Propositions 3.1 and 3.2 ought to be spectral measures, that is, they should satisfy $P(\Omega) = I$ (the identity operator on X) and $P(E \cap F) = P(E)P(F)$, for every $E, F \in \mathcal{B}(\Omega)$. It is shown that this is indeed the case. It is this fact, exploited in the final section, which makes certain types of integral transforms useful in determining criteria for spectrality of operators. The notation remains as in Section 2.

A map $U : \Gamma \to L(X)$ is said to be a representation of Γ in $L(X)$, if $\langle U(\cdot)(x), x' \rangle$ is measurable, for each $x \in X$ and $x' \in X'$, and if U is a homomorphism, that is,

$$U(\gamma_1\gamma_2) = U(\gamma_1)U(\gamma_2), \quad \gamma_1, \gamma_2 \in \Gamma.$$

If Γ has an identity element e, then it is assumed that $U(e) = I$.

The kernel $K : \Omega \times \Gamma \to \mathbb{C}$ is said to be multiplicative if

$$K(\cdot, \gamma_1\gamma_2) = K(\cdot, \gamma_1)K(\cdot, \gamma_2),$$

for every $\gamma_1 \in \Gamma$ and $\gamma_2 \in \Gamma$. If Γ has an identity element, then it is assumed that $K(\cdot, e)$ is the constant function 1.

If $U : \Gamma \to L(X)$ is the K-transform of a *spectral* measure and if the kernel K is *multiplicative*, then U is necessarily a homomorphism.

Proposition 4.1. *Let X be a quasi-complete, barrelled lcs and K a bounded, multiplicative kernel. Let $P : \mathcal{B}(\Omega) \to L(X)$ be a spectral measure. Then the K-transform of P is an equicontinuous representation of Γ in $L(X)$.*

Proof. Firstly, if Γ has an identity, then

$$\widehat{P}(e) = \int_\Omega K(w, e)dP(w) = \int_\Omega 1dP = I.$$

If $x \in X$, then for each $\gamma \in \Gamma$,

$$(P(\cdot)(x))\widehat{\ }(\gamma) = \int_\Omega K(w,\gamma)dP(w)(x) \in \beta\overline{bco}(P(\cdot)(x))(\mathcal{B}(\Omega)),$$

where β is a bound for K on $\Omega \times \Gamma$. Since X is barrelled and $(P(\cdot)(x))(\mathcal{B}(\Omega))$ is bounded, it follows that $\{\widehat{P}(\gamma);\ \gamma \in \Gamma\}$ is an equicontinuous part of $L(X)$.

To show that $\widehat{P}$ is a homomorphism, it suffices to show that

$$\langle \widehat{P}(\gamma_1\gamma_2)(x), x'\rangle = \langle \widehat{P}(\gamma_1)\widehat{P}(\gamma_2)(x), x'\rangle$$

for each $\gamma_1, \gamma_2 \in \Gamma$, $x \in X$ and $x' \in X'$. This follows easily from the definition of $\widehat{P}$ and the multiplicativity of K and P. ■

Conversely, given a bounded, multiplicative kernel and a representation of Γ in $L(X)$ satisfying the conditions (i) – (iii) of Proposition 3.1, a natural question to ask is whether the operator-valued measure which exists by Proposition 3.1, is a spectral measure. If the space of measures $\mathcal{S}(\Gamma)$ satisfies certain conditions, then this question has a positive answer. Firstly however, some technical lemmas are needed.

Lemma 4.2. *Let X be a quasi-complete, barrelled lcs and $P : \mathcal{B}(\Omega) \to L(X)$ an operator-valued measure. Then the K-transform of P is Pettis ν-integrable for each $\nu \in \mathcal{S}(\Gamma)$.*

Proof. The conditions on X imply that necessarily P is an equicontinuous measure and hence, that bounded measurable functions are P-integrable.

If $\nu \in \mathcal{S}(\Gamma)$, then (ix) of Section 1 implies $w \mapsto \int_E K(w,\gamma)d\nu(\gamma)$, $w \in \Omega$, is bounded and measurable for each $E \in \mathcal{B}(\Gamma)$. Hence, the operator $T_E^\nu = \int_\Omega \int_E K(w,\gamma)d\nu(\gamma)dP(w)$ exists in $L(X)$ and is easily shown to satisfy

$$\langle T_E^\nu(x), x'\rangle = \int_E \langle \widehat{P}(\gamma)(x), x'\rangle d\nu(\gamma),$$

for each $x \in X$ and $x' \in X'$. This shows that $\widehat{P}$ is Pettis ν- integrable and $(\nu\widehat{P})(E) = T_E^\nu$, $E \in \mathcal{B}(\Gamma)$. ■

Lemma 4.3. *Let K be a bounded, multiplicative kernel and $\mathcal{S}(\Gamma)$ be an algebra with respect to a convolution $* : \mathcal{S}(\Gamma) \times \mathcal{S}(\Gamma) \to \mathcal{S}(\Gamma)$ satisfying the identity*

$$\int_\Gamma \psi(\xi)d(\mu * \nu)(\xi) = \int_\Gamma \int_\Gamma \psi(st)d\mu(s)d\nu(t) = \int_\Gamma \int_\Gamma \psi(st)d\nu(t)d\mu(s), \tag{20}$$

for all $\mu, \nu \in \mathcal{S}(\Gamma)$ and every bounded measurable function ψ on Γ. Then

$$(\mu * \nu)\tilde{} = \tilde{\mu}.\tilde{\nu}, \tag{21}$$

for each $\mu \in \mathcal{S}(\Gamma)$ and $\nu \in \mathcal{S}(\Gamma)$.

If X is a quasi-complete, barrelled lcs and $U : \Gamma \to L(X)$ is an equicontinuous representation which is Pettis ν-integrable for each $\nu \in \mathcal{S}(\Gamma)$, then the map

$$T : \nu \mapsto \int_\Gamma U d\nu, \quad \nu \in \mathcal{S}(\Gamma),$$

is a homomorphism, that is,

$$T(\mu * \nu) = T(\mu)T(\nu) = T(\nu)T(\mu), \tag{22}$$

for every $\mu \in \mathcal{S}(\Gamma)$ and $\nu \in \mathcal{S}(\Gamma)$.

Proof. The identity (21) follows from (20) and the Fubini theorem.

Fix $\mu, \nu \in \mathcal{S}(\Gamma)$ and $x \in X$, $x' \in X'$. Since $\langle U(\cdot)(x), x'\rangle$ is bounded and measurable, it follows from (20) and the fact that U is a representation, that

$$\langle T(\mu * \nu)(x), x'\rangle = \int_\Gamma \int_\Gamma \langle U(s)U(t)(x), x'\rangle d\nu(t) d\mu(s) =$$

$$= \int_\Gamma \int_\Gamma \langle U(t)(x), U(s)'(x')\rangle d\nu(t) d\mu(s) = \int_\Gamma \langle T(\nu)(x), U(s)'(x')\rangle d\mu(s) =$$

$$= \int_\Gamma \langle U(s)T(\nu)(x), x'\rangle d\mu(s) = \langle T(\mu)T(\nu)(x), x'\rangle.$$

Since this is for every $x \in X$ and $x' \in X'$, the result follows. ∎

The following result is well known for Banach spaces (Lemma 3 of [5]). Its extension to more general spaces presents no difficulties.

Lemma 4.4. *Let X be a quasi-complete, barrelled lcs and $P : \mathcal{B}(\Omega) \to L(X)$ a regular, operator-valued measure. If there exists a dense subalgebra, D, of $C_0(\Omega)$, such that*

$$P(fg) = P(f)P(g),$$

for every $f \in D$ and $g \in D$, then P is multiplicative.

The converse of Proposition 4.1 can now be stated.

Proposition 4.5. *Let K be a bounded, multiplicative kernel and $\mathcal{S}(\Gamma)$ an algebra with respect to a convolution satisfying (20). Let X be a quasi-complete, barrelled lcs and $U : \Gamma \to L(X)$ a representation satisfying conditions (i) – (iii) of Proposition 3.1. Then there exists a unique equicontinuous, regular, multiplicative measure $P : \mathcal{B}(\Omega) \to L(X)$ such that $U = \widehat{P}$.*

If $P(\Omega) = I$, then P is a spectral measure. In particular, if Γ has an identity, then P is a spectral measure.

Proof. It follows from Proposition 3.1 that there exists a unique equicontinuous, regular measure $P : \mathcal{B}(\Omega) \to L(X)$, such that $U = \widehat{P}$. If $D = \widetilde{\mathcal{S}}(\Gamma)$, then it follows from (x) of Section 1 and Lemma 4.3 that D is a dense subalgebra of $C_0(\Omega)$. By Lemma 4.4, it suffices to show that

$$P(\widetilde{\mu}\widetilde{\nu}) = P(\widetilde{\mu})P(\widetilde{\nu}), \tag{23}$$

for each $\mu, \nu \in \mathcal{S}(\Gamma)$.

Fix $\mu, \nu \in \mathcal{S}(\Gamma)$. It follows from (21), the Fubini theorem and Lemma 4.2 that

$$P(\widetilde{\mu}\widetilde{\nu}) = \int_\Gamma U d(\mu * \nu).$$

But $\int_\Gamma U d(\mu * \nu) = \left(\int_\Gamma U d\mu\right)\left(\int_\Gamma U d\nu\right)$, by (22), and it is easily verified that $\int_\Gamma U d\mu = P(\widetilde{\mu})$ and $\int_\Gamma U d\nu = P(\widetilde{\nu})$. Hence, (23) follows.

If Γ has an identity, then $U(e) = I$ by definiton. However, also

$$U(e) = \int_\Omega K(w, e) dP(w) = \int_\Omega 1 dP = P(\Omega).$$ ∎

Remark 5. It was noted in each of the Examples 1.1, 1.2 and 1.4 that the kernel was bounded. Furthermore, in each of these cases the kernel is multiplicative and the space $\mathcal{S}(\Gamma)$ satisfies the requirements of Lemma 4.3; the operation is convolution of measures in each case. For Examples 1.1 and 1.4 the semi-group Γ has an identity. ∎

5. CRITERIA FOR SPECTRALITY

In this final section various criteria for spectrality of operators in locally convex spaces are given. These criteria follow from certain results of previous sections applied to specific kernels.

Let X be a lcs. An operator $T \in L(X)$ with real spectrum is said to be pseudo-hermitian if there exists a regular spectral measure $P : \mathcal{B}(\mathbb{R}) \to L(X)$ such that

$$T = \int_{\mathbb{R}} w dP(w). \tag{24}$$

In the notation of Example 1.1 for Fourier-Stieltjes transforms, let $\Omega = \mathbb{R}$. Then $\Gamma = \mathbb{R}$, the measure λ is Lebesgue measure and $\mathcal{F}(X)$ is the space of all bounded, weakly continuous, X-valued functions on Γ (see Remark 2). Let $u_k, k \in \mathbb{N}$, be the classical Riesz kernel on Ω and $w_k = \widehat{u}_k$, $k \in \mathbb{N}$, be the corresponding sequence of Fourier transforms. Let the maps $L_k, k \in \mathbb{N}$, be defined as in Example 1.1 and the corresponding maps $\Phi_k, k \in \mathbb{N}$, be given by (11). The following result was proved for reflexive Banach spaces in [3] and for arbitrary Banach spaces in [5]. For a lcs X it can be found in [11; Proposition 4.4]; the proof given there is the "same" as if Proposition 4.5 were applied to the particular setting of Example 1.1 just described.

Proposition 5.1. *Let X be a quasi-complete, barrelled lcs and $T \in L(X)$. Suppose that the representation $U(\gamma) = \exp(i\gamma T)$, $\gamma \in \mathbb{R}$, exists, is weakly continuous and satisfies*

$$\lim_{\gamma \to 0} \frac{1}{i\gamma}(I - \exp(-i\gamma T)) = T, \tag{25}$$

in $L(X)$. Suppose that for each $x \in X$, the map $U(\cdot)(x)$ belongs to $\mathcal{F}(X)$, the maps $L_k(U(\cdot)(x))$, $k \in \mathbb{N}$, are Pettis integrable and the corresponding family of maps (11), $k \in \mathbb{N}$, is weakly equicompact. Then the operator T is pseudo-hermitian.

Conversely, if T is a pseudo-hermitian operator, then the representation $U : \mathbb{R} \to L(X)$ given by $U(\gamma) = \exp(i\gamma T)$, $\gamma \in \mathbb{R}$, exists, is weakly continuous, equicontinuous and (25) is valid.

Remark 6. If X is a Banach space and $T \in L(X)$, then $f(T)$ exists as a member of $L(X)$ for every function f analytic in a neighbourhood of $\sigma(T)$. For locally convex spaces this is no longer the case. Even if f is an entire function, $f(T)$ may not exist. It is for this reason that the existence of $\exp(i\gamma T)$, $\gamma \in \mathbb{R}$, has to be *assumed* in Proposition 5.1. If T has compact spectrum in $\mathbb{C}$, then

$$\sup\left\{\limsup_{n\to\infty}(p(T^n))^{1/n};\ p \text{ is a continuous semi-norm on } L(X)\right\} < \infty,$$

(see [10]). This is sufficient for $\exp(i\gamma T)$, $\gamma \in \mathbb{R}$, to exist, be continuous and satisfy (25). However, there exist many operators with unbounded spectrum which also satisfy these conditions; see Proposition 4.5 of [11]. ∎

An operator $T \in L(X)$ with $\sigma(T) \subset \mathbb{T}$ is said to be pseudo-unitary if there exists a regular spectral measure $P : \mathcal{B}(\mathbb{T}) \to L(X)$ such that

$$T = \int_{\mathbb{T}} w dP(w).$$

The following result can be found in [5] for Banach spaces and in [11; Proposition 4.6] for lcs. It again follows from Proposition 4.5, this time applied to the setting of Example 1.1 with $\Omega = \mathbb{T}$ (the circle group), $\Gamma = \mathbb{Z}$, the measure λ being arc length on $\mathbb{T}$ and $\mathcal{F}(X)$ being the space of all X-valued bounded sequences on $\mathbb{Z}$. We let u_k, $k \in \mathbb{N}$, be the Fejér kernel on $\mathbb{T}$ and $w_k = \widehat{u}_k$, $k \in \mathbb{N}$, be the corresponding sequence of Fourier transforms. The representation $U : \mathbb{Z} \to L(X)$ is given by $U(n) = T^n$, $n \in \mathbb{Z}$, and for $x \in X$, the maps $L_k(U(\cdot)x)$, $k \in \mathbb{N}$, are the Fejér means given by

$$L_k(U(\cdot)x)(w) = k^{-1} \sum_{m=0}^{k-1} \sum_{n=-m}^{m} w^n T^n x, \quad w \in \mathbb{T}.$$

Proposition 5.2. *Let X be a quasi-complete, barelled lcs and $T \in L(X)$ have an inverse in $L(X)$. Suppose that $\{T^n;\ n \in \mathbb{Z}\}$ is an equicontinuous part of $L(X)$, and for each $x \in X$, the set*

$$\left\{ \frac{1}{k} \sum_{m=0}^{k-1} \sum_{n=-m}^{m} \widehat{\psi}(-n) T^n(x);\ k \in \mathbb{N},\ \psi \in C(\mathbb{T});\ \| \psi \|_\infty \leq 1 \right\},$$

is relatively weakly compact. Then there exists a regular spectral measure $P : \mathcal{B}(\mathbb{T}) \to L(X)$ such that

$$T^n = \int_{\mathbb{T}} w^n dP(w), \tag{26}$$

for each $n \in \mathbb{Z}$. In particular, T is pseudo-unitary.

Conversely, if T is a pseudo-unitary operator then (26) holds for every $n \in \mathbb{Z}$.

An operator $T \in L(X)$ with spectrum in $[0, \infty)$ is said to be pseudo-positive if there exists a regular spectral measure $P : \mathcal{B}([0,\infty)) \to L(X)$ such that

$$T = \int_{[0,\infty)} w dP(w).$$

In the notation of Example 1.2 for Laplace transforms, let $\Omega = [0, \infty)$ and $\Gamma = (0, \infty)$. Then $\mathcal{F}(X)$ is the space of all X-valued functions f on Γ, having weak derivatives of all orders and such that $\lim_{t\to\infty} f(t)$ exists weakly in X (see Remark 2). Let L_k, $k \in \mathbb{N}$, be the Widder differential operators as defined in Example 1.2 and let the corresponding maps $\Phi_k, k \in \mathbb{N}$, be given by (11). The following result is in [12]; the proof given there can be deduced directly from Proposition 4.5.

Proposition 5.3. *Let X be a quasi-complete, barrelled lcs and $T \in L(X)$. Then T is pseudo-positive, if and only if, the semi-group $U(\gamma) = \exp(-\gamma T)$, $\gamma > 0$, exists, satisfies the identities*

$$\lim_{\gamma\to 0+} U(\gamma) = I \quad \textit{and} \quad \lim_{\gamma\to 0+} \frac{1}{\gamma}(I - U(\gamma)) = T,$$

in $L(X)$, and for each $x \in X$, the conditions (i) - (iii) of Proposition 3.1 are satisfied.

As a final example we consider an application of the Stieltjes transform (cf. Example 1.3 and Remark 2), where $\Omega = [0,\infty)$, $\Gamma = (0,\infty)$ and $\mathcal{F}(X)$ is the space of all X-valued functions f on Γ, having weak derivatives of all orders and such that $\lim_{t\to 0+} tf(t)$ exists weakly in X. Let L_k, $k \in \mathbb{N}$, be the Widder differential operators as defined in Example 1.3. In this case the kernel K is not uniformly bounded.

Let the operator $T \in L(X)$ have spectrum in $[0,\infty)$. Then the resolvent operator $R(\gamma) = (\gamma + T)^{-1}$, of $(-T)$, is defined for each $\gamma > 0$. Our final result can be found in [4] for reflexive Banach spaces and in [13] for lcs; the proof given there is just an application of our Proposition 4.5 to the special setting of Example 1.3 just described.

Proposition 5.4. *Let X be a quasi-complete, barrelled lcs and $T \in L(X)$ have spectrum in $[0,\infty)$. Then T is pseudo-positive, if and only if, for each $x \in X$,*

(i) the functions $L_k(R(\cdot)(x))$, $k \in \mathbb{N}$, are Pettis integrable and

(ii) the maps $\Phi_k(R(\cdot)(x)) : C_0([0,\infty)) \to X$, $k \in \mathbb{N}$, given by

$$\Phi_k(R(\cdot)(x))(\psi) = \int_{[0,\infty)} \psi L_k(R(\cdot)(x))d\lambda, \quad \psi \in C_0([0,\infty)),$$

are weakly equicompact.

References

[1] Hewitt, E., *A new proof of Plancherel's theorem for locally compact abelian groups,* Acta Sci. Math. (Szeged), **24** (1963), 219-227.

[2] Hille, E. and R. S. Phillips, "Functional Analysis and Semigroups," Amer. Math. Soc. Colloq. Publ., No.XXXI, New York, 1957.

[3] Kantorovitz, S., *On the characterization of spectral operators,* Trans. Amer. Math. Soc., **111** (1964), 152-181.

[4] Kantorovitz, S., *Characterization of unbounded spectral operators with spectrum in a half-line,* Comment. Math. Helvetici, **56** (1981), 163-178.

[5] Kluvánek, I., *Characterization of Fourier-Stieltjes transforms of vector and operator-valued measures,* Czechoslovak Math. J., **17** (92), (1967), 261-276.

[6] Kluvánek, I., *Fourier transforms of vector-valued functions and measures,* Studia Math., **37** (1970), 1-12.

[7] Kluvánek, I. and G. Knowles, "Vector Measures and Control Systems," North Holland, Amsterdam, 1976.

[8] McKee, S. K., *Orthogonal expansion of vector-valued functions and measures,* Mat. Časopis, **22** (1972), 71-80.

[9] McKee, S. K., *Transforms of vector-valued functions and measures,* Mat. Časopis, **23** (1973), 5-13.

[10] Neubauer, G., *Zur Spektraltheorie in lokal konvexen Algebren,* Math. Ann., **142** (1961), 131-164.

[11] Ricker, W. J., *Extended spectral operators,* J. Operator Theory, **9** (1983), 269-296.

[12] Ricker, W. J., *Semigroups of operators and an application to spectral theory,* Math. Proc. Cambridge Phil. Soc., **96** (1984), 143-149.

[13] Ricker, W. J., *Characterization of Stieltjes transforms of vector measures and an application to spectral theory,* Hokkaido Math. J., **13** (1984), 299-309.

[14] Ricker, W. J., *Characterization of Poisson integrals of vector-valued functions and measures on the unit circle,* Hokkaido Math. J., **16** (1987), 29-42.

[15] Rudin, W., "Real and Complex Analysis," McGraw Hill, Ljubljana, 1970.

[16] Simon, A. B., *Cesàro summability on groups: characterization and inversion of Fourier transforms,* Function Algebras, Proc. Internat. Symp., Tulane University, (1965), 208-215.

[17] Whitford, A. K., *Laplace-Stieltjes transforms of vector-valued measures,* Mat. Časopis, **22** (1972), 156-163.

[18] Whitford, A. K., *Moment sequences in locally convex spaces,* Math. Ann. **205** (1973), 317-322.

[19] Widder, D. V., "The Laplace Transform," Princeton University Press, 1946.

[20] Wong, J. S. W., *On a characterization of Fourier transforms,* Monatsh. Math., **70** (1966), 74-80.

School of Mathematics
University of New South Wales
Kensington, 2033
Australia.

Received March 27, 1991
and in final form March 27, 1991

Semigroup Forum Vol. 45 (1992) 364–371

RESEARCH ARTICLE

Limited Semaphore Codes

M. Satyanarayana and Supriya Mohanty

Communicated by Gerard Lallement

An alphabet A is a set of symbols, called letters. A word over A is a finite concatenation of letters. A^* denotes a free monoid over A, which includes an empty word, represented by 1. A subset X of $A^+ = A\backslash 1$ is called a code over A if $x_1x_2\dots x_n = y_1y_2\dots y_m$ with $x_i, y_i \in X$, then $n = m$ and $x_i = y_i$ for every i, $1 \leq i \leq n$. A word (letter) in X is called an X-word (X-letter). A code X is called a uniformly synchronous code if for some non-negative integer d, $uxv \in X^*$ with $x \in X^d$, $u, v \in A^*$ implies $ux, xv \in X^*$, where X^* is a free monoid over X. The smallest d satisfying this property is called the synchronizing delay of X, denoted by $\sigma(X)$. A code X is right complete if, for any $u \in A^*$, there exists a $t \in A^*$ such that $ut \in X^*$. A code X satisfies the $F - d$ condition if $A^+X^dA^+ \cap X = \emptyset$, where d is a natural number. A code X is called a semaphore code iff $A^*XA^+ \cap X = \emptyset$ and X is right complete, equivalently, iff X is a right complete prefix code satisfying the $F - 1$ condition [1;109]. A code X is a synchronous code if there exists an $x \in A^*$ such that $A^*x \subseteq X^*$. x is called a synchronizing word for X. Clearly $x \in X^*$. It is well-known that synchronous codes are right complete prefix codes. A code X is called d-synchronous iff $A^*X^d \subseteq X^*$. Obviously d-synchronous codes are synchronous. Let m, n be any two nonnegative integers such that $m + n \neq 0$. A code X over A is said to be an (m,n)-limited code if for any sequence $\{u_0, u_1, \dots, u_{m+n}\}$ in A^*, $u_0u_1, u_1u_2, \dots u_{m+n-1}u_{m+n} \in X^*$ imply $u_m \in X^*$. A limited code is an (m,n)-limited code for some non-negative integers m, n with $(m+n) \neq 0$. We treat any subset of A to be a $(0,0)$-limited code. In this paper we shall determine when a limited code is a semaphore code. The class of semaphore codes is distinct from the class of synchronous codes in general. We provide some conditions when a semaphore code, not necessarily limited, is a synchronous code and vice versa. In particular we observe that limited semaphore codes are not only synchronous but also d-synchronous. We shall show that 1-synchronous codes are always semaphore codes as well as limited codes and determine the structure of certain 1-synchronous codes, which are proved exactly to be $(1,0)$-limited semaphore codes. Through this work we assume that no code is a subset of A. For all undefined terms here we refer the reader to [1].

Firstly we shall mention another characterization of semaphore codes in general and thus obtain a characterization of limited semaphore codes.

Lemma 1. *A code X is a semaphore code iff X is a right complete prefix code and whenever $uv \in X$ with $v \in A^+$, then there exists $t \in A^*$ such that $vt \in X$.*

Proof. Let X be a semaphore code. Evidently it is a right complete prefix code. Suppose $uv \in X$ with $v \in A^+$. We may assume $u \neq 1$, otherwise $v = v1 \in X$. Since X is right complete, $vs \in X^*$ for some $s \in A^*$. Since $v \neq 1$, we have $vs \neq 1$, and hence $vs \in X^+$. Then $vs = x_1x_2 \dots x_i \dots x_n$ with $x_i \in X$ and $n \geq 1$. We have then $v = x_1$ or $v = x_1r$ or $vr = x_1$, where

$r \neq 1$. If $v = x_1 r$, then $uv = ux_1 r \in A^+XA^+ \cap X$, which contradicts the $F-1$ property of semaphore codes. In the other two cases, clearly we have $vt \in X^*$ for some $t \in X^*$. Conversely, let X be a right complete prefix code satisfying the given condition. Assume $uxv \in X$ with $x \in X$ and $u, v \in A^+$. By assumption, $xvt \in X$ for some $t \in X^*$. This implies $vt = 1$ and so $v = 1$, which is not true. Thus $A^+XA^+ \cap X = \emptyset$, which proves that X is a semaphore code.

Lemma 2. *A limited code is prefix iff it is (m, o)-limited for some natural number m.*

Proof. It is noted in the example 2.2 [1; 330] that any (m, o)-limited code is a prefix code. Conversely, let X be an (m, n)-limited code which is prefix. Let $\{u_0, u_1, \ldots, u_{m+n}\}$ be a sequence in A^* such that $u_0u_1, u_1u_2, \ldots, u_{m+n-1}u_{m+n} \in X^*$. Since X is (m, n)-limited, $u_m \in X^*$. But X^* is right unitary by the prefix property of X [1;46]. Now $u_m, u_mu_{m+1} \in X^*$ imply $u_{m+1} \in X^*$. Continuing in this way we get $u_{m+n} \in X^*$, which proves that X is $(m+n, o)$-limited.

Theorem 3. *A limited code X is a semaphore code iff X is right complete, X is (m, o)-limited for some natural number m, and whenever $u_0u_1 \in X$ with $u_1 \neq 1$, then $u_1u_2, u_2u_3, \ldots, u_{m-1}u_m \in X$ with $u_m \in X \cup \{1\}$ for some $u_2, u_3, \ldots, u_m \in A^*$.*

Proof. Let X be a limited semaphore code. Then, by Lemma 2, X may be assumed to be (m, o)-limited. Suppose $u_0u_1 \in X$ for some $u_1 \in A^+$. Then, by Lemma 1, $u_1u_2 \in X$ for some $u_2 \in A^*$. This process continues if no one of $u_2, u_3, \ldots$, constructable by Lemma 1, is 1. Then we have $u_0u_1, u_1u_2, \ldots, u_{m-1}u_m \in X$, so that $u_m \in X^*$ since X is (m, o)-limited. Suppose $u_m \notin X \cup \{1\}$. Then $u_m = x_1x_2 \ldots x_n$ with $x_i \in X$ and $n > 1$. Then $u_{m-1}u_m = u_{m-1}(x_1)x_2 \ldots x_n \in A^+XA^+ \cap X$, contradicting the $F-1$ property of semaphore codes. Suppose now, that there exist $u_2, u_3 \ldots u_r$ with $r < m$, each being different from 1 and $u_{r+1} = 1$ such that $u_0u_1, u_1u_2, \ldots, u_ru_{r+1} \in X$. Then the set $u_{r+2} = u_r$; $u_{r+3} = 1; \ldots; u_m = 1$ or u_r according as u_{m-1} is u_r or 1. We have then $u_m \in X \cup \{1\}$ since $u_r \in X$.

For the converse, it suffices to prove that X satisfies the $F-1$ condition by virtue of Lemma 2. Let $uxv \in X$ with $u, v \in A^+$ and $x \in X$. From the given condition, we get $xvt \in X$ for some $t \in A^*$. But $x, xvt \in X$ imply $vt = 1$ by the prefix condition and so $v = 1$, which is a contradiction. Thus X satisfies the $F-1$ condition. ∎

The above theorem enables us to formalize a property given in the example 2.3 [1; 330] in the following:

Corollary 4. *Let X be a $(1,0)$-limited code. Then X is a semaphore code iff X is right complete and $uv \in X$ implies $v \in X \cup \{1\}$.*

Semaphore codes and synchronous codes share many common properties such as thin, right complete, prefix, maximal prefix, and maximal [1]. However, they form distinct classes. But we have the following interesting general result. In this context we may mention that maximal prefix codes containing letters, if finite, are synchronous [1; 137].

Theorem 5. *If X is a semaphore code over A containing at least one letter, then X is a synchronous code and every letter in X is a synchronizing word for X.*

Proof. Let $a \in X \cap A$. If a is not a synchronizing word for X, then for some $u \in A^+$, $ua \notin X^+$. Since X is a maximal prefix code, $A^* = X^*P$, where P is the set of all proper left factors of X-words [2; 117]. Thus $ua \in (P-1) \cup X^+(P-1)$. If $ua \in P-1$, then $uav \in X$ for some $v \in A^+$. Consequently $uav \in A^+XA^+ \cap X$, which contradicts the $F-1$ condition of semaphore codes. If $ua \in X^+(P-1)$, then $ua = x_1x_2\ldots x_nt$, where $x_i \in X$, $t \in P-1$ and $n \geq 1$. Then we get either $a = t$ or $ra = t$ for some $r \in A^+$. If $t = a$, then $as \in X$ for some $s \in A^+$. This contradicts that X is a prefix code since $a \in X$. If $ra = t$, then for some $s \in A^+$, $ras \in X \cap A^+XA^+$. This violates the $F-1$ property. Thus $ua \in X^*$, which asserts our result. ∎

Theorem 6. *If X is a 1-synchronous code, then X is a $(2,0)$-limited semaphore code with synchronizing delay 1.*

Proof. Let $uxv \in X^*$ with $v \in A^+$. Since X is 1-synchronous, $ux \in X^*$. Then by the prefix property of X, $v = 1$. Thus the $F-1$ property is satisfied and hence X is a semaphore code. Consider now $uxv \in X^*$ with $x \in X$. As above $ux \in X^*$. Then $ux, uxv \in X^*$ imply $v \in X^*$ by the prefix property of X. Hence $v \in X^*$. Thus $ux, xv \in X^*$, proving $\sigma(X) = 1$. This implies that X is a $(2,2)$-limited code by Proposition 2.5 [1; 332]. We claim that in fact X is $(2,0)$-limited. For, let $u_0u_1, u_1u_2 \in X^*$. By the right complete property of X, we have $u_2u_3, u_3u_4 \in X^*$ for some $u_3, u_4 \in A^*$. Since X is $(2,2)$-limited, $u_2 \in X^*$. Thus X is $(2,0)$-limited. ∎

However, within the family of limited codes we have

Theorem 7. *Limited semaphore codes are uniformly synchronous and synchronous (indeed d-synchronous).*

Proof. By Corollary 11 of [3], the concepts of being limited and uniformly synchronous coincide in the case of semaphore codes. Then by Theorem 3 of [3], limited semaphore codes are d-synchronous and so synchronous. ∎

We cite now an example of a limited synchronous code which is not a semaphore code. This example also shows that the converse of Theorem 5 is not true.

Example 8. Let $X = \{a, b^m a, c^m a, c^m b, b^{n_1}c^{m_1}a, b^{n_1}c^{m_1}b^{n_2}c^{m_2}a, \ldots, b^{n_1}c^{m_1} \ldots b^{n_k}c^{m_k}a\}$ where m and n_1 take all possible positive integral values and n_k $(k \neq 1)$ and m_k take all possible nonnegative integral values.

Clearly X is a prefix code. a is a synchronizing word for X. For, let $w \in A^*$. Then $w = a^{l_1}b^{m_1}c^{n_1} \ldots a^{l_r}b^{m_r}c^{n_r} \ldots a^{l_k}b^{m_k}c^{n_k}$, where l_r, m_r, $n_r \geq 0$ for all $1 \leq r \leq k$. If for all r, $l_r = 0$, then $wa \in X$. Let r be the first-index such that $l_r > 0$. Then $wa = b^{m_1}c^{n_1} \ldots a(a^{l_r-1})b^{m_r}c^{n_r} \ldots c^{n_k}a \in X^*$. Thus a is a synchronizing word. We claim that X is $(2,0)$-limited. For, let $u_0u_1, u_1u_2 \in X^*$. Suppose $u_2 \notin X^*$. If $u_1 \in X^*$, then $u_2 \in X^*$ by the right unitary property of X^* [1; 46]. Furthermore, if $u_2 = uav$ with $u, v \in A^*$, then $u_1u_2 = (u_1ua)v = X^*$ with $u_1ua \in X^*$ since a is a synchronizing word. Hence as above, $v \in X^*$ which implies $uav = (ua)v \in X^*$. So assume $u_1 \notin X^*$ and a is not a factor of u_2. Since $u_0u_1 \in X^*$, u_1 ends with either a or b. But, since a is a synchronizing word, and $u_1 \notin X^*$, u_1 ends with b. Let $u_1 = ub$ with $u \neq 1$. Then, since $u_0u_1 \in X^*$, and $c^m b$ is the only X-word ending with b, $u = tc^m$, where without loss of generality, we may assume that m is the highest power of c with which u ends. Thus $u_1 = ub = tc^m b$. Now either t ends with either a or b. Since a is a synchronizing word, and $u_1 \notin X^*$, t must end with b, and consequently $u_1 = sc^n bc^m b$, where $t = sc^n b$. Repeating this argument finitely many times we get $u_1 = \prod_{m \geq 1} c^m b \in X^*$, which is a contradiction. Hence $u_1 = b$. Now, since $u_1 u_2 \in X^*$, and a is not a factor of u_2, we have $u_2 = vb$ with $v \in A^*$. Therefore, $u_1 u_2 = bvb \in X^*$. Since every X-word beginning with b must end with a, a must be a factor of v, and hence a factor of u_2. This contradiction proves $u_2 \in X^*$. Finally, since $bcba = b(cb)a \in A^+ X A^+ \cap X$, X is not a semaphore code.

For the converse when a limited synchronous code is a semaphore code, we present below some partial results along with some general results. First we mention a result, which is interesting in its own right. This is used to characterize a class of 1-synchronous codes in the next theorem.

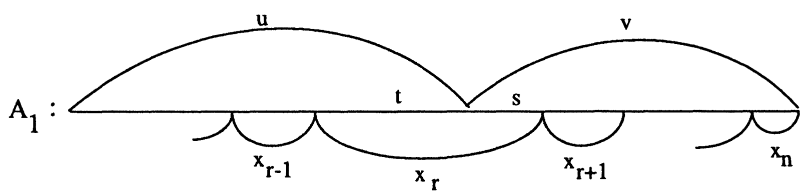
Proposition 9. *For a code X the following are equivalent:*

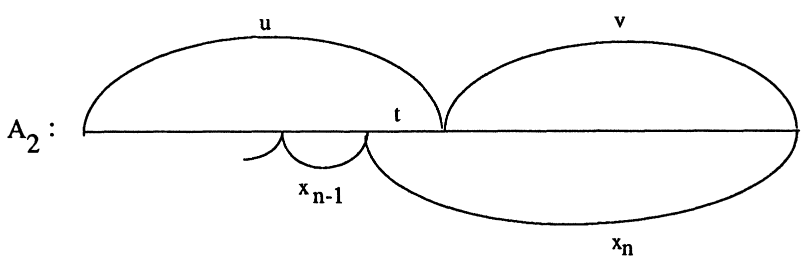
i) *X is a $(1,0)$-limited code with $\sigma(X) = 1$.*

ii) *If $uv \in X$, then $v \in X \cup \{1\}$.*

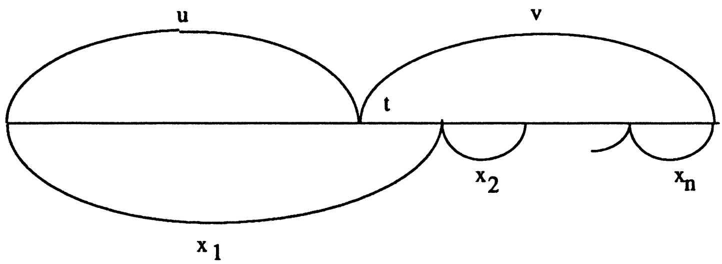
Proof. (i) $\Longrightarrow$ (ii). Let $uv \in X$. Then $v \in X^*$ since X is $(1,0)$-limited. Let $v = x_1 x_2 \ldots x_n$ with $x_i \in X$ and $n > 1$. Then $uv = u(x_1)x_2 \ldots x_n \in X^*$ implies $ux_1 \in X^*$ since $\sigma(X) = 1$. Since $ux_1 \neq 1$, uv is a code word as well as a product of at least two code words, which is absurd. Hence $v = x_1 \in X$.

(ii) $\Longrightarrow$ (i). Let $uv \in X^*$. Suppose $v \notin X^*$. Then $u \notin X^*$ since otherwise, by the prefix property, $v \in X^*$. Set $uv = x_1 x_2 \ldots x_n$ with $x_i \in X$. If $n = 1$, $uv \in X$ and thus by our hypothesis $v \in X \cup \{1\}$. Hence assume $n > 1$. We have now only the following three possibilities depicted as follows:





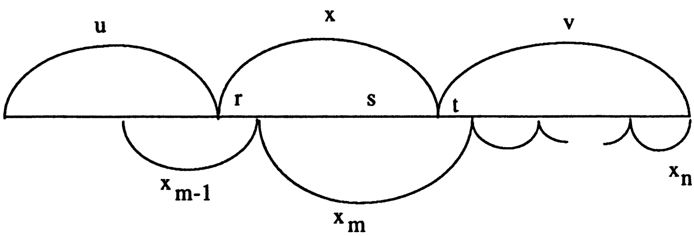
A_3 :



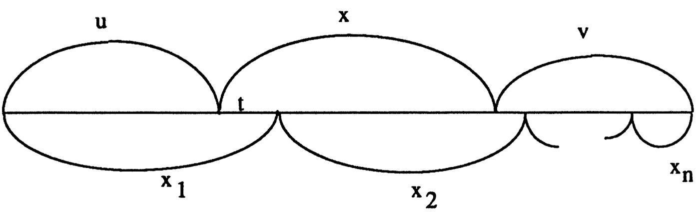
In the case A_1, $u = x_1x_2 \dots x_{r-1}t$ with $x_i \in X$ and $ts = x_r$. Then $s \in X$. Thus $v \in X^*$. In the case A_2, $u = x_1x_2 \dots x_{n-1}t$ with $x_i \in X$. Since $tv = x_n \in X$, $v \in X$. In the last case, $ut = x_1 \in X$, so that $t \in X$. Thus $v \in X^*$. Hence X is $(1,0)$-limited.

Firstly we observe that X satisfies the $F-1$ condition. If $uxv \in X$ with $u, v \in A^+$, $x \in X$, then by (ii), $x, xv \in X \cup \{1\}$, which implies $v = 1$ since X is a prefix code by Lemma 2. We claim now $\sigma(X) = 1$. Let $uxv \in X^*$ with $x \in X$. Since X is $(1,0)$-limited, $xv \in X^*$. It suffices to prove that $ux \in X^*$. Suppose $ux \notin X^*$. Now $uxv = x_1x_2 \dots x_n$ with $x_i \in X$ and $n > 1$ by the $F-1$ condition. Again by the $F-1$ condition, we are left with the following two choices:

A_1 :



A_2 :



In the case A_1, $rs \in X$ and $st \in X$. Then by (ii), $s, t \in X$. Hence the product of two code words is a code word, which is absurd. In the case A_2, $ut \in X$, which implies $t \in X$. But t is a left factor and an X-word. This contradicts the prefix property of X becauise $(1,0)$-limited codes are prefix by Lemma 2. Thus $ux \in X^*$ and $\sigma(X) = 1$.

Recall the following definition: A word u is said to be X-*primitive* with respect to a code X if $u^n \in X$, then $n = 1$.

Theorem 10. *For a code X over A the following are equivalent:*

i) *X is a 1-synchronous code in which every X-word ends with an X-letter.*

ii) *X is a $(1,0)$ limited semaphore code.*

iii) $X = (B')^*B$, *where* B *is a nonempty set of* X*-letters;* $A = B \cup B'$; $B \cap B' = \emptyset$; $B' \neq \emptyset$.

iv) X *is a right complete code satisfying the property:* $uv \in X$ *implies* $v \in X \cup \{1\}$.

v) X *is a semaphore code containing words of every length such that every* X*-word ends with an* X*-letter and every letter is* X*-primitive.*

Proof. (i) $\Longrightarrow$ (iii). $B \neq \emptyset$ by hypothesis. Since $X \not\subseteq A$, $B' \neq \emptyset$. If $a \in B$, then no X-word is of the form uav with $v \neq 1$, since $ua \in X^*$ and $uav \in X$ contradict the prefix property. Since every X-word ends with an X-letter, we must have for every $x \in X$, $x = b_1b_2 \ldots b_m a_1 a_2 \ldots a_r$ with $b_i \in B'$ and $a_i \in B$. But $r = 1$ since otherwise $x \in A^+XA^+$, contradicting the semaphore property of X mentioned in Theorem 6. Thus $X \subseteq (B')^*B$. Let $x = b_1b_2 \ldots b_m a$ with $b_i \in B'$ and $a \in B$. If $x \notin X$, then by the 1-synchronous property, $b_1b_2 \ldots b_m a = x_1x_2 \ldots x_r$ with $x_i \in X$, $r > 1$. Since every X-word ends with an X-letter $x_1 \neq b_1$ or $b_1b_2, \ldots$ or $b_1b_2 \ldots b_m$. Thus $x = b_1b_2 \ldots b_m a \in X$. Hence $X = (B')^*B$.

(ii) $\Longrightarrow$ (iii). We observe firstly that $a \in B$ iff a occurs as an end letter of an X-word. If $a \in B$, then $1a \in X$, so that a occurs as an end letter of an X-word. If $a \in A$, and if $ua \in X$, then $a \in X^*$, since X is $(1,0)$-limited and so $a \in X$. Thus $a \in B$. Consider any X-word. Then it is of the form ua with $a \in B$. A limited code X is circular by Proposition 2.2 [1; 330] and in a circular code X every word is X-primitive [1; 323]. Also X satisfies the $F-1$ condition. Hence no $b \in B$ occurs in u if the length of u is greater than 1. If $u \in B$, then we have that an X-word is a product of two X-words, which is absurd. Thus $u \in (B')^*$ and so $ua \in (B')^*B$. We claim now that ua with $u \in (B')^*$ and $a \in B$ is in X, which shows $X = (B')^*B$. Suppose $ua \notin X$. By the right complete condition, $uat \in X^*$. If $uat \in X$, then $t \neq 1$ by our assumption. Then $uat \in A^+XA^+ \cap X$, which is impossible. Thus we have $uat = x_1x_2 \ldots x_n$ with $x_i \in X$ and $n \geq 2$. Set $u = b_1^{n_1} \ldots b_r^{n_r}$ where $b_i \in B'$, $n_1 > 0$ and $n_j \geq 0$ for $j \geq 2$. If $b_1^{n_1} = x_1, s$, then x_1 is a power of b_1, which implies $b_1 = x_1 \in X$, since every word is X-primitive. But this is not true since $X \cap B' = \emptyset$. Hence $b_1^{n_1}$ is a left factor of x_1. Since $x_1 \neq b_1$ and since every X-word ends with a letter from B, we have $x_1 = ua$ or $x_1 = uat$. Because of the $F-1$ condition, the latter is not possible. Thus $ua \in X$, proving our assertion.

(iii) $\Longrightarrow$ (i) and (ii). These are straightforward verifications.

(iv) $\Longrightarrow$ (ii). By Proposition 9, X is $(1,0)$-limited with $\sigma(X) = 1$. Then by Proposition 7 of [3], X satisfies the $F-1$ condition. Thus X is a semaphore code.

(v) $\Longrightarrow$ (iii). Firstly we note that if $uat \in X$ with $a \in B$ then $t = 1$ by the $F-1$ condition. If x is an X-word of length 2, then $x \in B'B$ since by the prefix property, x begins with a letter from B' and x ends with an X-letter by the hypothesis. If $b \in B'$, $a \in B$ and if $ba \notin X$, then by the right complete condition, $bat \in X^+$ for some $t \in A^+$. From the above we have $bat = x_1x_2 \ldots x_n$ with $x_i \in X$ and $n \geq 2$. Since $b, ba \notin X$, $bas = x_1$ with $s \in A^+$. This again contradicts the $F-1$ condition. Thus $B'B$ is the set of all X-words of length 2. In general, consider an X-word x of length $(n+1)$. By hypothesis, $x = ua$ with $a \in B$. No letter in B occurs in u because of the prefix and $F-1$ conditions. Therefore $u \in (B')^n$. Thus $x \in (B')^nB$. Let $u \in (B')^n$ and $a \in B$. Suppose $ua \notin X$. By the right completion condition $uat \in X^*$ with $t \in A^*$. Since $ua \notin X$, as above $uat \notin X$. Therefore $uat = x_1x_2 \ldots x_n$

with $x_i \in X$ and $n \geq 2$. Set $u = b_1^{r_1} \ldots b_m^{r_m}$ with $b_i \in B'$ and $r_1 > 1$. Then $b_1^{r_1} = x_1 t$ implies that x_1 is a power of b_1. Since X is limited, we have $x_1 = b_1 \in B$, which is not true. Hence $x_1 = b_1^{r_1} t$. Since X-words end with letters from B, this means $x_1 = uas$. As noted above we have then $s = 1$. Thus $ua = x_1 \in X$. This proves (iii).

If X satisfies any one of the equivalent conditions (i), (ii) and (iii), then it is easy to verify the conditions (iv) and (v). In fact $\sigma(X) = 1$ by Theorem 10 of [3]. Combining this with Proposition 9, the above conditions are evident.

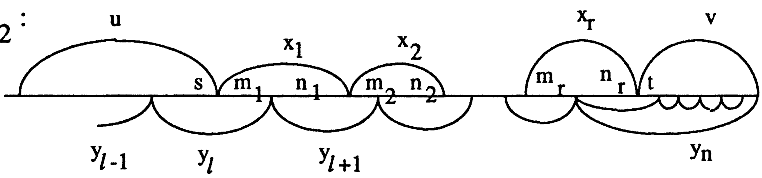
In Theorem 10 [3], we have proved that if an (m, o)-limited code X satisfies the $F-1$ condition, then $\sigma(X) \leq m$. Now we improve the upper bound of $\sigma(X)$ in the following:

Theorem 11. *Let X be a code satisfying the $F-1$ condition. Let r be any positive integer. If X is $(2r, 0)$-limited, then $\sigma(X) \leq r$, and if X is $(2r+1, 0)$-limited, $\sigma(X) \leq r+1$. Also, if X is $(1,0)$-limited, then $\sigma(X) = 1$.*

Proof. Let X be $(2r, 0)$-limited. Suppose $uxv \in X^*$ with $x \in X^r$. We claim $ux, xv \in X^*$. This is trivially true if u or v is 1. So assume $u, v \in A^+$. Since $uxv \in X$ contradicts the $F-1$ condition, we have then $uxv = y_1 y_2 \ldots y_n$ with $y_i \in X$ and $n > 1$, and $x = x_1 x_2 \ldots x_r$ with $x_i \in X$. If u or $ux_1 \ldots ux_i$ is in X^*, then $v \in X^*$ by the right unitary property of X^* and so $xv \in X^*$. So assume $u \notin X^*$ and $ux_1 \ldots ux_i \notin X^*$ for every i. Because of the $F-1$ condition, we have only the following choices represented pictorially as below, since u can either have y_1 as a left factor or be a left factor of y_1.

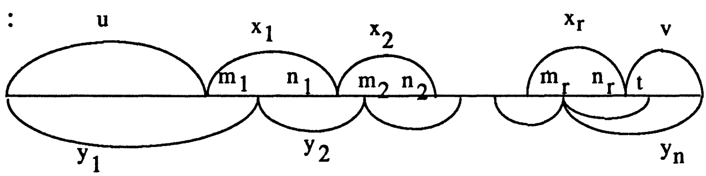
A_1 :

A_2 :



B_1 :

B_2 :



In the A_1- or the A_2-case, $sm_1, m_1n_1, \ldots, m_rn_r \in X$ so that $n_r \in X^*$ since X is $(2r, 0)$-limited. Then $n_r t$ or $n_r v \in X$ which implies $t = 1$ or $v = 1$. This is a contradiction. The B-cases can be treated in the same way, which proves our assertion. If X is $(2r+1, 0)$-limited, then X is $(2r+2, 0)$-limited by the remark in [1; 329] and hence in this case $\sigma(X) \leq r+1$. By observing that semaphore codes satisfy the $F-1$ condition and by combining Corollary 4 and Proposition 9, we have $\sigma(X) = 1$ if X is $(1, 0)$-limited.

Noting that semaphore codes satisfy the $F-1$ condition, we have

Corollary 12. *Let X be a semaphore code. Let r be any positive integer. If X is $(2r,0)$-limited, then $\sigma(X) \leq r$, and if X is $(2r+1,0)$-limited, then $\sigma(X) \leq r+1$. Also if X is $(1,0)$-limited or $(2,0)$-limited, then $\sigma(X) = 1$.*

References

[1] Berstel, J. and D. Perrin, *Theory of Codes*, Academic Press, Inc., 1985.

[2] Lallement, G., *Semigroups and Combinatorial Applications*, Wiley Publications, New York, (1979).

[3] Satyanarayana, M. and S. Mohanty, *Uniformly Synchronous Limited codes*, Semigroup Forum (to appear).

376 Chad Drive
Milpitas, California 95035
U.S.A.

Received April 6, 1991
and in final form September 20, 1991

Semigroup Forum Vol. 45 (1992) 372–384

RESEARCH ARTICLE

Semigroups and Resolvents of Bounded Variation, Imaginary Powers and H^∞ Functional Calculus

K. Boyadzhiev and R. deLaubenfels

Communicated by J. A. Goldstein

Abstract

Let $-A$ be a linear, injective operator, on a Banach space X. We show that $\exists$ an H^∞ functional calculus for A if and only if $-A$ generates a bounded strongly continuous holomorphic semigroup of uniform weak bounded variation, if and only if $A(z+A)^{-1}$ is of uniform weak bounded variation. This provides a sufficient condition for the imaginary powers of A, $\{A^{-is}\}_{s\in\mathbf{R}}$, to extend to a strongly continuous group of bounded operators; we also give similar necessary conditions.

I. Introduction

On a Hilbert space, the famous Spectral Theorem asserts an equivalence between an operator, A, having a $C_0(\mathbf{R})$ functional calculus, and iA generating a bounded, strongly continuous group. In this paper, we show that, on an arbitrary Banach space, generating a strongly continuous holomorphic semigroup of bounded variation is equivalent to having an H^∞ functional calculus (Section III). Being of bounded variation is analogous to being continuous, in the sense that it describes "good" behaviour of a function, but, unlike continuity, very little is known about semigroups of operators of bounded variation.

We also show that having an H^∞ functional calculus is equivalent to $A(z+A)^{-1}$ being of bounded variation in a sector (Section V).

Fractional powers of closed operators with certain growth assumptions on the resolvent constitute a wide branch in Analysis (see, for example, [10], [18], [29]). For operators of type θ (see below) one can define $\{A^{-z}\}_{Re(z)>0}$ as a holomorphic semigroup of possibly unbounded operators (see [14]) in the open right half-plane. Of particular interest is the behaviour of this holomorphic function near the imaginary axis $i\mathbf{R}$: sometimes it is possible to extend it continuously on $i\mathbf{R}$, so that the imaginary powers of A, $\{A^{-is}\}_{s\in\mathbf{R}}$ will be a strongly continuous group of bounded operators [13, 17.9.1]. (Its generator will be $\log A$.) For instance, the imaginary powers J^{-is} of the integration operator J on $L^2[0,1]$ constitute such a group (see [13, 23.16], [11], [16], [17]). These classical results and some related ones (for instance [15], [19]) have been complemented by new developments in the abstract setting, with many surprising applications to the Cauchy problem (see [5], [6], [26]).

1980 Mathematics Subject Classifications: 47A10, 47A60, 47D05, 47D10

The recent papers [5], [6], [26] are concerned primarily with the consequences of $\{A^{-is}\}_{s\in\mathbf{R}}$ being a strongly continuous group of bounded operators. In this paper, we give some simple necessary conditions and some sufficient conditions for this to occur, along with a characterization of a stronger condition, having an H^∞ functional calculus. Many of these results are in the spirit of recent work by McIntosh and Yagi, on a Hilbert space [22], [23], [31], [32] (see also [8], [9] and [28] for recent work on H^∞ functional calculi). The existence of the imaginary powers A^{-is} is related to the theory of the analytic generator – (see [3], [21]), since A and A^{-1} are analytic generators of $\{A^{is}\}_{s\in\mathbf{R}}$ and $\{A^{-is}\}_{s\in\mathbf{R}}$, respectively. Thus we have conditions determining when a given operator is an analytic generator.

Let us summarize our results informally as follows, by considering the following conditions, where we write $\{e^{tB}\}_{t\geq 0}$ to mean a strongly continuous semigroup generated by B.

(1) A has an H^∞ functional calculus.

(2) $\{A^{-is}\}_{s\in\mathbf{R}}$ is a strongly continuous group of bounded operators.

(3_q) $\{e^{-tA}\}_{t\geq 0}$ is holomorphic and $\exists\, \psi > 0$ such that

$$\int_0^\infty \left|tx^*\left(Ae^{-te^{i\phi}A}x\right)\right|^q \frac{dt}{t} < M_q\|x^*\|^q\,\|x\|^q\,,$$

$\forall\, x^*\in X^*,\ x\in X,\ |\phi|<\psi$.

(4_q) $\{e^{-tA}\}_{t\geq 0}$ is holomorphic and $\exists\, \psi > 0$ such that

$$\int_0^\infty \left\|tAe^{-te^{i\phi}A}x\right\|^q \frac{dt}{t} < M_q\|x\|^q\,,$$

$\forall\, x\in X,\ |\phi|<\psi$.

(5_q) $\{e^{-tA}\}_{t\geq 0}$ is holomorphic and

$$\int_0^\infty \|tAe^{-tA}\|^q \frac{dt}{t} < \infty\,,$$

(6) $\exists\, \psi > 0$ such that

$$\int_0^\infty \left|x^*\left(A(te^{i\phi}+A)^{-2}x\right)\right| dt < M\|x^*\|\,\|x\|\,,$$

$\forall\, x\in X,\ x^*\in X^*,\ |\phi|<\psi$.

In [22] and [23] it is shown that, on a Hilbert space, (1), (2) and (4_2) are equivalent. In this paper, on a general Banach space, we show that (1) is equivalent to (3_1), and is equivalent to (6). We also show that (2) implies (3_q), $\forall\, q\geq 2$, and that (4_2) implies (2). Finally, we show that (5_2) is equivalent to A being bounded with bounded inverse.

Note that (3_1) is saying that $t\mapsto e^{-tA}$ is (uniformly weakly) a semigroup of bounded variation (see Definition 2.5) and (6) is saying that $t\mapsto A(t+A)^{-1}$ is of bounded variation.

Our techniques are completely different than those in [9], [22] and [31].

II. Hypotheses, Definitions and Terminology

All operators are linear, on a Banach space X. We will write $D(A)$ for the domain of the (possibly unbounded) operator A, $Im(A)$ for the image, $sp(A)$ for the spectrum, $\rho(A)$ for the resolvent set. We will write $L(X)$ for the space of all bounded operators from X into itself.

We will assume throughout this paper that A is injective and has dense range.

Definition 2.1. We will write S_θ for the sector $\{re^{i\phi} \mid r > 0, |\phi| < \theta\}$, and H_ε for the strip $\{z \mid |Im(z)| < \varepsilon\}$.

Definition 2.2. Suppose $0 \leq \theta < \pi$. We will say that the operator A is of *type θ* if $\forall\ \phi > \theta$, $sp(A) \subseteq \overline{S_\phi}$ and $\{\|w(w-A)^{-1}\| \mid w \notin \overline{S_\phi}\}$ is bounded.

Definition 2.3. For any complex z, there are numerous (equivalent) definitions of the *fractional power*, A^z. We will use the following, $\forall\ x \in D(A) \cap Im(A)$, when A is of type less than π.

(a). When $Re(z) \in (0,1)$, define

$$A^z x = \frac{\sin(z\pi)}{z\pi} \int_0^\infty t^z (A+t)^{-2} Ax\, dt .$$

This formula follows from an integration by parts (see [18] for the initial representation of A^z):

$$\begin{aligned} A^z x &= \frac{\sin(z\pi)}{\pi} \int_0^\infty t^{z-1}(A+t)^{-1} Ax\, dt \\ &= \frac{\sin(z\pi)}{z\pi} \left[t^z (A+t)^{-1} Ax \Big|_0^\infty + \int_0^\infty t^z A(A+t)^{-2} x\, dt \right] \\ &= \frac{\sin(z\pi)}{z\pi} \int_0^\infty t^z A(A+t)^{-2} x\, dt . \end{aligned}$$

When $-A$ generates a bounded strongly continuous holomorphic semigroup, a similar integration by parts, using the fractional powers representation in [18], Proposition 11.4, page 325, gives us, again for $Re(z) \in (0,1)$,

(b).

$$\Gamma(1-z)(A^z x) = \int_0^\infty t^{-z}(Ae^{-tA}x)\, dt .$$

For $|Re(z)| < 1$, $z \neq 0$, we also have, for A of type less than π, (see [18] or [26]),

(c).

$$\begin{aligned} A^z x = \frac{\sin(z\pi)}{\pi} \Bigg[z^{-1}x - (1+z)^{-1}A^{-1}x &+ \int_0^1 t^{z+1}(t+A)^{-1}A^{-1}x\, dt \\ &+ \int_1^\infty t^{z-1}(t+A)^{-1}Ax\, dt \Bigg] . \end{aligned}$$

Of course, when $z = 0$, $A^z \equiv I$, the identity operator.

Definition 2.4. Suppose Ω is an open subset of the complex plane whose closure is not the entire plane. We will say that A has an *$H^\infty(\Omega)$-functional calculus* if $sp(A) \subseteq \overline{\Omega}$ and $\exists$ a continuous algebra homomorphism, $f \mapsto f(A)$, from $H^\infty(\Omega)$ into $L(X)$, such that $(\lambda - A)^{-1} = f_\lambda(A)$, $\forall\ \lambda \notin \overline{\Omega}$, where $f_\lambda(z) \equiv (\lambda - z)^{-1}$.

Definition 2.5. We will say that the holomorphic operator-valued function $F : S_\theta \to L(X)$ is *uniformly weakly of bounded variation* if $\forall\ \psi < \theta$, $\exists\ M_\psi < \infty$ such that

$$\int_0^\infty \left|x^*(F'(te^{i\phi})x)\right|\, dt \leq M_\psi \|x^*\|\, \|x\|\ ,$$

$\forall\ x^* \in X^*,\ x \in X,\ |\phi| < \psi$.

III. Semigroups and H^∞ Functional Calculus

Theorem 3.1. *Suppose $0 < \theta \leq \frac{\pi}{2}$. Then the following are equivalent.*

(a) *A has an $H^\infty(S_\phi)$-functional calculus, $\forall\ \phi, \frac{\pi}{2} > \phi > \frac{\pi}{2} - \theta$.*

(b) *$-A$ generates a bounded strongly continuous holomorphic semigroup of angle θ, that is uniformly weakly of bounded variation.*

Remark 3.2. Generating a bounded strongly continuous holomorphic semigroup is equivalent to $\{tAe^{-tA}\}_{t\geq 0}$ being in $L^\infty\left([0,\infty), \frac{dt}{t}\right)$, while being of weak bounded variation is equivalent to $\{tAe^{-tA}\}_{t\geq 0}$ being weakly in $L^1\left([0,\infty), \frac{dt}{t}\right)$.

Having an H^∞ functional calculus implies the existence of bounded imaginary powers. In Section IV, we will show that having bounded imaginary powers implies that $\{tAe^{-tA}\}_{t\geq 0}$ is weakly in $L^q\left([0,\infty), \frac{dt}{t}\right), \forall\ q \geq 2$ (Theorem 4.1).

Thus, Theorems 3.1 and 4.1 demonstrate how fine the distinctions between generating a bounded strongly continuous holomorphic semigroup, having bounded imaginary powers and having an H^∞ functional calculus are.

Example 3.3. Let RHP be defined to be the open right half-plane $Re(z) > 0$ in the complex plane, $X \equiv \{f \in C_0(i\mathbf{R}) \cap C(\overline{RHP}) \mid f$ is holomorphic on RHP, $\lim_{|z|\to\infty, z\in RHP} f(z) = 0\}$, $-A \equiv \frac{d}{dz}$, the generator of

$$e^{-zA}f(w) \equiv f(z+w)\ .$$

The semigroup $\{e^{-zA}\}_{z\in S_{\frac{\pi}{2}}}$ is stable, that is, $\lim_{t\to\infty} \|e^{-tA}f\| = 0$, $\forall\ f \in X$. This implies that A has dense range.

Suppose $x^* \in X^*$. Then $\exists$ a signed measure μ_{x^*} such that

$$x^*(f) = \int_{\mathbf{R}} f(is)d\mu_{x^*}(s),\ \forall\ f \in X\ .$$

Thus for any $g \in X$, $z \in RHP$,

$$|x^*(Ae^{-zA}g)| = \left|\int_{\mathbf{R}} g'(is+z)\, d\mu_{x^*}(s)\right| = |f'(z)|\ ,$$

if

$$f(z) \equiv \int_{\mathbf{R}} g'(is+z)\, d\mu_{x^*}(s) .$$

This function f is in X and $\|f\|_\infty \le \|g\|_\infty \|x^*\|$.

Thus Theorem 3.1 is saying that A has an $H^\infty(S_\phi)$ functional calculus, $\forall\ \phi > 0$, if and only if functions in the unit ball of X are uniformly of bounded variation on sectors S_ψ, $0 < \psi < \frac{\pi}{2}$, that is, $\exists\ M_\psi < \infty$ such that

$$\int_{\mathbf{R}} \left| f'(te^{i\delta}) \right| dt \le M_\psi \|f\|_\infty , \tag{$*$}$$

for $f \in X$, $|\delta| < \psi$.

It is an open question whether $(*)$ is true.

Lemma 3.4. *Suppose $m > 0$, iB generates a strongly continuous group of bounded operators $\{e^{isB}\}_{s\in\mathbf{R}}$ and, for $j = 1, 2$, $\exists\ g_j : \mathbf{R} \to \mathbf{C}$ such that g_j is continuously differentiable, with $g_j'(s) = O(e^{m|s|})$, support of g_1 contained in $[0,\infty)$, support of g_2 contained in $(-\infty, 1]$ and $\forall\ x^* \in X^*$, $x \in X$, $\exists$ signed measures $E_{j,x^*,x}$, on $\mathbf{R}$, of bounded variation, such that the map $(x^*, x) \mapsto E_{i,x^*,x}$ is bilinear, for $i = 1, 2$ and*

$$x^*\left(e^{isB}x\right) = g_1(s)\left(\int_{\mathbf{R}} e^{isr} dE_{1,x^*,x}(r)\right) + g_2(s)\left(\int_{\mathbf{R}} e^{isr} dE_{2,x^*,x}(r)\right) . \tag{3.5}$$

Then, $\forall\ k > m$, $\exists$ an $H^\infty(H_k)$ functional calculus for B.

Proof. We will write f_z for the translate of f, $f_z(w) \equiv f(z+w)$. Fix $k > m$, and let $\mathcal{A}_k \equiv \{f \in H^\infty(H_k) \mid \exists\ M$ such that $|zf(z)| \le M$, $\forall\ z \in H_k\}$. For any $f \in \mathcal{A}_k$, by [27, Theorem IX.13, p. 18],

$$\sup_{\|x\|\le 1} \int_{\mathbf{R}} \|e^{isB}x\|\, |\mathcal{F}f(s)|\, ds < \infty .$$

Then it may be shown (see Davies [4], for this construction when $\{e^{isB}\}_{s\in\mathbf{R}}$ is bounded) that

$$f(B) \equiv \int_{\mathbf{R}} e^{isB}(\mathcal{F}f)(s)\, ds \tag{$*$}$$

defines an algebra homomorphism from $\mathcal{A}_k$ into $L(X)$ and, if $h_{r,n}(s) \equiv (r-s)^{-n}$, then, whenever $|Im(r)| > k$, $n \in N$,

$$h_{r,n}(B) = (r-B)^{-n} \tag{$**$}$$

for all such r, n.

Fix $x^* \in X^*$, $x \in X$. We will show that $\exists\ M_{x^*,x} < \infty$ such that

$$|x^*(f(B)x)| \le M_{x^*,x} \|f\|_\infty , \tag{$***$}$$

$\forall\ f \in \mathcal{A}_k$, where the supremum of f is taken over H_k.

Using (3.5) and $(*)$, we have, for $f \in \mathcal{A}_k$,

$$x^*(f(B)x) = \int_{\mathbf{R}} \left[\int_{\mathbf{R}} e^{isr} g_1(s) \mathcal{F} f(s)\, ds \right] dE_{1,x^*,x}(r) + \int_{\mathbf{R}} \left[\int_{\mathbf{R}} e^{isr} g_2(s) \mathcal{F} f(s)\, ds \right] dE_{2,x^*,x}(r) \,. \qquad (* * **)$$

Again writing f_z for the translate f, $f_z(w) \equiv f(z+w)$, and choosing a between m and k, this may be rewritten as

$$x^*(f(B)x) = \int_{\mathbf{R}} \left[f_{-ai}(s) * \mathcal{F}^{-1}(g_1(s)e^{-as}) \right](r)\, dE_{1,x^*,x}(r) + \int_{\mathbf{R}} \left[f_{ai}(s) * \mathcal{F}^{-1}(g_2(s)e^{as}) \right](r)\, dE_{2,x^*,x}(r) \,, \qquad (* * **)$$

for each $f \in \mathcal{A}_k$.

Thus, by the properties of convolution operators, to prove $(* * *)$ it is sufficient to have $\mathcal{F}^{-1}(g_1(s)e^{-as})$ and $\mathcal{F}^{-1}(g_2(s)e^{as})$ in L^1. This follows from the fact that $g_1(S)e^{-as}$ and $g_2(s)e^{as}$ are continuously differentiable and in $L^2(\mathbf{R})$, with derivatives in $L^2(\mathbf{R})$.

This proves $(* * *)$. By the Uniform Boundedness Theorem, applied twice, $\{\|f(B)\| \mid f \in \mathcal{A}_k, \|f\|_\infty \leq 1\}$ is bounded. Thus the map $f \mapsto f(B)$ is a bounded algebra homomorphism from $\mathcal{A}_k$, as a subspace of $H^\infty(H_k)$, into $L(X)$.

We may now use $(* * **)$ to define a bounded linear map, $f \mapsto f(B)$, from $H^\infty(H_k)$ into $L(X, X^{**})$, extending the map from $\mathcal{A}_k$ into $L(X)$. We must show that this extension is an algebra homomorphism from $H^\infty(H_k)$ into $L(X)$.

For any $f \in H^\infty(H_k)$, $r > k$, define

$$f_r(w) \equiv r f(w)(r + iw)^{-1} \,.$$

For the remainder of the proof, fix $f \in H^\infty(H_k)$. Suppose $g \in \mathcal{A}_k$. Then, for $x^* \in X^*$, $x \in X$, by dominated convergence,

$$\begin{aligned}
[(fg)(B)x](x^*) &= \lim_{r \to \infty} x^*[(f_r g)(B)x] \\
&= \lim_{r \to \infty} x^*[f_r(B)g(B)x] \\
&= [f(B)g(B)x](x^*) \,.
\end{aligned}$$

In particular, choosing $g(w) \equiv ((k+1) + iw)^{-1}$, this tells us that

$$f(B)((k+1) + iB)^{-1} = \frac{1}{k+1} f_{k+1}(B) \,,$$

so that, since $f_{k+1} \in \mathcal{A}_k$, $f(B) : D(B) \to X$; since $D(B)$ is dense in X and $f(B) \in L(X, X^{**})$, this implies that $f(B) \in L(X)$.

Now let $g \in H^\infty(H_k)$ be arbitrary. Using dominated convergence again,

$$\begin{aligned}
x^*[(fg)(B)x] &= \lim_{r \to \infty} x^*[(fg_r)(B)x] \\
&= \lim_{r \to \infty} x^*[f(B)g_r(B)x] \\
&= \lim_{r \to \infty} (f(B)^* x^*)(g_r(B)x) \\
&= (f(B)^* x^*)(g(B)x) = x^*[f(B)g(B)x] \,,
\end{aligned}$$

so that the map $f \mapsto f(B)$ is an algebra homomorphism from $H^\infty(H_k)$ into $L(X)$. Combined with $(**)$, this concludes the proof. ∎

Remark 3.6. If $g_j(s) \equiv 1$, a similar (but simpler) argument shows that B is a scalar-type spectral operator, when X is reflexive. Lemma 3.4 appears to shed some light on the relationship between scalar operators and operators with an H^∞ functional calculus.

Proof of Theorem 3.1. (b) $\to$ (a). Definition 2.3(c) implies that $\lim_{r\to 0+} A^{r+is}x = A^{is}x$, $\forall\, x \in D(A)\cap Im(A)$. Thus Definition 2.3(b) and hypothesis (b) imply that

$$x^*(A^{-is}x) = \frac{1}{\Gamma(1+is)}\int_0^\infty t^{is}x^*(Ae^{-tA}x)\,dt\ , \tag{$*$}$$

$\forall\, x^* \in X^*$, $x \in D(A)\cap Im(A)$.

By the bounded variation of $\{e^{-tA}\}_{t\geq 0}$, $\{\|A^{-is}x\| \mid x \in D(A)\cap Im(A),$ $\|x\| \leq 1\}$ is bounded, $\forall\, s \in \mathbf{R}$. Since $D(A)\cap Im(A)$ is dense, this implies that A^{-is} extends to a bounded operator, such that $(*)$ holds.

Weak, hence strong, continuity of $\{A^{-is}\}_{s\in\mathbf{R}}$ follows from $(*)$ and dominated convergence.

Let $B \equiv -\log(A)$, that is, iB is the generator of the strongly continuous group $\{A^{-is}\}_{s\in\mathbf{R}}$. Fix $\phi > \frac{\pi}{2}-\theta$. It is sufficient to show that B has an $H^\infty(H_\phi)$ functional calculus, since we may then define an $H^\infty(S_\phi)$ functional calculus for A by $f(A) \equiv (f\circ g)(B)$, where $g(z) \equiv e^{-z}$.

Choose ψ such that $\theta > \psi > \frac{\pi}{2} - \phi$ and fix $x \in X$, $x^* \in X^*$. We have shown that iB generates a strongly continuous group, given by

$$x^*(e^{isB}x) = \frac{1}{\Gamma(1+is)}\int_0^\infty t^{is}x^*(Ae^{-tA}x)\,dt\ .$$

We construct g_j and $E_{j,x^*,x}$, of Lemma 3.4, as follows. First, we claim that there exist $E_{j,x^*,x}$ as in Lemma 3.4, f_j such that $f_1(s)$ and $f_1'(s)$ are $O(e^{-\psi s}e^{|s|\frac{\pi}{2}})$, $f_2(s)$ and $f_2'(s)$ are $O(e^{\psi s}e^{|s|\frac{\pi}{2}})$, and

$$x^*(e^{isB}x) = f_j(s)\int_{\mathbf{R}} e^{isr}dE_{j,x^*,x}(r)\ , \tag{$**$}$$

for $j = 1, 2$, $x^* \in X^*$, $x \in X$.

To prove $(**)$, first note that, since A has dense range and $\{e^{-tA}\}_{t\geq 0}$ is a bounded holomorphic strongly continuous semigroup of angle $\theta > \psi$, $\exists\, \delta > \psi$ such that $\lim_{|z|\to\infty} \sup_{|\arg(z)|<\delta} \|e^{-zA}x\| = 0$ (this follows from [24], A-IV, Corollary 1.14, and the fact that $\{e^{-tA}\}_{t\geq 0}$, $\{e^{-te^{i\delta}A}\}_{t\geq 0}$ and $\{e^{-te^{-i\delta}A}\}_{t\geq 0}$ are bounded strongly continuous holomorphic semigroups). The Cauchy integral formula, applied to $Ae^{-zA}x = \frac{d}{dz}e^{-zA}x$, implies that $\sup_{|\arg(z)|<\psi} \|zAe^{-zA}x\| \to 0$, as $|z| \to \infty$.

Using this fact and the residue theorem, we may replace $[0,\infty)$ by $e^{i\psi}[0,\infty)$, as follows.

$$\begin{aligned} x^*(e^{isB}x) &= \frac{1}{\Gamma(1+is)}\int_{e^{i\psi}[0,\infty)} z^{is}x^*(Ae^{-zA}x)\,dz \\ &= \frac{e^{-s\psi}e^{i\psi}}{\Gamma(1+is)}\int_0^\infty w^{is}x^*(Ae^{-e^{i\psi}wA}x)\,dw, \end{aligned}$$

thus we may choose $f_1(s) \equiv \frac{e^{-s\psi}e^{i\psi}}{\Gamma(1+is)}$, $dE_{1,\phi,x}(r) \equiv x^*(Ae^{-e^{i\psi}e^r A}x)e^r\,dr$. It is clear how we may similarly choose f_2 and $E_{2,\phi,x}$. Thus, since $\psi < \theta$ and $[(\Gamma(1+is))^{-1}]$ and its derivative are $O(e^{|s|\frac{\pi}{2}})$, this proves $(**)$.

Now choose $h \in C^\infty(R)$ such that $h \equiv 1$ on $[1,\infty)$, $\equiv 0$ on $(-\infty,0]$, then let $g_1 \equiv hf_1$, $g_2 \equiv (1-h)f_2$ and we may apply Lemma 3.4.

(a) $\to$ (b). Fix $\psi \in (0,\theta)$. Let $\delta \equiv \frac{\pi}{2} - \frac{1}{2}(\theta+\psi)$. $\exists\, M_\psi < \infty$ such that $\|f(A)\| \le M_\psi \|f\|_\infty$, $\forall\, f \in H^\infty(S_\delta)$.

Fix $x \in X$, $x^* \in X^*$. The map $f \mapsto x^*(f(A)x)$ is a linear functional on $H^\infty(S_\delta) \cap C_0(\overline{S_\delta})$, thus $\exists$ a signed measure $\mu_{x^*,x}$, of variation less than $M_\psi\|x^*\|\,\|x\|$, such that,

$$x^*(f(A)x) = \int_{S_\delta} f(z)\,d\mu_{x^*,x}(z),\ \forall\, f \in H^\infty(S_\delta)\cap C_0(\overline{S_\delta}) \qquad (***)$$

(this follows from the Hahn-Banach and Riesz theorems – see [7]).

Letting $f_\lambda(z) \equiv \lambda(\lambda+z)^{-1}$, dominated convergence and $(***)$ imply that $\lim_{\lambda\to\infty} x^*(\lambda(\lambda+A)^{-1}x) = x^*(x)$, $\forall\, x^* \in X^*$, $x \in X$, so that A is densely defined. Using the same f_λ, $(***)$ also implies that $\{\|w(w+A)^{-1}\| \mid w \in S_{\frac{\pi}{2}+\psi}\}$ is bounded. Thus $-A$ generates a bounded holomorphic strongly continuous group of angle θ.

When $|\phi| < \psi$, letting $f(z) \equiv e^{-te^{i\phi}z}$ in $(***)$, note that, since $(\lambda - A)^{-1} = f_\lambda(A)$, $\forall\, \lambda \in S_\psi$ (see Definition 2.4), it follows, by comparing Laplace transforms, that $f(A) = e^{-te^{i\phi}A}$. Thus

$$\begin{aligned}
\int_0^\infty |x^*(Ae^{-te^{i\phi}A}x)|\,dt &= \int_0^\infty \left|\frac{d}{dt}\left(x^*(e^{-te^{i\phi}A}x)\right)\right| dt \\
&= \int_0^\infty \left|\int_{S_\delta} ze^{-te^{i\phi}z}\,d\mu_{x^*,x}(z)\right| dt \\
&\le \int_{S_\delta}\int_0^\infty \left|ze^{-te^{i\phi}z}\right| dt\,d|\mu_{x^*,x}| \\
&= \int_{S_\delta} \frac{|z|}{|Re(e^{i\phi}z)|}\,d|\mu_{x^*,x}| \le \frac{1}{\cos(\psi+\delta)}|\mu_{x^*,x}|(S_\delta)\ ,
\end{aligned}$$

proving (b). ■

Remark 3.7. It can easily be checked that the above functional calculus is consistent with the calculus constructed in [26], p. 434, via Mellin's transform.

IV. Semigroups and Imaginary Powers

It is well-known that, if A has an $H^\infty(S_\phi)$-functional calculus, then $\{A^{is}\}_{s\in\mathbf{R}}$ extends to a strongly continuous group of bounded operators. On a Hilbert space, the converse is true ([22], [23]), but on a general Banach space it is not.

We do not quite obtain a characterization of A having bounded imaginary powers. We obtain a necessary condition, in terms of the weak operator topology (Theorem 4.1) and we obtain a sufficient condition, that is the same as the necessary condition, except that the weak operator topology is replaced with the strong operator topology (Theorem 4.2). When these topologies are replaced by the operator norm topology, A must be bounded, with bounded inverse (Proposition 4.3).

Theorem 4.1. *Suppose A is of type less than π, $0 \le \theta < \frac{\pi}{2}$ and $\{A^{-is}\}_{s\in\mathbf{R}}$ is a strongly continuous group of bounded operators such that $\forall\ \phi > \theta$, $\exists\ M_\phi < \infty$ such that*

$$\|A^{-is}\| \le M_\phi e^{\phi|s|}, \quad \forall\ s \in \mathbf{R}\ .$$

Then A is of type θ and $\forall\ q \ge 2$, $0 \le \psi < (\frac{\pi}{2} - \theta)$, $\exists\ M_{q,\psi} < \infty$ such that

$$\int_0^\infty \left| tx^* \left(Ae^{-te^{i\delta}A}x \right) \right|^q \frac{dt}{t} < M_{q,\psi} \|x^*\|^q \|x\|^q\ ,$$

$\forall\ x^* \in X^*$, $x \in X$, $|\delta| < \psi$.

When X is a Hilbert space, the following result is in [22].

Theorem 4.2. *Suppose X is reflexive, A is of type $\theta < \frac{\pi}{2}$ and whenever $0 \le \psi < (\frac{\pi}{2} - \theta)$, $\exists\ M_\psi < \infty$ such that*

$$\int_0^\infty \left\| tAe^{-te^{i\delta}A}x \right\|^2 \frac{dt}{t} < M_\psi \|x\|^2 \quad \text{and} \quad \int_0^\infty \left\| tA^* e^{-te^{i\delta}A^*}x \right\|^2 \frac{dt}{t} < M_\psi \|x\|^2\ ,$$

$\forall\ x \in X$, $|\delta| < \psi$.

Then $\{A^{-is}\}$ extends to a strongly continuous group of bounded operators such that $\forall\ \phi > \theta$, $\exists\ M_\phi < \infty$ such that

$$\|A^{-is}\| \le M_\phi e^{\phi|s|}, \quad \forall\ s \in \mathbf{R}\ .$$

If X is a Hilbert space, then the converse is true.

This condition essentially replaces the weak operator topology (as in Theorem 4.1) with the strong operator topology. Hence it is natural to try to use the operator norm topology. However, this reduces to the relatively trivial situation of bounded operators.

It is not hard to show, using the Riesz-Dunford functional calculus, that when A is bounded, with bounded inverse, then $\int_0^\infty \|tAe^{-tA}\|^2 \frac{dt}{t} < \infty$. It is interesting that the converse is also true.

Proposition 4.3. *Suppose $-A$ generates a bounded strongly continuous holomorphic semigroup, $\{e^{-tA}\}_{t\ge0}$. Then the following are equivalent.*

(a) $\int_0^\infty \|tAe^{-tA}\|^2 \frac{dt}{t} < \infty$.

(b) *A is bounded and $0 \in \rho(A)$.*

Proof of Theorem 4.1. By [Theorem 2, 26], A is of type θ. Fix $\psi < (\frac{\pi}{2} - \theta)$. Let $\varepsilon \equiv \frac{1}{2}(\theta + \frac{\pi}{2} - \psi)$. For $|\delta| < \psi$, since $\theta < \varepsilon$, it follows that $\|\Gamma(1+is)((e^{i\delta}A)^{-is}\| \le M_\varepsilon e^{(\varepsilon+\psi-\frac{\pi}{2})|s|}$, $\forall\ s \in \mathbf{R}$. Since $\varepsilon < (\frac{\pi}{2} - \psi)$, this allows us to use, for instance, the functional calculus in [26], as follows:

$$\begin{aligned} x^*((e^{i\delta}A)e^{-e^r e^{i\delta}A}x)e^r &= \int_{\mathbf{R}} \Gamma(1+is)x^*((e^r e^{i\delta}A)^{-is}x)\frac{ds}{2\pi}\ , \\ &= \int_{\mathbf{R}} e^{-isr}\Gamma(1+is)x^*((e^{i\delta}A)^{-is}x)\frac{ds}{2\pi}\ , \end{aligned}$$

$\forall\ x \in X$, $x^* \in X^*$, $r \in \mathbf{R}$.

By the Hausdorff-Young theorem, if $\frac{1}{p}+\frac{1}{q}=1$, then

$$\int_0^\infty \left|tx^*(Ae^{-te^{i\delta}A}x)\right|^q \frac{dt}{t} = \int_{\mathbf{R}} \left|e^r x^*((e^{i\delta}A)e^{-e^r e^{i\delta}A}x)\right|^q dr$$
$$\leq \left[\int_{\mathbf{R}} \left|\Gamma(1+is)x^*((e^{i\delta}A)^{-is}x)\right|^p ds\right]^{\frac{q}{p}}$$
$$\leq \left[\int_{\mathbf{R}} \left|M_\varepsilon e^{|s|(\varepsilon+\psi-\frac{\pi}{2})}\right|^p ds\right]^{\frac{q}{p}} \|x^*\|^q\|x\|^q ,$$

concluding the proof, since $(\varepsilon+\psi-\frac{\pi}{2})<0$. ■

Proof of Theorem 4.2. For $x^* \in X^*$, $x \in D(A^2)$, $\phi > \theta$, let $\psi \equiv (\frac{\pi}{2}-\phi)$. For $|\delta| < \psi$, we calculate as follows, using Definition 2.3(b).

$$\left|x^*((e^{i\delta}A)^{-is}x)\right| = \left|\frac{1}{\Gamma(2+is)}\int_0^\infty t^{is+1}x^*(A^2e^{-te^{i\delta}A}x)\,dt\right|$$
$$\leq \frac{1}{|\Gamma(2+is)|}\left|\int_0^\infty \left(\sqrt{t}A^*e^{-\frac{t}{2}(e^{i\delta}A)^*}x^*\right)\left(\sqrt{t}Ae^{-\frac{t}{2}e^{i\delta}A}x\right)dt\right|$$
$$\leq \frac{1}{|\Gamma(2+is)|}\left(\int_0^\infty t\|A^*e^{-\frac{t}{2}e^{-i\delta}A^*}x^*\|^2\,dt\right)^{\frac{1}{2}}\left(\int_0^\infty t\|Ae^{-\frac{t}{2}e^{i\delta}A}x\|^2\,dt\right)^{\frac{1}{2}}$$
$$\leq \frac{1}{|\Gamma(2+is)|}M_\psi\|x^*\|\,\|x\| .$$

Since $D(A^2)$ is dense, this implies that A^{-is} extends to a bounded operator, $\forall\ s \in \mathbf{R}$ with $e^{s\delta}\|A^{-is}\| \leq M_\psi e^{\frac{\pi}{2}|s|}$, whenever $|\delta| < \psi$; this implies that $\|A^{-is}\| \leq M_\psi e^{(\frac{\pi}{2}-\psi)|s|} = M_\psi e^{\phi|s|}$.

When X is a Hilbert space, the converse follows from the (vector-valued) Plancherel theorem, just as Theorem 4.1 followed from the Hausdorff-Young theorem. ■

Proof of Proposition 4.3. (b) $\rightarrow$ (a). Since $0 \in \rho(A)$ and $\{e^{-tA}\}_{t\geq 0}$ is a bounded holomorphic strongly continuous semigroup, $\|Ae^{tA}\| \leq \|A\|\,\|e^{-tA}\|$ is exponentially decaying.

(a) $\rightarrow$ (b). As in the proof of Theorem 4.2, $\{A^{-is}\}_{s\in\mathbf{R}}$ is a strongly continuous group, given by

$$A^{-is} = \frac{1}{\Gamma(2+is)}\int_0^\infty t^{is+1}A^2e^{-tA}\,dt .$$

Since $\|t^{is+1}A^2e^{-tA}\| \leq |t^{is+1}|\,\|Ae^{-\frac{t}{2}A}\|^2$, dominated convergence now implies that $A^{-is} \rightarrow \int_0^\infty tA^2e^{-tA}\,dt = I$, in the operator norm, as $s \rightarrow 0$. This implies that the generator, $i\log A$, is bounded, hence A is bounded and $0 \in \rho(A)$, ■

V. Resolvents and H^∞ Functional Calculus

In this section we characterize having an $H^\infty(S_\theta)$ functional calculus in terms of the resolvent. This characterization has the advantage that θ need only be less than π, rather than $\frac{\pi}{2}$, as in Section III.

Theorem 5.1. *Suppose* $0 < \theta \le \pi$. *Then the following are equivalent:*

(a) *A has an $H^\infty(S_\phi)$-functional calculus,* $\forall\, \phi > \pi - \theta$.

(b) *A is of type $(\pi-\theta)$ and $\{A(z+A)^{-1}\}_{z\in S_\theta}$ is uniformly weakly of bounded variation.*

Example 5.2. If $-A \equiv \frac{d}{dx}$, the generator of left translation, on $C_0([0,\infty)) \equiv \{f \in C([0,\infty)) | \lim_{x\to\infty} f(x) = 0\}$, then a calculation shows that

$$(A(z+A)^{-1}f)(0) = \int_0^\infty e^{-sz} f'(s)\, ds \ ,$$

for any f in the domain of A. Thus, Theorem 5.1 states that A having an $H^\infty(S_\phi)$ functional calculus, $\forall\, \phi > \frac{\pi}{2}$, is equivalent to the Laplace transforms of the derivatives of functions in the unit ball of $C_0([0,\infty))$ being uniformly of bounded variation. (Note that A has dense range, because translation is stable.)

Proof. (a) $\to$ (b). Fix $\psi \in (0,\theta)$. Let $\delta \equiv \pi - \frac{1}{2}(\theta+\psi)$. $\exists\, M_\psi < \infty$ such that $\|f(A)\| \le M_\psi \|f\|_\infty$, $\forall\, f \in H^\infty(S_\delta)$.

Fix $x \in X$, $x^* \in X^*$. The map $f \mapsto x^*(f(A)x)$ is a linear functional on $H^\infty(S_\delta)$, thus $\exists$ a signed measure $\mu_{x^*,x}$, of variation less than $M_\psi\|x^*\|\,\|x\|$, such that

$$x^*(f(A)x) = \int_{S_\delta} f(z)\, d\mu_{x^*,x}(z), \quad \forall\, f \in H^\infty(S_\delta)$$

(see [7]).

$\exists\, K_\psi < \infty$ such that $\int_0^\infty |z(z+e^{i\phi}t)^{-2}|\, dt \le K_\psi$, when $z \in S_\delta$, $|\phi| < \psi$. Thus, when $|\phi| < \psi$, letting $f(z) \equiv z(z+e^{i\phi}t)^{-2}$,

$$\begin{aligned}\int_0^\infty \left|x^*(A(A+e^{i\phi}t)^{-2}x)\right| dt &= \int_0^\infty \left|\int_{S_\delta} z(z+e^{i\phi}t)^{-2}\, d\mu_{x^*,x}(z)\right| dt \\ &\le \int_{S_\delta}\int_0^\infty \left|z(z+e^{i\phi}t)^{-2}\right| dt\, d\,|\mu_{x^*,x}| \\ &\le K_\psi\, |\mu_{x^*,x}|(S_\delta) \le K_\psi M_\psi \|x^*\|\,\|x\| \ .\end{aligned}$$

(b) $\to$ (a). For $x \in D(A) \cap Im(A)$, we use Definition 2.3(a), to obtain, as in the proof of Theorem 3.1,

$$x^*(A^{-is}x) = \frac{\sin(is\pi)}{is\pi} \int_0^\infty t^{-is} x^*(A(A+t)^{-2}x)\, dt \ .$$

Thus, as in Theorem 3.1, $\{A^{-is}\}_{s\in\mathbf{R}}$ extends to a strongly continuous group of bounded operators; call its generator iB. We have

$$x^*\left(e^{-isB}x\right) = \frac{\sin(is\pi)}{is\pi} \int_0^\infty t^{-is} x^*(A(A+t)^{-2}x)\, dt \ ,$$

$\forall\, x^* \in X^*$, $x \in X$.

The remainder of the proof is exactly like the proof of Theorem 3.1, after its expression for $x^*(e^{-isB}x)$, using Lemma 3.4, with $\frac{1}{\Gamma(1+is)}$ replaced by $\frac{\sin(is\pi)}{is\pi}$. ∎

References

[1] Boyadzhiev, K and R. deLaubenfels, *H^∞-functional calculus for perturbations of generators of holomorphic semigroups*, Houston J. Math. **17** (1991), 131–147.

[2] Champeney, D. C., "A Handbook of Fourier Transforms," Cambridge University Press, Cambridge 1987.

[3] Cioranescu, I. and L. Zsido, *Analytic generators for one-parameter groups*, Tohoku Math. J. **28** (1976), 327–362.

[4] Davies, E. B., "One-Parameter Semigroups," Academic Press, London, 1980.

[5] Dore, G. and A. Venni, *On the closedness of the sum of two closed operators*, Math. Z. **196** (1987), 189–201.

[6] Dore, G. and A. Venni, *Some results about complex powers of closed operators*, J. Math. Anal. and Appl. **149** (1990), 124–136.

[7] Dunford, N. and J. T. Schwartz, "Linear Operators," Part I, Interscience, New York (1958).

[8] Duong, X. T., *H^∞ functional calculus of elliptic operators with C^∞ coefficients on L^p spaces of smooth domains*, J. Austral. Math. Soc. (Ser. A) **48** (1990), 113–123.

[9] Duong, X. T., *H_∞ functional calculus of second order elliptic PDE on L^p spaces*, Miniconference on Operators in Analysis, Proc. of the Center for Math. Analysis, ANU, Canberra **24** (1989), 91–102.

[10] Fattorini, H. O., "The Abstract Cauchy Problem," Addison Wesley, Reading, Mass., 1983.

[11] Fisher, M. J., *Imaginary powers of the indefinite integral*, Amer. J. Math. **93** (1971), 317–328.

[12] Goldstein, J. A., "Semigroups of Operators and Applications," Oxford, New York, 1985.

[13] Hille, E. and R. S. Phillips, "Functional Analysis and Semigroups," Colloq. Publ. Amer. Math. Soc. 1957.

[14] Hughes, R. J., *Semigroups of unbounded linear operators in Banach space*, Trans. Amer. Math. Soc. **230** (1977), 113–145.

[15] Hughes, R. J. and S. Kantorovitz, *Boundary values of holomorphic semigroups of unbounded operators and similarity of certain perturbations*, J. Funct. Anal. **29** (1978), 253–273.

[16] Kalisch, G. K., *On fractional integrals of purely imaginary order in L^p*, Proc. Amer. Math. Soc. **18** (1967), 136–139.

[17] Kober, H., *On a theorem of Schur and on fractional integrals of purely imaginary order*, Trans. Amer. Math. Soc. **50** (1941), 160–174.

[18] Komatsu, H., *Fractional powers of operators*, Pac. J. Math. **19** (1966), 285–346.

[19] Love, E. R., *Fractional derivatives of imaginary order*, J. London Math. Soc. (2) **3** (1971), 241–259.

[20] MacRobert, T. M., "Functions of a Complex Variable," Macmillan, London, 1962.

[21] Marschall, E., *On the analytic generator of a group of operators*, Indiana Univ. Math. J. **35** (1986), 289–309.

[22] McIntosh, A., *Operators which have an H^∞ functional calculus*, Miniconference on Operator Theory and PDE, Proc. of the Center for Math. Analysis, ANU, Canberra **14** (1986), 210–231.

[23] McIntosh, A. and A. Yagi, *Operators of type ω without a bounded H^∞-functional calculus*, Miniconference on Operators in Analysis 1989, Proc. of the Center for Math. Analysis, ANU, Canberra **24** (1989).

[24] Nagel, R., "One-parameter Semigroups of Positive Operators," Lecture Notes in Mathematics 1184, Springer, Berlin, 1986.

[25] Pazy, A., "Semigroups of Linear Operators and Applications to Partial Differential Equations," Springer, New York, 1983.

[26] Prüss, J. and H. Sohr, *On operators with bounded imaginary powers in Banach spaces*, Math. Z. **203** (1990), 429–452.

[27] Reed, M. and B. Simon, "Methods of Mathematical Physics," Vol. II, Academic Press, New York, 1975.

[28] Ricker, W., *Spectral properties of the Laplace operator in $L^p(\mathbf{R})$*, Osaka J. Math. **25** (1988), 399–410.

[29] Ross. B., (ed.) "Fractional Calculus and its Applications," Lecture Notes in Math. **457** (1975), Springer-Verlag.

[30] van Casteren, J. A., "Generators of strongly continuous semigroups," Research Notes in Math., 115, Pitman, 1985.

[31] Yagi, A., *Coincidence entre des espaces d'interpolation et des domaines de puissances fractionnaires d'operateurs*, C. R. Acad. Sci. Paris (Ser. I), **299** (1984), 173–176.

[32] Yagi, A., *Applications of the purely imaginary powers of operators in Hilbert spaces*, J. Func. Anal. **73** (1987), 216–231.

Department of Mathematics
Ohio Northern University
Ada, Ohio 45810, USA

Department of Mathematics
Ohio University
Athens, Ohio 45701, USA

Received March 26, 1991
and in final form April 29, 1992

Semigroup Forum Vol. 45 (1992) 385–394

RESEARCH ARTICLE

The lower central series of the free partially commutative group

G. Duchamp and D. Krob

Communicated by G. Lallement

Abstract.

This paper is devoted to the study of the lower central series of the free partially commutative group $F(A,\vartheta)$ in connection with the associated free partially commutative Lie algebra. Using a convenient Magnus transformation, we show that the quotients of the lower central series of $F(A,\vartheta)$ are free abelian groups and that $F(A,\vartheta)$ can be fully ordered.

0. Introduction

The *free partially commutative monoid* $M(A,\vartheta)$ was introduced by P. Cartier and D. Foata in 1969 (see [6]) for the study of combinatorial problems in connection with word rearrangements. Since that time this monoid was the subject of many studies which were principally motivated by the fact that $M(A,\vartheta)$ can be used as a model for concurrent computing (see [7,10]).

In this paper, we continue the algebraic exploration of the free partially commutative structures we began in [11,12,13]. We are specially interested here with the *free partially commutative group* $F(A,\vartheta)$ (see [7,10]) which is closely related to $M(A,\vartheta)$ since it is defined by the presentation

$$F(A,\vartheta) = < A;\, aba^{-1}b^{-1},\, (a,b) \in \vartheta >$$

Observe that the commutation relators are very typical since they can both be interpreted as monoidal relators ($ab = ba$) and as Lie relators ($[a,b] = 0$) (they are in a certain sense the only ones of this kind (see [13])). Thus it is not surprising that the corresponding free partially commutative Lie algebra $L(A,\vartheta)$ plays an important role in the study of $F(A,\vartheta)$ and that the usual Magnus theory (cf [22]) can be extended to $F(A,\vartheta)$. Note also that the Magnus transformation (i.e. the embedding of $F(A,\vartheta)$ into the group of units of a suited filtered algebra) reproduces "grosso modo" the Lie algebra / Lie group situation, the infinitesimal elements being here the iterated commutators.

Let us end this introduction by giving the structure of our paper that we divided in three sections. In the first one, we just recall some basic facts concerning free partially commutative structures. The second part is devoted to the construction and study of a graded isomorphism between the graded Lie $\mathbb{Z}$-algebra associated with the quotients of the lower central series of $F(A,\vartheta)$ and the Lie algebra $L(A,\vartheta)$ itself. This section ends with a "partially commutative collecting process" in $F(A,\vartheta)$ based on linear decompositions within the Lie algebra $L(A,\vartheta)$. The third part deals with some applications. In particular, we show how to construct orders compatible with the group structures of the Magnus group and of $F(A,\vartheta)$. It follows immediately that their algebras (over fields) can be embedded into skew fields. This section ends with an example of a family of non-isomorphic residually nilpotent groups (cf [17]) that have isomorphic quotients in their lower central series.

1. Preliminaries

1.1. Free partially commutative monoid and group

Let A be an alphabet. Then a *partial commutation relation* ϑ on A is just a symmetric and irreflexive subset of $A \times A$. We can represent ϑ by its *commutation graph* which is the non-directed graph built over A where two letters $a, b \in A$ are related by an edge iff $(a,b) \in \vartheta$. Then the *free partially commutative monoid* on A associated with ϑ is the monoid denoted by $M(A,\vartheta)$ which is defined by the monoid presentation

$$M(A,\vartheta) = < A;\ ab = ba,\ (a,b) \in \vartheta >$$

We denote by $\overline{w}$ the image in $M(A,\vartheta)$ of a word $w \in A^*$ by the natural projection of A^* onto $M(A,\vartheta)$. Since the length of any $u, v \in A^*$ such that $\overline{u} = \overline{v}$ is clearly the same, we can define the *length* (or *degree*) $|w|$ of a word $\overline{w} \in M(A,\vartheta)$ as the length of any representative $w \in A^*$ of $\overline{w}$. We denote by $M_n(A,\vartheta)$ the set of the elements of length n in $M(A,\vartheta)$.

We also denote by $F(A,\vartheta)$ the *free partially commutative group* (cf [7,10]) which is defined by the group presentation

$$F(A,\vartheta) = < A;\ ab = ba,\ (a,b) \in \vartheta >$$

It can be shown (see [7,10]) that the monoid $M(A,\vartheta)$ embeds into the group $F(A,\vartheta)$ in a natural way.

1.2. Partially commutative formal power series

In the sequel, K will be a fixed non-trivial ring. We refer the reader to [4] for all generalities concerning graduations and filtrations. Let us now introduce the K-algebra $K\!<\!< A,\vartheta >\!>$ of the *partially commutative formal power series* over A which is the "large" K-algebra of $M(A,\vartheta)$ (cf [2]). Every series S in $K\!<\!< A,\vartheta >\!>$ can be represented as follows

$$S = \sum_{w \in M(A,\vartheta)} (S|w)\, w$$

The *support* of S is the set $\mathrm{supp}(S) = \{\, w \in M(A,\vartheta),\ (S|w) \neq 0 \,\}$. A series S in $K\!<\!< A,\vartheta >\!>$ is a *partially commutative polynomial* iff its support is finite. We denote by $K\!< A,\vartheta >$ the sub K-algebra of $K\!<\!< A,\vartheta >\!>$ which consists of all polynomials. It can be equipped with the degree graduation defined by

$$\forall\, n \geq 0,\ K_n\!< A,\vartheta > = \{\, P \in K\!< A,\vartheta >,\ \mathrm{supp}(P) \subset M_n(A,\vartheta) \,\}$$

We also define the *valuation* $\nu(S)$ of a series $S \in K\!<\!< A,\vartheta >\!>$ as the minimal length in $\mathbb{N} \cup \{+\infty\}$ of the elements of the support of S. We can now introduce the valuation filtration of $K\!<\!< A,\vartheta >\!>$ defined by

$$\forall\, n \geq 0,\ K^n\!<\!< A,\vartheta >\!> = \{\, S \in K\!<\!< A,\vartheta >\!>,\ \nu(S) \geq n \,\}$$

The *augmentation ideal* $\mathfrak{M}(A,\vartheta) = K^1\!<\!< A,\vartheta >\!>$ is the ideal which consists of the series $S \in K\!<\!< A,\vartheta >\!>$ such that $(S|1) = 0$. As a series of $K\!<\!< A,\vartheta >\!>$ is invertible iff its constant term is invertible in K, we can give:

Definition 1.1. The following subset of $K\!<\! A, \vartheta \!>$

$$\mathfrak{Mg}(A,\vartheta) = \{\, 1+m,\ m \in \mathfrak{M}(A,\vartheta)\,\} = 1 + \mathfrak{M}(A,\vartheta)$$

is a group called the *partially commutative Magnus group.*

Note 1.1. We can equip $\mathfrak{Mg}(A,\vartheta)$ with the filtration induced by the valuation filtration, i.e. defined by $\mathfrak{Mg}^n(A,\vartheta) = 1 + K^n\!<\!< A,\vartheta >\!>$ for every $n \geq 0$.

1.3. The free partially commutative Lie algebra

The *free partially commutative Lie K-algebra* $L(A,\vartheta)$ is the Lie algebra which is defined by the Lie algebra presentation

$$L(A,\vartheta) = <A;\ [a,b] = 0,\ (a,b) \in \vartheta >$$

It can be shown that $L(A,\vartheta)$ is isomorphic to the Lie subalgebra of $K\!<\!A,\vartheta\!>$ generated by the letters of A (see [11,12]). Hence $L(A,\vartheta)$ can be graded by the graduation $(L_n(A,\vartheta))_{n\geq 0}$ induced by the degree graduation of $K\!<\!A,\vartheta\!>$, i.e. defined by $L_n(A,\vartheta) = L(A,\vartheta) \cap K_n\!<\!A,\vartheta\!>$ for every $n \geq 0$. Let us now define by induction the following families of $L(A,\vartheta)$

$$\mathcal{A}_1 = A \quad \text{and} \quad \forall\, n \geq 2,\ \mathcal{A}_n = \{\, [u,v],\ \exists\, p,q \in \mathbb{N}^*,\ p+q = n,\ u \in \mathcal{A}_p,\ v \in \mathcal{A}_q \,\}$$

An element of $\mathcal{A}_n$ will be called a *Lie element* of order n. The Lie elements of order n form a generating family of the K-module $L_n(A,\vartheta)$. We will also denote by $\mathcal{A}$ the family of all Lie elements.

Let us now recall the following important result proved independently in [11], [12] and also by P. Lalonde in [20] (see also [13,14,19]):

Theorem 1.1. *The K-modules $L(A,\vartheta)$ and $L_n(A,\vartheta)$ are free.*

Remark 1.1. When A is a finite alphabet, we will denote by $l(n)$ the rank of the free K-module $L_n(A,\vartheta)$. It is independent of K and can be computed by formulas that generalize the usual Witt's formulas (see [12,19,20]).

1.4. Magnus transformations

For every letter $a \in A$, let $r(a)$ be a one-letter series in $K^2\!<\!< a >\!>$. Then the following property holds in $K\!<\!< A,\vartheta >\!>$:

$$(a,b) \in \vartheta \quad \Longrightarrow \quad (1+a+r(a))(1+b+r(b)) = (1+b+r(b))(1+a+r(a))$$

Hence the universal property of $F(A,\vartheta)$ allows us to give:

Definition 1.2. Let $r = (r(a))_{a\in A}$ be a family of $(K^2\!<\!< a >\!>)_{a\in A}$. Then we will call *partially commutative Magnus transformation* associated with r the group morphism μ_r from $F(A,\vartheta)$ into $\mathfrak{Mg}(A,\vartheta)$ which is defined by

$$\forall\, a \in A,\ \mu_r(a) = 1 + a + r(a)$$

Remark 1.2. Although the study of a general Magnus transformation μ_r can be reduced to the study of the Magnus transformation μ associated with the zero family $r = (0)_{a\in A}$ using the same argument as in [5] II. 5. N°4, it is useful to have all Magnus transformations at our disposal. For instance, the exponential transformation is a Magnus transformation (when $\mathrm{char}(K) = 0$) which is used in the study of the Hausdorff group (cf [5]).

Let us now give the following theorem proved in [15] which generalizes the classical corresponding result for free groups. It shows that a Magnus transformation realizes always an embedding of the free partially commutative group into the Magnus group, hence into the algebra $K\!<\!< A,\vartheta >\!>$. This result was also discovered independently of us by C. Droms ([8]).

Theorem 1.2. *The Magnus morphism μ_r is into.*

1.5. Commutators

Let us recall that the *commutator* (g,h) of two elements g,h of a group G is defined by $(g,h) = ghg^{-1}h^{-1} \in G$. We refer then to [2] or [16] for the definition of the lower central series of a group. Note that similarly to the Lie case, we can define in $F(A,\vartheta)$ the following families of commutators:

$$\mathcal{C}_1 = A \quad \text{and} \quad \forall\, n \geq 2,\ \mathcal{C}_n = \{\, (u,v),\ \exists\ p,q \in \mathbb{N}^*,\ p+q=n,\ u \in \mathcal{C}_p,\ v \in \mathcal{C}_q \,\}$$

"substituing parentheses to brackets". For every $n \geq 1$, an element of $\mathcal{C}_n$ will be called an *homogeneous commutator* of degree n. We will denote by $\mathcal{C}$ the family of all homogeneous commutators.

Let us finally give the following interesting result (whose proof is immediate by induction on n) which precises the action of a Magnus transformation on the lower central series of $F(A,\vartheta)$:

Proposition 1.3. *Let μ_r be the Magnus morphism associated with a family r of $(K^2\!<\!<a>\!>)_{a\in A}$ and let $(F_n(A,\vartheta))_{n\geq 1}$ be the lower central series of $F(A,\vartheta)$. Then we have for every $n \in \mathbb{N}^*$ and for every $g \in F_n(A,\vartheta)$:*

$$\mu_r(g) \in 1 + l_n(g) + \mathbb{Z}^{n+1}\!<\!<A,\vartheta>\!> \quad \textit{where} \quad l_n(g) \in L_n(A,\vartheta)$$

2. The lower central series of $F(A,\vartheta)$

2.1. The quotients of the lower central series of $F(A,\vartheta)$

Let $(F_n(A,\vartheta))_{n\geq 1}$ be the lower central series of $F(A,\vartheta)$. Then the quotient groups $F_n(A,\vartheta)/F_{n+1}(A,\vartheta)$ are all abelian groups, i.e. $\mathbb{Z}$-modules. Thus we can consider the graded $\mathbb{Z}$-module $\mathcal{G}$ defined by

$$\mathcal{G} = \bigoplus_{n\in\mathbb{N}} F_n(A,\vartheta)/F_{n+1}(A,\vartheta)$$

Using the classical technique of Lazard and Magnus (cf [21,22]), it is possible to equip $\mathcal{G}$ with a Lie $\mathbb{Z}$-algebra structure as follows. Let g and h be respectively in $F_p(A,\vartheta)$ and in $F_q(A,\vartheta)$ with $p,q \geq 1$. Then $(g,h) = ghg^{-1}h^{-1} \in F_{p+q}(A,\vartheta)$ according to a classical property of the lower central series (see [2] or [22]). Thus we can equip $\mathcal{G}$ with a Lie bracket defined for every $p,q \in \mathbb{N}^*$, for every $g \in F_p(A,\vartheta)/F_{p+1}(A,\vartheta)$ and $h \in F_q(A,\vartheta)/F_{q+1}(A,\vartheta)$ by the relation

$$[g,h] = (g,h) \in F_{p+q}(A,\vartheta)/F_{p+q+1}(A,\vartheta)$$

Observe also that the following property is obviously satisfied in $\mathcal{G}$:

$$(a,b) \in \vartheta \quad \Longrightarrow \quad [aF_2(A,\vartheta), bF_2(A,\vartheta)] = (a,b)F_3(A,\vartheta) = F_3(A,\vartheta) = 0$$

Hence, according to the universal property of $L_{\mathbb{Z}}(A,\vartheta)$ (cf [11,12]), there exists a unique Lie $\mathbb{Z}$-morphism α from $L_{\mathbb{Z}}(A,\vartheta)$ into $\mathcal{G}$ such that

$$\forall\, a \in A,\ \alpha(a) = aF_2(A,\vartheta) \in \mathcal{G}$$

We can now give the following theorem whose proof follows by using the same argument than in the free case (see [5] II. 5. N°4). Note also that its proof is independent from the injectivity of the Magnus transformation.

Theorem 2.1. *The Lie $\mathbb{Z}$-morphism α is an isomorphism of graded Lie algebras from $L_{\mathbb{Z}}(A,\vartheta)$ graded by $(L_n(A,\vartheta))_{n\geq 1}$ into $\mathcal{G}$.*

As an immediate consequence of th. 2.1, we obtain that the two following *abelian* groups are isomorphic for every $n \geq 1$

$$F_n(A,\vartheta)/F_{n+1}(A,\vartheta) \simeq L_n(A,\vartheta)$$

Moreover, when A is a finite alphabet, we have according to th. 1.1

$$F_n(A,\vartheta)/F_{n+1}(A,\vartheta) \simeq \mathbb{Z}^{l(n)}$$

2.2. The Magnus transformation is filtered

The next result shows that any Magnus morphism μ_r is a filtered morphism. Its proof is exactly the same as in the free group case (see [5] II. 5. N° 4). It is to be noted that it does not use the injectivity of μ_r.

Theorem 2.2. *The following relation holds for every $n \geq 1$:*

$$F_n(A,\vartheta) = \mu_r^{-1}(\mathfrak{Mg}^n(A,\vartheta))$$

It follows clearly from th. 2.2 that $\mu_r(g) \in \mathfrak{Mg}^n(A,\vartheta) - \mathfrak{Mg}^{n+1}(A,\vartheta)$ for every $g \in F_n(A,\vartheta) - F_{n+1}(A,\vartheta)$. Hence the element $l_n(g)$ involved in prop. 1.3 is in particular not 0 when $g \in F_n(A,\vartheta) - F_{n+1}(A,\vartheta)$. Moreover, using the same method than in [5] II.5., it follows from th. 2.2 that we have

$$\bigcap_{n\geq 1} F_n(A,\vartheta) = \mu_r^{-1}(1) \qquad (\mathcal{R})$$

It follows immediately now from this relation and from th. 1.2 that we have:

Theorem 2.3. *The free partially commutative group $F(A,\vartheta)$ is a residually nilpotent group.*

2.3. The collecting process

We shall now show how the usual collecting process for residually nilpotent groups (see [17] and [16] for the free group case) works here. Hence let us first define a *parenthesing* mapping π which maps every element $a \in \mathcal{A}_n$ into $\mathcal{C}_n$ modulo $F_{n+1}(A,\vartheta)$. It is inductively defined by

$$\forall\, a \in A,\ \pi(a) = a\,[F_2(A,\vartheta)]$$

$$\forall\, a = [u,v] \in \mathcal{A}_n \text{ with } u \in \mathcal{A}_p,\ v \in \mathcal{A}_q,\ p+q = n,\ \pi(a) = (\pi(u),\pi(v))[F_{n+1}(A,\vartheta)]$$

It is easy to prove by induction on n that π is just the restriction to $\mathcal{A}$ of the Lie morphism α of section 2.1. Thus our definition really makes sense.

Note 2.1. It is also possible to define with a similar induction a bracketting mapping β from $\mathcal{C}$ into $\mathcal{A}$. Such a definition is consistent since it can be proved that $\beta(c)$ is the Lie element $l_n(c)$ of prop. 1.3 for every $c \in \mathcal{C}_n$.

We can now give the following proposition which precises an algorithmical partially commutative version of the classical collecting process.

Proposition 2.4. *For every $n \geq 1$, let $(l_{n,i})_{i\in I_n}$ be a basis of the free $\mathbb{Z}$-module $L_n(A,\vartheta)$ indexed by a totally ordered set I_n and let $c_{n,i}$ be a representative in $F_n(A,\vartheta)$ of $\pi(l_{n,i})$ for every $i \in I_n$. Then every element $g \neq 1$ of $F(A,\vartheta)$ has a decomposition for every $m \geq 0$ of the form*

$$g = \prod_{k=n}^{n+m} \prod_{i \in I_k} (c_{k,i})^{\alpha_{k,i}} \ [F_{n+m+1}(A,\vartheta)] \qquad \text{with } ((\alpha_{k,i})_{i\in I_k})_k \in \prod_{k=n}^{n+m} \mathbb{Z}^{(I_k)}$$

where n denotes the unique integer such that $g \in F_n(A,\vartheta) - F_{n+1}(A,\vartheta)$ and where every product respects the order of its indexing set.

Proof. Let now $g \neq 1$ be in $F(A,\vartheta)$. By th. 2.3, there exists $n \in \mathbb{N}^*$ such that $g \in F_n(A,\vartheta) - F_{n+1}(A,\vartheta)$. According to the results of section 2.2, there exists $l_n(g) \in L_n(A,\vartheta) - \{0\}$ such that $\mu(g) \in 1 + l_n(g) + \mathbb{Z}^{n+1}\ll A,\vartheta\gg$. We can then decompose $l_n(g)$ on the K-basis $(l_{n,i})_{i\in I_n}$ of $L_n(A,\vartheta)$

$$l_n(g) = \sum_{i\in I_n} \alpha_i\, l_{n,i} \qquad \text{with } (\alpha_i)_i \in \mathbb{Z}^{(I_n)}$$

Let us define now g_n (the product respects the order of I_n) by

$$g_n = \prod_{i\in I_n} (c_{n,i})^{\alpha_i} \ \in F_n(A,\vartheta)$$

Since $F_n(A,\vartheta)/F_{n+1}(A,\vartheta)$ is abelian, we have $\mu(g_n^{-1}g) \in 1 + \mathbb{Z}^{n+1}\ll A,\vartheta\gg$, i.e. $\mu(g_n^{-1}g) \in \mathfrak{Mg}^{n+1}(A,\vartheta)$. Hence it follows from th. 2.2 that

$$g_n^{-1}g \in F_{n+1}(A,\vartheta) \iff g = \prod_{i\in I_n} (c_{n,i})^{\alpha_i} \ [F_{n+1}(A,\vartheta)]$$

It is now easy to obtain the proposition by using an obvious induction. ■

Example 2.1. Let $A = \{a,b,c\}$ be equipped with $\vartheta = \{(a,b),(b,a)\}$ and let us consider the element $g = ac^{-1}b$ for instance. In order to develop g according to the collecting process, we must first construct a $\mathbb{Z}$-basis of the Lie algebra $L(A,\vartheta)$. The method used in [12] gives us

$$L(A,\vartheta) = L(\mathcal{T}) \oplus L(a,c) \qquad \text{where } \mathcal{T} = \{\, b, [c,b], [c,[c,b]], \ldots \}$$

It is easy to compute the image of g by the Magnus transformation μ:

$$\mu(g) = (1+a)(1+c)^{-1}(1+b) \in 1 + (a-c+b) + K^2\ll a\gg$$

Let us take the alphabet A ordered by $a<b<c$ as a set of representatives in $F_1(A,\vartheta)$ of the basis A of $L_1(A,\vartheta)$. Thus we have

$$g = abc^{-1}(cb^{-1}a^{-1}ac^{-1}b) = abc^{-1}(c,b^{-1}) = abc^{-1}\ [F_2(A,\vartheta)]$$

To continue the process, we must now compute

$$\mu(cb^{-1}c^{-1}b) = (1+c)(1-b+b^2+\ldots)(1-c+c^2+\ldots)(1+b)$$

which belongs to $1 + bc - cb + K^3\ll A,\vartheta\gg$. Thus, using the method given by prop. 2.4, we immediately obtain

$$g = abc^{-1}.(c,b)\ [F_3(A,\vartheta)]$$

3. Some applications

3.1. $F(A,\vartheta)$ can be fully ordered

We refer to [3] (VI. 29) for the definition of a fully ordered group. Let now G be a group equipped with a *strictly decreasing* family $(G_n)_{n\geq 1}$ of *normal* subgroups such that the three following conditions hold

$$i)\ \forall\, n \in \mathbb{N}^*,\ \forall\, g \in G,\ \forall\, p \in G_n,\ \exists\, h \in G_n,\ (hg,p) \in G_{n+1}$$

$$ii)\ G_1 = G \qquad iii)\ \bigcap_{n\geq 1} G_n = \{1\}$$

Then we can give the following theorem

Theorem 3.1. *Under the previous assertions, let us suppose that the group G_n/G_{n+1} can be fully ordered by an order $<_n$ for every $n \geq 1$. Then there exists a unique order making G a fully ordered group such that the canonical projections π_n from G_n onto G_n/G_{n+1} are increasing.*

Proof. We recall (see [3] VI.29) that the definition of a full order that is compatible with the group structure of G is equivalent to the definition of a subset $P \subset G$ (the "strictly positive cone") satisfying the properties

$$a)\ P \cap P^{-1} = \emptyset \qquad b)\ \forall\, g \in G,\ gPg^{-1} \subset P$$
$$c)\ P.P \subset P \qquad d)\ P \cup P^{-1} = G - \{1\}$$

According to conditions $ii)$ and $iii)$, we can define an "order function" by setting $\alpha(g) = \sup\{\, n \in \mathbb{N},\ g \in G_n \,\} \in \mathbb{N} \cup \{+\infty\}$ for every $g \in G$. Since $(G_n)_{n\geq 1}$ is a strictly decreasing family of normal subgroups of G, we have

$$(1)\ \ \forall\, g \in G,\ \alpha(g) = \alpha(g^{-1}) \qquad (2)\ \ \forall\, g,h \in G,\ \alpha(hgh^{-1}) = \alpha(g)$$
$$(3)\ \ \forall\, g,h \in G,\ \alpha(gh) \geq \inf(\alpha(g),\alpha(h))$$
$$(4)\ \ \forall\, g,h \in G,\ \alpha(g) < \alpha(h) \implies \alpha(gh) = \alpha(hg) = \alpha(g)$$

Let us first suppose that G is fully ordered by an order $<_G$ such that all the projections π_n are increasing. Then it is easy to prove that we must have

$$g >_G 1 \iff \pi_{\alpha(g)}(g) >_{\alpha(g)} 1 \qquad (*)$$

since the orders on G and on $G_{\alpha(g)}/G_{\alpha(g)+1}$ are full. Thus if the desired order exists, it is unique. Conversely let us define $P = \{\, g \in G - \{1\},\ \pi_{\alpha(g)}(g) >_{\alpha(g)} 1 \,\}$. It is then easy to check that P satisfies to conditions $a)$ and $d)$. Moreover condition $b)$ follows also easily from properties $i)$ and (2). Hence let us just prove that P satisfies condition $c)$.

Condition $c)$: Let g,h be in P. Hence we have here $\pi_{\alpha(g)}(g) >_{\alpha(g)} 1$ and $\pi_{\alpha(h)}(h) >_{\alpha(h)} 1$. Let us suppose first that $\alpha(g) = \alpha(h)$. Then we have

$$\pi_{\alpha(g)}(g).\pi_{\alpha(h)}(h) = \pi_{\alpha(g)}(gh) >_{\alpha(g)} 1 \qquad (\mathcal{R})$$

But property (3) shows that we have here $\alpha(gh) \geq \alpha(g)$. Relation $(\mathcal{R})$ implies that $\alpha(gh) = \alpha(g)$ since otherwise we would have $\alpha(gh) > \alpha(g)$ and hence $\pi_{\alpha(g)}(gh) = 1$ which contradicts $(\mathcal{R})$. Therefore relation $(\mathcal{R})$ says exactly that gh is in P. Let us suppose now that $\alpha(g) \neq \alpha(h)$. Since the other case is similar, we can suppose for instance that $\alpha(g) > \alpha(h)$. Then according to property (4), we have $\alpha(gh) = \alpha(h)$ and we get here

$$\pi_{\alpha(h)}(g) = 1 \quad \text{and} \quad \pi_{\alpha(h)}(h) >_{\alpha(h)} 1 \implies \pi_{\alpha(h)}(gh) >_{\alpha(h)} 1$$

As $\alpha(gh) = \alpha(h)$, this last inequality just means that gh is in P. Hence we proved in all cases that $gh \in P$. Thus condition $c)$ holds.

Since P satisfies conditions $a)$, $b)$, $c)$ and $d)$, we can define a full order on G such that P becomes the positive cone of G for this order. It is then straightforward to check that the projections π_n are all increasing with this order on G. Thus this ends our proof. ∎

Remark 3.1. The order we defined on G can described as follows: if $g \in G - \{1\}$, there is a maximal index n such that $g \in G_n$. Then $\pi_n(g) = gG_{n+1} \neq 1$ and we can decide if $g > 1$ in the quotient G_n/G_{n+1} which is fully ordered.

Notes 3.1. 1) Th. 3.1 can be easily extended (with some technical modifications) to the case of a family of groups $(G_\lambda)_{\lambda \in \Lambda}$ where Λ is an arbitrary ordinal.

2) Let us consider for every $n \geq 1$ the group morphism Φ_n from G into $\mathrm{Aut}(G_n/G_{n+1})$ which associates with every $g \in G$ the inner automorphism of G_n/G_{n+1} naturally defined by g. Then it is easy to see that the technical condition $i)$ is satisfied if $\Phi_n(G) = \Phi_n(G_n)$ for every $n \in \mathbb{N}^*$.

3) The technical condition $i)$ is satisfied if the sequence $(G_n)_{n \geq 1}$ satisfies the property $(G_n, G) \subset G_{n+1}$ for every $n \in \mathbb{N}$, which is stronger than the condition given in 2) and is exactly equivalent to the fact that G_n contains the n-th term of the lower central series of G for every $n \geq 1$.

Corollary 3.2. *There exists an order which makes the partially commutative free group $F(A, \vartheta)$ a fully ordered group with $M(A, \vartheta)$ as positive elements.*

Proof. Since any free abelian group can be fully ordered by a lexicographic ordering for instance, the corollary follows easily from th. 3.1, from the results of section 2.1 and from the above note 3). ∎

Remark 3.2. There are several full orders on $F(A, \vartheta)$ which are given by the previous corollary. The proof of th. 3.1 gives us an algorithm to see how two elements $g, h \in F(A, \vartheta)$ are related by an order $<$ defined by cor. 3.2. It suffices to develop gh^{-1} by the collecting process described in section 2.3. Thus we obtain a decomposition of the form

$$gh^{-1} = \prod_{c \in \mathcal{B}_n} c^{\alpha_c} \ [F_{n+1}(A, \vartheta)] \qquad \text{with } (\alpha_c)_c \in \mathbb{Z}^{(\mathcal{B}n)} - \{0\}$$

where $\mathcal{B}_n$ denotes the basis of the free abelian group $F_n(A, \vartheta)/F_{n+1}(A, \vartheta)$ which is used for lexicographically ordering $\mathbb{Z}^{(\mathcal{B}n)}$ in cor. 3.2. Thus $gh^{-1} > 1$ if and only if $(\alpha_c)_{c \in \mathcal{B}_n} > 0$ in the previous lexicographic order.

Corollary 3.3. *Let K be a ring such that $(K, +)$ is a fully ordered commutative group. Then there exists an order which makes the Magnus group $\mathfrak{Mg}(A, \vartheta)$ a fully ordered group.*

Proof. Since it is easy to check that the filtration $(\mathfrak{Mg}^n(A, \vartheta))_{n \geq 1}$ of the Magnus group satisfies to all the preliminary conditions of th. 3.1, our corollary follows easily from this theorem and from the obvious fact that

$$\forall\, n \geq 1,\ \mathfrak{Mg}^n(A, \vartheta)/\mathfrak{Mg}^{n+1}(A, \vartheta) \simeq (K_n{<}A, \vartheta{>}, +) \simeq (K^{M_n(A, \vartheta)}, +)$$

since, as $(K, +)$ is a fully ordered abelian group, it suffices to put a well order on $M_n(A, \vartheta)$ in order to define a lexicographic order on $K^{M_n(A, \vartheta)}$. ∎

Remark 3.3. As every Magnus morphism μ_r is injective, we can transfer to $F(A, \vartheta)$ every order defined by cor. 3.3 on a Magnus group $\mathfrak{Mg}(A, \vartheta)$ by defining $g < h \iff \mu_r(g) < \mu_r(h)$ for every $g, h \in F(A, \vartheta)$. Hence $F(A, \vartheta)$ becomes a fully ordered group by a different method from cor. 3.2. However it can be checked that we obtain the same orders on $F(A, \vartheta)$ when $K = \mathbb{Z}$.

Corollary 3.4. *Let K be a field. Then the polynomial ring $K\!<\!A,\vartheta\!>$ and the group K-algebra $K[F(A,\vartheta)]$ can be embedded into a Malcev series skew field.*

Proof. Let us fully order $F(A,\vartheta)$ with $<$ according to cor. 3.2. Then we can consider the Malcev series skew field (see [18] or [23])

$$K_M[[F(A,\vartheta)]] = \{\, S \in F(A,\vartheta)^K ,\ \mathrm{supp}(S) \text{ is well ordered for } < \}$$

in which $K[F(A,\vartheta)]$ and $K\!<\!A,\vartheta\!>$ embed obviously. ■

Note 3.2. S. Varrichio used cor. 3.4 for proving the decidability of the equality of two partially commutative rational series over a domain (cf [24]).

3.2. A family of examples

The methods we developped in [12] and here allow us to construct easily non-isomorphic groups such that all the quotients of the lower central series of these groups are isomorphic abelian groups. Such examples were also given by G. Baumslag (see [1]) where non-isomorphic groups with an isomorphic lower central sequence are presented. However, using th. 2.3, it is easy to see that two free partially commutative groups with the same central sequence have the same Lie algebra and hence are isomorphic (see [13]).

Proposition 3.5. *Let ϑ, ϑ' two be commutation relations over a finite alphabet A such that their commutation graphs are non-isomorphic but have the same number of n-cliques for every $n \geq 1$. Then the two free partially commutative groups $F(A,\vartheta)$ and $F(A,\vartheta')$ are non-isomorphic, but have isomorphic quotients in their lower central series for every $n \geq 1$:*

$$F_n(A,\vartheta)/F_{n+1}(A,\vartheta) \simeq F_n(A,\vartheta')/F_{n+1}(A,\vartheta') \qquad (QCS)$$

Proof. Since the graphs (A,ϑ) and (A,ϑ') have the same number of n-cliques for every n, it follows from the Witt's calculus of [12] that $L_n(A,\vartheta)$ and $L_n(A,\vartheta')$ have the same rank for every $n \geq 1$. It follows immediately now from the results of section 2.1 that the successive quotients of the lower central series of $F(A,\vartheta)$ and $F(A,\vartheta')$ are isomorphic abelian groups. On the other hand, $F(A,\vartheta)$ and $F(A,\vartheta')$ are non-isomorphic since two free partially commutative groups are isomorphic iff their commutation graphs are isomorphic (see [9]). ■

Example 3.1. The two following graphs are the smallest examples of commutation relations that satisfy the conditions of prop. 3.5:

c

$\vartheta =$ a ——— b and $\vartheta' =$ a ——— b ——— c ——— d

d

and hence that give us an example of two non-isomorphic groups which are isomorphic in their lower central series quotients.

References

[1] Baumslag G., *Groups with the same lower central sequence as a relatively free group. I The groups*, Trans. AMS, **129** (1967), 308–321.

[2] Bourbaki N., "Algèbre", Chap. 1 à 3, CCLS, 1970.

[3] Bourbaki N., "Algèbre", Chap. 4 à 7, Masson, 1981.

[4] BourbakiN., "Algèbre commutative", Chap. 1 à 4, Masson, 1985.

[5] Bourbaki N., "Algèbres et groupes de Lie", Chap. 2 et 3, CCLS, 1972.

[6] Cartier P., Foata D., "Problèmes combinatoires de commutation et de réarrangements", Lecture Notes in Math., **85**, Springer, 1969.

[7] Choffrut C., *Free partially commutative monoid*, LITP Report 86-20 (1986).

[8] Droms C., *Residual nilpotence of graph groups* (unpublished).

[9] Droms C., *Isomorphisms of graph groups*, Proc. A.M.S. **100**, 3 (1987), 407–408.

[10] Duboc C., *Commutations dans les monoïdes libres: un cadre théorique pour l'étude du parallélisme*, Thèse, Université de Rouen, LITP Report 86-25 (1986).

[11] Duchamp G., *On the free partially commutative Lie algebra*, LITP Report 89-74 (1989).

[12] Duchamp G., Krob D., *The free partially commutative Lie algebra : bases and ranks*, LITP Report 90-87 (1990), To appear in "Advances in Math.".

[13] Duchamp G., Krob D., *Free partially commutative structures*, LITP Report 90-65 (1990), To appear in "Journal of Algebra".

[14] Duchamp G., Krob D., *Factorisations dans le monoïde partiellement commutatif libre*, C. R. Acad. Sci. Paris, I, **312** (1991), 189–192.

[15] Duchamp G., Krob D., *Partially commutative Magnus transformations*, LITP Report (1991).

[16] Hall M., "Theory of groups", Chelsea, 1976.

[17] Hall P., "The Edmonton Notes on nilpotent groups", Queen Mary College Math. Notes, University of Alberta, 1969.

[18] Krob D., *Somes examples of formal series used in non-commutative algebra*, Theor. Comp. Sci., **79** (1991), 111–135.

[19] Lalonde P., *Empilements de Lyndon et bases des algèbres de Lie libres*, (Submitted to Theor. Comput. Sci.).

[20] Lalonde P., "Contribution à l'étude des empilements", Publications du LACIM, **4**, UQAM, Montréal, 1991.

[21] Lazard M., "Groupes, anneaux de Lie et problème de Burnside", Inst. Mat. dell. Universita Roma, 1960.

[22] Magnus W., Kharass A., Solitar D., "Combinatorial group theory", Dover, 1976.

[23] Neumann B.H., *On ordered division rings*, Trans. A.M.S., **66**(1949), 202–252.

[24] Varricchio S., *On the decidability of the equivalence problem for partially commutative rational power series*, LITP Report 90.97 (1990).

Gérard Duchamp - Daniel Krob
Laboratoire d'Informatique de Rouen
CNRS (LITP)
Université de Rouen
Faculté des Sciences
76130 Mont Saint-Aignan - FRANCE

Received June 27, 1991
and in final form November 24, 1991

Semigroup Forum Vol. 45 (1992) 395–397

SHORT NOTE

Perfect Completely Semisimple Inverse Semigroups

S. M. Goberstein

Communicated by N. R. Reilly

Let S be a semigroup. Given subsets A and B of S, define as usual $AB = \{ab : a \in A, b \in B\}$. Then for any congruence ε on S and $a, b \in S$, $(a\varepsilon)(b\varepsilon) \subseteq (ab)\varepsilon$. Following Wagner [5], we say that a congruence ε on S is *perfect* if for all $a, b \in S$, $(a\varepsilon)(b\varepsilon) = (ab)\varepsilon$. This definition is motivated by the fact that if G is a group, H a normal subgroup of G and A, B any two cosets of H in G, then AB is also a coset of H in G, so every congruence on a group is perfect. A semigroup S is called *perfect* if all congruences on S are perfect. Perfect semigroups were studied by Fortunatov in a series of papers. In particular, he determined the structure of perfect Clifford semigroups [2, Theorem 5]. Since perfectness is a "group-like" property of a semigroup and since inverse semigroups represent one of the most important generalizations of groups, it is natural to study, in addition to perfect Clifford semigroups, other classes of perfect inverse semigroups. The structure of finite inverse perfect semigroups was described by Hamilton and Tamura [3]. It ought to be mentioned that finiteness of the semigroups was used at several crucial points in the proofs of the main results of [3]. In this note, using certain facts about general perfect semigroups due to Fortunatov [2], we determine the structure of perfect completely semisimple inverse semigroups and thus generalize simultaneously [2, Theorem 5] and some of the principal results of [3].

In what follows we adopt the terminology and notation of [4]. To indicate that a semigroup S is a semilattice Y of semigroups S_α $(\alpha \in Y)$, we write $S = [Y; S_\alpha]$. If S is a strong semilattice Y of semigroups S_α determined by a system of homomorphisms $\phi_{\alpha,\beta}$ for all $\alpha \geq \beta$ (that is, $\phi_{\alpha,\alpha} = \iota_{S_\alpha}$, $\phi_{\alpha,\beta}\phi_{\beta,\gamma} = \phi_{\alpha,\gamma}$ and $ab = (a\phi_{\alpha,\alpha\beta})(b\phi_{\beta,\alpha\beta})$ for any $a \in S_\alpha$ and $b \in S_\beta$), it is usually denoted by $[Y; S_\alpha, \phi_{\alpha,\beta}]$ (see [4, II.2.2 and II.2.3]). If $S = [Y; S_\alpha, \phi_{\alpha,\beta}]$ is such that Y is a chain and each homomorphism $\phi_{\alpha,\beta}$ is surjective, we say that S is *chain-surjective* (in [2] such strong semilattices of semigroups were called "right"). Recall [4, II.2.6] that S is a Clifford semigroup iff $S = [Y; G_\alpha, \phi_{\alpha,\beta}]$ where each G_α is a group and for any $\alpha \geq \beta$ and $a \in G_\alpha$, $a\phi_{\alpha,\beta} = ae$ where e is the identity of G_β.

Result 1 (from Fortunatov [2, Theorem 2 and its proof, Propositions 6 and 7]). *Let S be a perfect semigroup. Then $S = [Y; S_\alpha]$ and the following statements hold:*

(i) *Y is a chain;*

(ii) *if S does not have a zero, then for any $\alpha \in Y$, S_α is a simple semigroup;*

(iii) *if S contains a zero 0, then Y has a least element α_0, S_α is a simple semigroup for every $\alpha \neq \alpha_0$, and either $S_{\alpha_0} = \{0\}$ or S_{α_0} is a 0-minimal ideal of S whose zero is not adjoined;*

(iv) *if* $\alpha > \beta$, *then for any* $a \in S_\alpha$ *and* $b \in S_\beta$, *there exist* $x, y \in S_\beta$ *such that* $b = ax = ya$.

Result 2 (Fortunatov [2, Theorem 5]). *A Clifford semigroup* $S = [Y; G_\alpha, \phi_{\alpha,\beta}]$ *is perfect iff it is chain-surjective.*

Let S be an inverse semigroup, and let $\Omega(S)$ be the translational hull of S. By [4,V.1.2] we can identify $\Omega(S)$ with the semigroup of all right translations of S, which are linked to some left translations of S, and note that each $\omega \in \Omega(S)$ is linked to exactly one left translation which we also denote by ω (if $x \in S$, then the image of x under ω considered as a right [left] translation is denoted by $x\omega$ [by ωx]).

Recall that a semigroup S is said to be *completely semisimple* [4, I.6.19] if all principal factors of S are completely 0-simple or completely simple. By a *Brandt semigroup* we mean a completely 0-simple inverse semigroup whose zero is not adjoined (this differs slightly from the usual definition – cf. [4, II.3.1]). The following theorem extends Result 2 and [3, Theorem 1.11 and Lemma 2.2] to the class of completely semisimple inverse semigroups.

Theorem 3. *Let S be a completely semisimple inverse semigroup. Then S is perfect iff it is one of the following:* (1) *a chain-surjective Clifford semigroup;* (2) *a Brandt semigroup;* (3) *an ideal extension of a Brandt semigroup B by C^0 where $C = [Z; G_\alpha : \phi_{\alpha,\beta}]$ is a chain-surjective Clifford semigroup and for each $\alpha \in Z$, there is a homomorphism $\psi_\alpha : G_\alpha \to \Omega(B)$ such that:*

(a) *for any $b \in B$ and $g \in G_\alpha (\alpha \in Z)$, $bg = b(g\psi_\alpha)$ and $gb = (g\psi_\alpha)b$;*

(b) *for every $\alpha \in Z$, $G_\alpha \psi_\alpha$ is a subgroup of the symmetric group on B;*

(c) $\phi_{\alpha,\beta}\psi_\beta = \psi_\alpha$ *for all* $\alpha \geq \beta$.

Proof. Suppose that S is perfect. Then, by Result 1(i), $S = [Y; S_\alpha]$ where Y is a chain. Assume that either S has no zero or its zero is adjoined. Since S is completely semisimple and inverse, it follows from Result 1((ii) and (iii)) that for every $\alpha \in Y$, S_α is a group. Thus S is a perfect Clifford semigroup and, by Result 2, it is chain-surjective. Now assume that S has a zero which is not adjoined. Since S is completely semisimple and inverse, Result 1(iii) implies that Y has a least element α_0, S_{α_0} is a Brandt semigroup and S_α is a group for every $\alpha \neq \alpha_0$. If $Y = \{\alpha_0\}$, S is a Brandt semigroup. Otherwise denote $B = S_{\alpha_0}$, $Z = Y \backslash \{\alpha_0\}$, $G_\alpha = S_\alpha$ for each $\alpha \in Z$, and $C = S \backslash B$. Then C is a Clifford semigroup, $C = [Z; G_\alpha, \phi_{\alpha,\beta}]$, and S is an ideal extension of B by C^0. By [2, Proposition 5] C is perfect and hence chain-surjective according to Result 2.

Let $\tau = \tau(S{:}B)$ be the canonical homomorphism of S into $\Omega(B)$ [4, I.9.2 and I.9.3]. Thus $(s\tau)b = sb$ and $b(s\tau) = bs$ for all $s \in S$ and $b \in B$. For each $\alpha \in Z$, set $\psi_\alpha = \tau | G_\alpha$ and denote the identity of G_α by e_α. Then (a) holds simply by definition of ψ_α. Take any $\alpha \in Z$ and $g \in G_\alpha$. Since S is perfect and B is an ideal of S, we have $Bg = gB = B$ [2, Proposition 2]. Hence $g\psi_\alpha$ maps B onto B. By Result 1(iv), for any $b \in B$, $b = e_\alpha b = be_\alpha$. It follows that if $b_1(g\psi_\alpha) = b_2(g\psi_\alpha)$ for some $b_1, b_2 \in B$, then $b_1 = b_1 e_\alpha = b_1 gg^{-1} = b_2 gg^{-1} = b_2 e_\alpha = b_2$. Thus $g\psi_\alpha$ is a permutation on B, so (b) holds. Now take $\beta \leq \alpha$. For any $b \in B$,

$$b(g\psi_\alpha) = (be_\beta)g = b(e_\beta g) = b(g\phi_{\alpha,\beta}) = b[(g\phi_{\alpha,\beta})\psi_\beta],$$

so $g\psi_\alpha = (g\phi_{\alpha,\beta})\psi_\beta$. Hence (c) also holds.

By [2, Theorem 1] completely [0]-simple semigroups are perfect. This and Result 2 imply that in order to prove the converse it is sufficient to assume that S has the structure described in case (3). Take any congruence ε on S. Let $\varepsilon_B = \varepsilon \cap B \times B$, $\varepsilon_\alpha = \varepsilon \cap G_\alpha \times G_\alpha$ $(\alpha \in Y)$ and $\overline{\varepsilon} = \varepsilon_B \cup \left(\bigcup_{\alpha \in Y} \varepsilon_\alpha \right)$. If $b \in B$ and $g \in G_\alpha$ for some $\alpha \in Y$, then $(b\varepsilon_B)(g\varepsilon_\alpha) = (bg)\varepsilon_B$ and $(g\varepsilon_\alpha)(b\varepsilon_B) = (gb)\varepsilon_B$ due to (a) and (b). From this and from Result 2 together with [2, Theorem 1], it follows that $\overline{\varepsilon}$ is a perfect congruence on S. According to [1, Propositions 18 and 19], perfectness of $\overline{\varepsilon}$ implies that of ε if for any $\alpha \in Z$, the following conditions are satisfied:

1) if $b_0\varepsilon \cap G_\alpha \neq \emptyset$ for *some* $b_0 \in B$, then $b\varepsilon \cap G_\alpha \neq \emptyset$ for *every* $b \in B$;
2) if $\beta < \alpha$ and $g_0\varepsilon \cap G_\alpha \neq \emptyset$ for *some* $g_0 \in G_\beta$, then $g\varepsilon \cap G_\alpha \neq \emptyset$ for *every* $g \in G_\beta$.

Condition 2) holds since $\phi_{\alpha,\beta}$ is surjective. Now suppose that $b_0 \varepsilon g$ for some $b_0 \in B$ and $g \in G_\alpha$. Since $g\psi_\alpha$ is a permutation on B, there exists $x \in B$ such that $xb_0 = 0$ but $x(g\psi_\alpha) \neq 0$ (cf. [3, the proof of Lemma 1.8]). Then $(xb_0, xg) = (0, x(g\psi_\alpha)) \in \varepsilon$, so by [4, II.3.10 (i)], $\varepsilon_B = B \times B$. It follows that $b \varepsilon g$ for every $b \in B$ and hence condition 1) holds. This completes the proof. ■

Using [4, II.2.8 and II.3.7] and Theorem 3, one can easily formulate an isomorphism criterion for perfect completely semisimple inverse semigroups and thus generalize results from [3, Section 2]. Moreover, Theorem 3 can be used to extend the description of congruences on finite inverse perfect semigroups given in [3, Theorems 3.3 and 3.7] to perfect completely semisimple inverse semigroups. Since these generalizations are routine, we do not give them here.

References

[1] Fortunatov, V. A., *Perfect semigroups decomposable in a semilattice of rectangular groups*, Studies in Algebra, Saratov Univ. Press, **No. 2**, (1970), 67–78 (in Russian).

[2] Fortunatov, V. A., *Perfect semigroups*, Izv. Vysš. Učebn. Zaved. Matematika, **No. 3**, (1972), 80–90 (in Russian).

[3] Hamilton, H. and T. Tamura, *Finite inverse perfect semigroups and their congruences*, J. Austral. Math. Soc. (Ser. A), **32** (1982), 114–128.

[4] Petrich, M., "Inverse Semigroups", Wiley, New York (1984).

[5] Wagner, V. V., *Algebraic topics of the general theory of partial connections in fiber bundles*, Izv. Vysš. Učebn. Zaved. Matematika, **No. 11**, (1968), 26–32 (in Russian).

Department of Mathematics & Statistics
California State University, Chico
Chico, California 95929-0525
USA

Received November 14, 1990
and in final form March 1, 1991

Semigroup Forum Vol. 45 (1992) 398

ANNOUNCEMENT

The INTERNATIONAL CONFERENCE ON SEMIGROUPS AND ALGEBRAS OF COMPUTER LANGUAGES will be held in Qingdao, China, May 25 – 28, 1993.

The conference is jointly organized by Qingdao University, Lanzhou University, Shandong Normal University and Qinghai Normal University.

The purpose of the conference is to discuss recent developments in the fields of codes, orders, free monoids, transformation and algebraic semigroups, automata, formal languages, word problems, and combinatorics.

Qingdao is situated on the shore of the Yellow Sea and is well known in China for swimming, seafood, and beers. After the conference there will be a two-day long excursion (29 – 30 May) to the Mountain Tai to visit the famous Confucian Temple and home town of Confucius. A possible tour to Mt. Laoshan can also be arranged for participants.

Everyone interested in the conference is invited to write for information to:

DR. K. P. SHUM
DEPARTMENT OF MATHEMATICS
CHINESE UNIVERSITY OF HONG KONG
SHATIN, N. T.
HONG KONG
EMAIL: B121715@vax.csc.cuhk.hk
TELEX: 60301 CUHK HX
CABLES: SINOVERSITY
FAX: (852) 603-5154

or to:

PROFESSOR G. F. ZHOU
DEPARTMENT OF COMPUTER AND INFORMATION SCIENCE
QINGDAO UNIVERSITY
QINGDAO 26671
P. R. CHINA
FAX: (0086) 532-514713