ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 44

FASC. 3—4

SZEGED, 1982

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

44. KÖTET
FASC. 3—4

SZEGED, 1982

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

Acta Sci. Math. 44 (1982), 185—213

Strong subband-parcelling extensions of orthodox semigroups

MÁRIA B. SZENDREI

The notion of subband-parcelling and strong subband-parcelling congruences on orthodox semigroups was introduced in [4] by generalizing the common properties of congruences appearing in a number of structure theorems concerning orthodox semigroups. For a list of such structure theorems the reader is referred to [4]. In particular, the concept of strong subband-parcelling congruences includes the idempotent separating congruences and the least inverse semigroups congruences on orthodox semigroups which play singificant roles in the theory of orthodox semigroups.

An orthodox semigroup T is said to be a strong subband-parcelling extension of the semigroup S if there exists a strong subband-parcelling congruence $\varkappa$ on T such that $T/\varkappa$ is isomorphic to S.

The aim of the present paper is to describe all strong subband-parcelling extensions of orthodox semigroups. All strong subband-separating extensions of orthodox semigroups are characterized in [5]. It is worth dealing with this special case separately because a much simpler construction is needed than in the general case. On the other hand, in the class of all inverse semigroups every subband-parcelling congruence is subsemilattice-separating.

In Sections 2 and 3 we introduce the construction which will be used in Section 4 to describe the strong subband-parcelling extensions of orthodox semigroups. At the end of this section we apply our results to characterize an orthodox semigroup as an extension of an inverse semigroup by the least inverse semigroup congruence. Thus we obtain a structure theorem for orthodox semigroups which describes orthodox semigroups by means of their bands of idempotents and greatest inverse semigroup homomorphic images as it was done also by YAMADA ([7]). However, our construction seems to be easier to apply in certain cases than the quasi-direct product used by him.

The notions and notations of [1] and [2] are used.

Received on September 4, 1980.

1. Preliminary notions and results

The concept of a subband-parcelling congruence on an orthodox semigroup was introduced in [4] as follows.

Let B be a band. Suppose δ is a congruence on B with $\delta \subseteqq \mathcal{D}$ and $\bar{B}$ is a subband in B with the property that $\bar{B}$ is a union of δ-classes. For brevity, if $\bar{B}$ and δ satisfy these conditions then we say that $\bar{B}, \delta$ is an *associated pair* in B. Let T be an orthodox semigroup with band of idempotents B.

Definition. The congruence relation $\varkappa$ on T is called $(\bar{B}, \delta)$*-parcelling* if the following conditions are fulfilled:

(d_1) $\delta \subseteqq \varkappa | B$,

(d_2) every $\varkappa$-class containing an idempotent element contains an element of $\bar{B}$, and

(d_3) the elements of $\bar{B}$ belonging to a $\varkappa$-class form a δ-class which is the greatest one in this $\varkappa$-class. (By the order of δ-classes we mean the natural order of B/δ.)

In particular, if δ is the identical congruence then $\varkappa$ is called $\bar{B}$*-separating*. In this case (d_1) is satisfied trivially and (d_3) means that every element of $\bar{B}$ is the greatest idempotent in the $\varkappa$-class containing it.

The following proposition characterizes the subband-parcelling congruences.

Proposition 1.1. *Let T be an orthodox semigroup with band of idempotents B. The congruence relation $\varkappa$ on T is subband-parcelling if and only if there exists a greatest $\mathcal{D}$-class in the band of idempotents of each idempotent $\varkappa$-class and their union is a subband in B.*

Proof. Suppose first that $\bar{B}, \delta$ is an associated pair in B and $\varkappa$ is a $(\bar{B}, \delta)$-parcelling congruence. Let K be an idempotent $\varkappa$-class in T. Since T is regular K contains an idempotent element and hence, by (d_2) and (d_3), there exists a greatest δ-class in the band of idempotents E of K and this δ-class is just $\bar{B} \cap E$. Since $\delta \subseteqq \varkappa | B$ by (d_1), $\delta | E \subseteqq \mathcal{D}_E$ is implied by $\delta \subseteqq \mathcal{D}_B$. Therefore $e_1 \mathcal{D}_E \leqq e_2 \mathcal{D}_E$ follows from $e_1 \delta \leqq e_2 \delta$ for every pair e_1, e_2 in E which shows that $\bar{B} \cap E$ is the greatest $\mathcal{D}$-class in E. Clearly, the union of these $\mathcal{D}$-classes for all $\varkappa$-classes is just $\bar{B}$.

Conversely, assume that $\varkappa$ is a congruence on T with the properties that the band of idempotents of each idempotent $\varkappa$-class contains a greatest $\mathcal{D}$-class and their union is a subband $\bar{B}$ in B. Consider the congruence $\delta = \varkappa | B \cap \mathcal{D}_B$ on B. Then the greatest $\mathcal{D}$-class in the band of idempotents of an idempotent $\varkappa$-class is the greatest δ-class and $\bar{B}, \delta$ is an associated pair in B. One can easily see that $\varkappa$ is a $(\bar{B}, \delta)$-parcelling congruence. The proof is complete.

The most important properties of subband-parcelling congruences proved in

[4] are drawn up in the following results. If $\bar{B}, \delta$ is an associated pair in B then $\bar{\delta}$ is used to mean $\delta|\bar{B}$.

Theorem 1.2 ([4] Theorem 2.5). *Suppose T is an orthodox semigroup with band of idempotents B and $\bar{B}, \delta$ is an associated pair in B. If there exists a $(\bar{B}, \delta)$-parcelling congruence $\varkappa$ on T then B is a band $\bar{B}/\bar{\delta}$ of the bands $F_x=\{b\in B: b\delta\leqq x$ and $b\delta\leqq y$ implies $x\leqq y$ for every y in $\bar{B}/\bar{\delta}\}$ $(x\in\bar{B}/\bar{\delta})$ with greatest δ-class x. The $\varkappa$-class containing the δ-class x is an orthodox subsemigroup in T with band of idempotents F_x.*

Remark. F_x is given in [4], Theorem 2.5, in a slightly modified form. The equivalence of these characterizations can be easily checked.

Lemma 1.3. *Suppose T is an orthodox semigroup with band of idempotents B and $\bar{B}, \delta$ is an associated pair in B. Assume that T has a $(\bar{B}, \delta)$-parcelling congruence. Let t and t^* be inverses of each other in T such that $tt^*\in F_{\bar{x}}$ and $t^*t\in F_{\bar{y}}$. Then, for every x and y in $\bar{B}/\bar{\delta}$ with $x\mathcal{R}\bar{x}$ and $y\mathcal{L}\bar{y}$, there exists an inverse t' of t such that $tt'\in F_x$ and $t't\in F_y$.*

Theorem 1.4 ([4] Theorem 2.9). *Let T be an orthodox semigroup with band of idempotents B and $\bar{B}, \delta$ an associated pair in B. Suppose T to have a $(\bar{B}, \delta)$-parcelling congruence. Then $S_B=\{t\in T: e\mathcal{R}t\mathcal{L}f$ for some e, f in $\bar{B}\}$ is an orthodox subsemigroup in T. The band of idempotents in S_B is $\bar{B}$ and the inverses of the elements in S_B belong to S_B.*

The subsemigroup S_B plays a significant role in the case of strong subband-parcelling congruences.

Definition. The $(\bar{B}, \delta)$-parcelling congruence $\varkappa$ on the orthodox semigroup T is called *strong* if every $\varkappa$-class contains an element of S_B.

Obviously, the (B, δ)-parcelling congruences are strong for $S_B=T$.

The following result is important in describing the strong subband-parcelling extensions of orthodox semigroups.

Proposition 1.5 ([4] Theorem 2.10). *Assume that T is an orthodox semigroup with band of idempotents B and $\bar{B}, \delta$ is an associated pair in B. Let $\varkappa$ be a strong $(\bar{B}, \delta)$-parcelling congruence on T. Consider two elements t and t' in T which are inverses of each other. Then there exist elements s and s' in S_B with $s\varkappa t$ and $s'\varkappa t'$ such that s and s' are inverses of each other.*

2. Semidirect product of a partial left band and a right orthodox partial semigroup

The structure of completely regular semigroups was characterized among others by WARNE in [6]. We will apply his result in the special case of bands.

Let Y be a semilattice and I_α a left zero semigroup for every α in Y. A partial groupoid on $I=\bigcup\{I_\alpha:\alpha\in Y\}$ is called a *lower associative semilattice Y of left zero semigroups* I_α $(\alpha\in Y)$ if (i) $I_\alpha\cap I_\beta=\square$ whenever $\alpha\neq\beta$, (ii) the product of elements a in I_α and b in I_β is defined if and only if $\alpha\geqq\beta$, (iii) if $\alpha\geqq\beta$ then $I_\alpha I_\beta\subseteqq I_\beta$ and (iv) if $\alpha\geqq\beta\geqq\gamma$ and $a\in I_\alpha, b\in I_\beta, c\in I_\gamma$ then $a(bc)=(ab)c$. The notion of an *upper associative semilattice of right zero semigroups* is obtained dually.

Let I be a lower associative semilattice Y of left zero semigroups I_α $(\alpha\in Y)$ and J an upper associative semilattice Y of right zero semigroups J_α $(\alpha\in Y)$. For every u in J, let A_u be a transformation of I and, for every a in I, let B_a be a transformation of J such that $aA_u\in I_{\alpha\beta}$ and $uB_a\in J_{\alpha\beta}$ provided $a\in I_\alpha$ and $u\in J_\beta$. Moreover, let the following conditions be fulfilled:

(W1) if $a\in I_\alpha$, $b\in I_\beta$ with $\alpha\geqq\beta$ and $u\in J$ then

(a) $uB_{ab}=uB_aB_b$,

(b) $(ab)A_u=aA_u\cdot bA_{uB_a}$;

(W2) if $u\in J_\alpha$, $v\in J_\beta$ with $\alpha\leqq\beta$ and $a\in I$ then

(a) $aA_{uv}=aA_vA_u$,

(b) $(uv)B_a=uB_{aA_v}\cdot vB_a$.

A pair A, B satisfying these conditions is termed an *(I, J)-pair*.

Define a multiplication on the set $E=\bigcup\{I_\alpha\times J_\alpha:\ \alpha\in Y\}$ by

$$(a,u)(b,v)=(a\cdot bA_u,\ uB_b\cdot v).$$

One can show that E is a band with respect to this multiplication. This band is called a *semidirect product of I and J* and is denoted by $\mathscr{B}(I, J; A, B)$.

Theorem 2.1 (WARNE [6]). *Every band is isomorphic to a semidirect product of some I and J where I is a lower associative semilattice Y of left zero semigroups and J is an upper associative semilattice Y of right zero semigroups for some semilattice Y.*

First we generalize the notion of an upper associative semilattice of right zero semigroups by introducing the notion of a right orthodox partial semigroup. We need the definition of the spined product of partial groupoids. The notion of spined product of semigroups is due to KIMURA [3].

Let both S and T be partial groupoids which are semilattices Y of full subgroupoids S_α and T_α $(\alpha\in Y)$, respectively. By the *spined product $S\otimes_Y T$ of S and T*

over Y we mean the subdirect product of S and T whose underlying set is

$$\cup\{S_\alpha \times T_\alpha : \alpha \in Y\}.$$

Let $\bar{Y}$ be a semilattice. Let $\bar{J}$ be an upper associative semilattice $\bar{Y}$ of right zero semigroups $\bar{J}_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$. For every $\bar{\alpha}$ in $\bar{Y}$, suppose $Y_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$ to be a semilattice with identity $\bar{\alpha}$. Assume that Y is a semilattice $\bar{Y}$ of semilattices $Y_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$ such that $\bar{Y}$ is a subsemilattice in Y. Let J be a partial groupoid with respect to the operation denoted by "$\cdot$". We say that an element σ in J is *idempotent* if $\sigma\cdot\sigma$ is defined and $\sigma\cdot\sigma=\sigma$. Suppose $J=\cup\{J^{\bar{j}}_\alpha:(\bar{j},\alpha)\in\bar{J}\otimes_{\bar{Y}}Y\}$ where $J^{\bar{j}}_\alpha\cap J^{\bar{k}}_\beta=\square$ provided $(\bar{j},\alpha)\neq(\bar{k},\beta)$. Introduce the following notation: for every α in $Y_{\bar{\alpha}}$, let $J_\alpha=\cup\{J^{\bar{j}}_\alpha:\bar{j}\in\bar{J}_{\bar{\alpha}}\}$. Assume that there exists a unary operation "$'$" on J such that the following hold: for arbitrary elements ϱ in $J^{\bar{j}}_\alpha$, σ in $J^{\bar{k}}_\beta$ and τ in $J^{\bar{l}}_\gamma$ with $\varrho'\in J_{\alpha'}$, $\sigma'\in J_{\beta'}$, and $\tau'\in J_{\gamma'}$ we have

(B1) α and α' are contained in the same $Y_{\bar{\alpha}}$ and $\varrho''\in J_\alpha$;
(B2) $\varrho\cdot\sigma$ is defined in J if and only if $\alpha'\leqq\beta$ and, in this case, $\varrho\cdot\sigma\in J^{\bar{j}\cdot\bar{k}}_\alpha$ and $(\varrho\cdot\sigma)'\in J_\xi$ for some ξ with $\xi\leqq\beta'$;
(B3) if $\alpha'\leqq\beta$ and $\beta'\leqq\gamma$ then $(\varrho\cdot\sigma)\cdot\tau=\varrho\cdot(\sigma\cdot\tau)$;
(B4) $\varrho\cdot\varrho'\cdot\varrho=\varrho$ and $\varrho'\cdot\varrho\cdot\varrho'=\varrho'$;
(B5) the idempotents in J_α form a right zero semigroup for each α in Y;
(B6) if ϱ and σ are idempotents with $\alpha\leqq\gamma$ and $\beta\leqq\gamma'$ then $(\varrho\cdot\tau)'\in J_\beta$ if and only if $(\sigma\cdot\tau')'\in J_\alpha$.

Note that both sides of the equality in (B3) are defined by (B2). Parentheses are not needed in (B4) by (B3). Moreover, J_α contains an idempotent for every α in Y as $\varrho\in J_\alpha$ implies by (B4) and (B2) that $\varrho\cdot\varrho'$ is an idempotent in J_α. In property (B2) the product $\bar{j}\cdot\bar{k}$ is defined in $\bar{J}$ since $\alpha\in Y_{\bar{\alpha}}$, $\beta\in Y_{\bar{\beta}}$ imply $\alpha'\in Y_{\bar{\alpha}}$ whence it follows by $\alpha'\leqq\beta$ that $\bar{\alpha}\leqq\bar{\beta}$. Since ϱ and σ are idempotent in (B6), we have $\alpha'\leqq\alpha$ and $\beta'\leqq\beta$. Thus the products in (B6) are defined.

A partial groupoid J fulfilling the above conditions is termed a *right orthodox partial semigroup over* $\bar{J}\otimes_{\bar{Y}}Y$. Dually, one can define a *left orthodox partial semigroup I over* some $\bar{I}\otimes_{\bar{Y}}Y$ where $\bar{I}$ is a lower associative semilattice $\bar{Y}$ of left zero semigroups. The duals of properties (B1)—(B6) will be referred to as (B1)*—(B6)*. If the elements of J or I are idempotent then we call them a *partial right band* and a *partial left band*, respectively.

Suppose $\varrho\in J_\alpha$ and $\varrho'\in J_{\alpha'}$. An element $\sigma\in J_\beta$ with $\sigma'\in J_{\beta'}$ is called an *inverse* of ϱ provided $\alpha'\leqq\beta$, $\beta'\leqq\alpha$ and $\varrho\sigma\varrho=\varrho$, $\sigma\varrho\sigma=\sigma$. Property (B4) means that the operation "$'$" picks out an inverse of each element. Observe that $\varrho\cdot\sigma$ and $\sigma\cdot\varrho$ are idempotent provided ϱ and σ are inverses of each other.

In what follows we draw up the basic properties of a right orthodox partial semigroup J in several lemmas. For brevity, introduce the following notations. If

$\alpha \leqq \beta$ in Y then we denote this fact also by $J_\alpha \leqq J_\beta$. If $\sigma \in J_\alpha$ then $J(\sigma)$ is used to mean J_α.

Lemma 2.2. *If* $\varrho \in J_\alpha$, $\varrho', \sigma \in J_{\alpha'}$, *and* $\sigma' \in J_{\alpha''}$ *then* $(\varrho \cdot \sigma)' \in J_{\alpha''}$.

Proof. By (B1) and (B2), the product $\varrho' \cdot (\varrho \cdot \sigma)$ is defined and $J((\varrho' \cdot (\varrho \cdot \sigma))') \leqq \leqq J((\varrho \cdot \sigma)')$. On the other hand, (B3) ensures that $(\varrho' \cdot \varrho) \cdot \sigma = \varrho' \cdot (\varrho \cdot \sigma)$. Moreover, the product $\varrho \cdot ((\varrho' \cdot \varrho) \cdot \sigma)$ is also defined by (B2) and properties (B3), (B4) imply it to be equal to $\varrho \cdot \sigma$. Again by (B2), we have $J((\varrho \cdot \sigma)') \leqq J((\varrho' \cdot (\varrho \cdot \sigma))')$. Hence we obtain that $J((\varrho \cdot \sigma)') = J(((\varrho' \cdot \varrho) \cdot \sigma)')$. It follows from (B2), (B3) and (B4) that $\varrho' \cdot \varrho$ and $\sigma' \cdot \sigma$ are idempotents in $J_{\alpha'}$ and $J_{\alpha''}$, respectively. Thus, by (B6), we have $J(((\varrho' \cdot \varrho) \cdot \sigma)') = J_{\alpha''}$ as $((\sigma' \cdot \sigma) \cdot \sigma')' = \sigma'' \in J_{\alpha'}$, by (B4) and (B1). Hence $J((\varrho \cdot \sigma)') = J_{\alpha''}$ which was to be proved.

Lemma 2.3. *If* $\varrho \in J_\alpha$, $\varrho' \in J_{\alpha'}$, $\sigma \in J_\beta$, $\sigma' \in J_{\beta'}$ *and* $\alpha' \leqq \beta$ *then* $(\varrho \cdot \sigma)' \in J_{\beta'}$ *implies* $\alpha' = \beta$.

Proof. By (B2), $\varrho \cdot \sigma \in J_\alpha$. Thus $(\varrho' \cdot (\varrho \cdot \sigma))' = ((\varrho' \cdot \varrho) \cdot \sigma)' \in J_{\beta'}$ by Lemma 2.2. Since $\varrho' \cdot \varrho \in J_{\alpha'}$, $\sigma' \cdot \sigma \in J_{\beta'}$ and they are idempotent property (B6) ensures $((\sigma' \cdot \sigma) \cdot \sigma')' \in J_{\alpha'}$. On the other hand, (B1) and (B4) imply $((\sigma' \cdot \sigma) \cdot \sigma')' = \sigma'' \in J_\beta$. Hence $\alpha' = \beta$.

Lemma 2.4. *If* ϱ *and* ϱ^* *are inverses of each other in* J *then* $J(\varrho) = J(\varrho^{*\prime})$ *and* $J(\varrho^*) = J(\varrho')$. *In particular,* $J(\varepsilon) = J(\varepsilon')$ *provided* ε *is idempotent.*

Proof. By definition, $J(\varrho') \leqq J(\varrho^*)$, $J(\varrho^{*\prime}) \leqq J(\varrho)$ and $\varrho \cdot \varrho^* \cdot \varrho = \varrho$, $\varrho^* \cdot \varrho \cdot \varrho^* = \varrho^*$. Lemma 2.3 implies that $J((\varrho \cdot \varrho^*)') = J(\varrho)$. On the other hand, by (B2), we have $J((\varrho \cdot \varrho^*)') \leqq J(\varrho^{*\prime})$. Thus $J(\varrho) \leqq J(\varrho^{*\prime})$ whence we conclude the equality $J(\varrho) = J(\varrho^{*\prime})$. Similarly, starting with the equality $\varrho^* \cdot \varrho \cdot \varrho^* = \varrho^*$, the equality $J(\varrho^*) = J(\varrho')$ yields.

This lemma shows that in the case of partial right [left] bands the operation "′" can be chosen to be the identity transformation. In what follows, the operation "′" is always assumed to be identical in the case of partial right [left] bands.

Lemma 2.5. *The inverses of the idempotent elements in* J *are also idempotent.*

Proof. Suppose $\varepsilon \in J_\alpha$ is idempotent and ξ is an inverse of ε. By Lemma 2.4, $\varepsilon' \in J_\alpha$ and $\xi, \xi' \in J_\alpha$. Since $\varepsilon \cdot \varepsilon = \varepsilon$ we have $\xi = \xi \cdot \varepsilon \cdot \xi = (\xi \cdot \varepsilon) \cdot (\varepsilon \cdot \xi)$. Here $\xi \cdot \varepsilon$ and $\varepsilon \cdot \xi$ are idempotents in J_α. Thus, by (B5), their product ξ is also idempotent.

Lemma 2.6. *Let* ϱ^* *and* ϱ^{**} *be two inverses of* $\varrho \in J_\alpha$. *Then* $\varrho^* = \varrho^{**} \cdot \varepsilon$ *for some idempotent element* ε *in* J_α.

Proof. Assume that $\varrho' \in J_{\alpha'}$. Then, by Lemma 2.4, $\varrho^*, \varrho^{**} \in J_{\alpha'}$ and

$\varrho^{*\prime}, \varrho^{**\prime} \in J_\alpha$. Moreover, the elements $\varrho^* \cdot \varrho$ and $\varrho^{**} \cdot \varrho$ are idempotents in $J_{\alpha'}$. (B5) implies $(\varrho^* \cdot \varrho) \cdot (\varrho^{**} \cdot \varrho) = \varrho^{**} \cdot \varrho$. On the other hand, we obtain by (B3) that $(\varrho^* \cdot \varrho) \cdot (\varrho^{**} \cdot \varrho) = \varrho^* \cdot (\varrho \cdot \varrho^{**} \cdot \varrho) = \varrho^* \cdot \varrho$ whence we have $\varrho^* \cdot \varrho = \varrho^{**} \cdot \varrho$. Thus $\varrho^* = \varrho^* \cdot \varrho \cdot \varrho^* = \varrho^{**} \cdot (\varrho \cdot \varrho^*)$ where $\varrho \cdot \varrho^*$ is an idempotent element in J_α.

Lemma 2.7. *If $\sigma \in J_\alpha$, $\sigma' \in J_{\alpha'}$ and $\varrho = \sigma \cdot \varepsilon$ for some idempotent ε in $J_{\alpha'}$ then any inverse σ^* of σ is an inverse of ϱ.*

Proof. By Lemmas 2.2 and 2.4 we have $(\sigma \cdot \varepsilon)' \in J_{\alpha'}$. Moreover, $\sigma^* \cdot \sigma$ is an idempotent element in $J_{\alpha'}$ since $\sigma^* \in J_{\alpha'}$ by Lemma 2.4. Thus (B5) ensures $\varepsilon \cdot (\sigma^* \cdot \sigma) = \sigma^* \cdot \sigma$. By applying (B3) we obtain that

$$\varrho \cdot \sigma^* \cdot \varrho = (\sigma \cdot \varepsilon) \cdot \sigma^* \cdot (\sigma \cdot \varepsilon) = \sigma \cdot \big(\varepsilon \cdot (\sigma^* \cdot \sigma)\big) \cdot \varepsilon = \sigma \cdot (\sigma^* \cdot \sigma) \cdot \varepsilon = \sigma \cdot \varepsilon = \varrho$$

and

$$\sigma^* \cdot \varrho \cdot \sigma^* = \sigma^* \cdot (\sigma \cdot \varepsilon) \cdot \sigma^* = (\sigma^* \cdot \sigma) \cdot \varepsilon \cdot (\sigma^* \cdot \sigma) \cdot \sigma^* = (\sigma^* \cdot \sigma) \cdot (\sigma^* \cdot \sigma) \cdot \sigma^* = \sigma^*.$$

The proof is complete.

Define a relation "$\sim$" on J by $\sigma \sim \tau$ if and only if there exists a common inverse of σ and τ in J.

Lemma 2.8. (i) *The relation $\sim$ is an equivalence.*

(ii) *Let $\sigma, \tau \in J$. Then $\sigma \sim \tau$ if and only if $J(\sigma) = J(\tau)$, $J(\sigma') = J(\tau')$ and there exists an idempotent ε in $J(\sigma')$ with $\sigma = \tau \cdot \varepsilon$.*

(iii) *Let $\sigma, \tau \in J$. Then $\sigma \sim \tau$ if and only if the sets of inverses of σ and τ are equal.*

Proof. (ii) and (iii) immediately follow from Lemmas 2.6 and 2.7. Statement (i) is clear by (iii).

Lemma 2.9. *Suppose $\varrho \in J_\alpha$, $\varrho' \in J_{\alpha'}$ with $\alpha' \leqq \alpha$ and $\sigma \in J_\alpha$. Then $\varrho \cdot \sigma = \sigma$ if and only if ϱ is idempotent.*

Proof. If ϱ is idempotent then $\alpha' = \alpha$ by Lemma 2.4. Since $\sigma \cdot \sigma' \in J_\alpha$ is also idempotent we have $\varrho \cdot (\sigma \cdot \sigma') = \sigma \cdot \sigma'$ by (B5). Consequently, (B3) and (B4) imply $\varrho \cdot \sigma = \big(\varrho \cdot (\sigma \cdot \sigma')\big) \cdot \sigma = \sigma \cdot \sigma' \cdot \sigma = \sigma$. Conversely, suppose $\varrho, \sigma \in J_\alpha$, $\varrho' \in J_{\alpha'}$, $\sigma' \in J_{\beta'}$ and $\alpha' \leqq \alpha$. Let

$$\varrho \cdot \sigma = \sigma. \tag{1}$$

Then $(\varrho \cdot \sigma)' = \sigma' \in J_{\beta'}$. Thus $\alpha' = \alpha$ follows from Lemma 2.3. Since $\varrho' \cdot \varrho$ and $\sigma \cdot \sigma'$ are idempotents in J_α property (B5) implies the equality $(\sigma \cdot \sigma') \cdot (\varrho' \cdot \varrho) = \varrho' \cdot \varrho$.

By applying (1) and (B3) we obtain that

$$\varrho = \varrho\cdot\varrho'\cdot\varrho = \varrho\cdot\big((\sigma\cdot\sigma')\cdot(\varrho'\cdot\varrho)\big) = \varrho\cdot\sigma\cdot\big(\sigma'\cdot(\varrho'\cdot\varrho)\big) = \sigma\cdot\big(\sigma'\cdot(\varrho'\cdot\varrho)\big) =$$
$$= (\sigma\cdot\sigma')\cdot(\varrho'\cdot\varrho) = \varrho'\cdot\varrho.$$

Hence ϱ is, in fact, idempotent.

Lemma 2.10. *Let* $\varrho, \sigma\in J$ *with* $J(\varrho)=J(\sigma')$. *Then* $\sigma^*\cdot\varrho=\sigma^{**}\cdot\varrho$ *holds for arbitrary inverses* σ^*, σ^{**} *of* σ.

Proof. By Lemma 2.6, $\sigma^*=\sigma^{**}\cdot\varepsilon$ for some idempotent ε in $J(\sigma')$. Since $\varepsilon\cdot\varrho=\varrho$ by Lemma 2.9 we infer that $\sigma^*\cdot\varrho=\sigma^{**}\cdot\varepsilon\cdot\varrho=\sigma^{**}\cdot\varrho$.

Lemma 2.11. *If* $\varrho\sim\sigma$ *and* τ *is a common inverse of* ϱ *and* σ *then* $\varrho\cdot\tau\cdot\sigma=\sigma$.

Proof. Clearly, $\varrho\cdot\tau\in J(\varrho)=J(\sigma)$ is idempotent. Hence Lemma 2.9 immediately implies the required equality.

In [6] Warne has introduced the concept of a semidirect product of a lower associative semilattice Y of left zero semigroups and an upper associative semilattice Y of right groups. We gave the definition before Theorem 2.1 for the special case of right zero semigroups instead of right groups. We generalize this concept by defining a semidirect product of a partial left band over $\bar{I}\otimes_Y Y$ and a right orthodox partial semigroup over $\bar{J}\otimes_Y Y$.

Suppose we are given a semilattice $\bar{Y}$, a lower associative semilattice $\bar{Y}$ of left zero semigroups $\bar{I}_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$ denoted by $\bar{I}$ and an upper associative semilattice $\bar{Y}$ of right zero semigroups $\bar{J}_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$ denoted by $\bar{J}$. Moreover, let $Y_{\bar{\alpha}}$ be a semilattice with identity $\bar{\alpha}$ for all $\bar{\alpha}$ in $\bar{Y}$. Suppose Y is a semilattice $\bar{Y}$ of semilattices $Y_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$ such that $\bar{Y}$ is a subsemilattice in Y. Let I be a partial left band over $\bar{I}\otimes_Y Y$ and J a right orthodox partial semigroup over $\bar{J}\otimes_Y Y$. Suppose that $\bar{A}, \bar{B}$ is an $(\bar{I}, \bar{J})$-pair.

Assume that $A=\{A_\sigma : \sigma\in J\}$ is a system of transformations of I and $B=\{B_a : a\in I\}$ is a system of transformations of J such that the following are valid:

(C1) if $a\in I_\alpha^{\bar{i}}$, $\varrho\in J_\beta^{\bar{j}}$ and $\varrho'\in J_{\beta'}$ then
  (a) $aA_\varrho\in I_{\alpha_1}^{\bar{i}\bar{A}_{\bar{j}}}$, $\varrho B_a\in J_{\alpha_1}^{\bar{j}\bar{B}_{\bar{i}}}$ and $(\varrho B_a)'\in J_{\alpha_1'}$ where $\alpha_1\leqq\beta$ and $\alpha_1'\leqq\alpha$,
  (b) $\alpha_1=\alpha_1'=\alpha\beta$ provided ϱ is idempotent,
  (c) $\varrho B_a\sim\varrho$ whenever $\alpha=\beta'$,
  (d) if $\alpha<\beta'$ then α_1 is the element of Y for which $(\varepsilon\cdot\varrho')'\in J_{\alpha_1}$ holds provided $\varepsilon\in J_\alpha$ is idempotent;

(C2) if $a\in I_\alpha$, $b\in I_\beta$ with $\alpha\geqq\beta$ and $\varrho\in J$ then
  (a) $\varrho B_{ab}=\varrho B_a B_b$,
  (b) $(a\cdot b)A_\varrho = aA_\varrho\cdot bA_{\varrho B_a}$;

(C3) if $\varrho\in J_\alpha$, $\sigma\in J_\beta$ with $\alpha\leqq\beta$ and $a\in I$ then

(a) $aA_{\varrho\cdot\sigma}=aA_\sigma A_\varrho$,

(b) $(\varrho\cdot\sigma)B_a=\varrho B_{aA_\sigma}\cdot\sigma B_a$.

The pair A, B with these properties is called an *(I, J)-pair over $\bar{A}, \bar{B}$.*

Note that $\alpha_1\in Y_{\bar{\alpha}\bar{\beta}}$ in (C1) (a) provided $\alpha\in Y_{\bar{\alpha}}$ and $\beta\in Y_{\bar{\beta}}$ as $\bar{I}\bar{A}_{\bar{J}}\in\bar{I}_{\bar{\alpha}\bar{\beta}}$ and $\bar{J}\bar{B}_{\bar{I}}\in\bar{J}_{\bar{\alpha}\bar{\beta}}$. The idempotent ε in (C1) (d) can be chosen arbitrarily in J_α for $J((\varepsilon_1\cdot\varrho')')=J((\varepsilon_2\cdot\varrho')')$ by (B6) provided ε_1 and ε_2 are idempotents in J_α. Moreover, one can easily check by (C1) (a) that the right hand sides of the equalities in (C2) (b) and (C3) (b) are defined in I and J, respectively.

Let us define a multiplication on the set $\bigcup\{I_\alpha\times J_\alpha:\alpha\in Y\}$ by

$$(a,\varrho)(b,\sigma)=(a\cdot bA_\varrho,\ \varrho B_b\cdot\sigma). \tag{2}$$

Suppose $a\in I_\alpha$, $\varrho\in J_\alpha$ and $b\in I_\beta$, $\sigma\in J_\beta$. Then, by (C1) (a), we have $bA_\varrho\in I_{\alpha_1}$, $\varrho B_b\in J_{\alpha_1}$ and $(\varrho B_b)'\in J_{\beta_1}$ where $\alpha_1\leqq\alpha$ and $\beta_1\leqq\beta$. Therefore the products $a\cdot bA_\varrho$ and $\varrho B_b\cdot\sigma$ are defined in I and J, respectively, and we have $a\cdot bA_\varrho\in I_{\alpha_1}$, $\varrho B_b\cdot\sigma\in J_{\alpha_1}$ by (B2). Thus (2) is, in fact, a multiplication on the required set. The groupoid thus defined is called a *semidirect product of I and J* and is denoted by $\mathfrak{B}(I, J;\ A, B)$.

Before proving that $\mathfrak{B}(I, J;\ A, B)$ is an orthodox semigroup we verify six lemmas for the (I, J)-pairs.

Lemma 2.12. *If $a\in I_\alpha$ and $\varrho\in J_\beta$, $\varrho'\in J_{\beta'}$ with $\beta'\leqq\alpha$ then $\varrho B_a\sim\varrho$.*

Proof. By property (C3) (b), we have

$$\varrho B_a=(\varrho\cdot(\varrho'\cdot\varrho))B_a=\varrho B_{aA_{\varrho'\cdot\varrho}}\cdot(\varrho'\cdot\varrho)B_a. \tag{3}$$

Since $\varrho'\cdot\varrho\in J_{\beta'}$ is idempotent we have $aA_{\varrho'\cdot\varrho}\in I_{\beta'}$, $(\varrho'\cdot\varrho)B_a\in J_{\beta'}$ and $((\varrho'\cdot\varrho)B_a)'\in J_{\beta'}$ by (C1) (b). Thus $\varrho B_{aA_{\varrho'\cdot\varrho}}\sim\varrho$ follows from (C1) (c). Moreover, owing to (B2), we have $a\cdot aA_{\varrho'\cdot\varrho}\in I_{\beta'}$. Hence (C1) (c) ensures both $(\varrho'\cdot\varrho)B_{a\cdot aA_{\varrho'\cdot\varrho}}\sim\varrho'\cdot\varrho$ and $(\varrho'\cdot\varrho)B_aB_{aA_{\varrho'\cdot\varrho}}\sim(\varrho'\cdot\varrho)B_a$. Making use of Lemma 2.8 (i) we obtain by (C2) (a) that $\varrho'\cdot\varrho\sim(\varrho'\cdot\varrho)B_a$. Then Lemma 2.5 implies $(\varrho'\cdot\varrho)B_a$ to be an idempotent in $J_{\beta'}$. We have seen that $\varrho B_{aA_{\varrho'\cdot\varrho}}\sim\varrho$. Applying Lemma 2.8 (ii) one can easily infer that (3) implies $\varrho B_a\sim\varrho$.

Lemma 2.13. *Let $a\in I_\alpha$ and $\varrho\in J_\beta$, $\varrho'\in J_{\beta'}$. Suppose $\varrho B_a\in J_{\alpha_1}$. If $\beta'\not\leqq\alpha$ then $\alpha_1<\beta$.*

Proof. In the equality (3) which clearly holds by (C3) (b) now we have $aA_{\varrho'\cdot\varrho}\in I_{\alpha\beta'}$, and $(\varrho'\cdot\varrho)B_a\in J_{\alpha\beta'}$ where $\alpha\beta'<\beta'$ as $\beta'\not\leqq\alpha$. By (B2), we obtain from (3) that $J_{\alpha_1}=J(\varrho B_a)=J(\varrho B_{aA_{\varrho'\cdot\varrho}})$. Hence, owing to property (C1) (d), we have

$J_{\alpha_1}=J((\varepsilon\cdot\varrho')')$ where $\varepsilon\in J_{\alpha\beta'}$ is idempotent. (B1), (B2) and Lemma 2.3 ensure that $\alpha_1<\beta$.

Lemma 2.14. *Let $\alpha_1, \alpha_2, \beta$ and $\beta'\in Y$ such that $\alpha_1\leqq\alpha_2\leqq\beta'$. Assume that $\varrho\in J_\beta$ and $\varrho'\in J_{\beta'}$. Then*

(i) *$I(a_1A_\varrho)\leqq I(a_2A_\varrho)$ provided $a_1\in I_{\alpha_1}$ and $a_2\in I_{\alpha_2}$.*

(ii) *$J((\varepsilon_1\cdot\varrho')')\leqq J((\varepsilon_2\cdot\varrho')')$ provided ε_1 and ε_2 are idempotents belonging to J_{α_1} and J_{α_2}, respectively.*

Proof. The product $a_2\cdot a_1$ is defined in I by (B2)* and $a_2\cdot a_1\in I_{\alpha_1}$. Therefore properties (C1) (a), (c) and (d) imply $I((a_2\cdot a_1)A_\varrho)=I(a_1A_\varrho)$. However, we have $(a_2\cdot a_1)A_\varrho=a_2A_\varrho\cdot a_1A_{\varrho B_{a_2}}$ by (C2) (b) whence it follows by (B2)* and (C1) (a) that $I((a_2\cdot a_1)A_\varrho)=I(a_1A_{\varrho B_{a_2}})\leqq I(a_2A_\varrho)$. Thus (i) is verified. Taking into consideration (C1) (c) and (d) statement (ii) is an immediate consequence of (i).

Lemma 2.15. *Let $a\in I_\alpha$, $\varrho\in J_\beta$ and $\varrho'\in J_{\beta'}$ such that $\alpha\leqq\beta'$. If $\varrho B_a\in J_{\alpha_1}$ and ε is an idempotent in J_{α_1} then $\varepsilon\cdot\varrho\sim\varrho B_a$.*

Proof. If $\varrho B_a\in J_{\alpha_1}$ then $\alpha_1\leqq\beta$ and hence $\varepsilon\cdot\varrho$ is defined. Moreover, $aA_\varrho\in I_{\alpha_1}$. Thus $\varepsilon B_{aA_\varrho}\sim\varepsilon$ by (C1) (c). Hence we infer by Lemma 2.5 that $\varepsilon B_{aA_\varrho}$ is idempotent. Then Lemma 2.9 ensures that $\varepsilon B_{aA_\varrho}\cdot\varrho B_a=\varrho B_a$. Therefore, by (C3) (b), we have $(\varepsilon\cdot\varrho)B_a=\varrho B_a$. Properties (C1) (c) and (d) imply that $(\iota\cdot\varrho')'\in J_{\alpha_1}$ provided ι is an idempotent in J_α. By (B6), this implies $(\varepsilon\cdot\varrho)'\in J_\alpha$. Thus, in consequence of (C1) (c), we have $(\varepsilon\cdot\varrho)B_a\sim\varepsilon\cdot\varrho$ which ensures that $\varrho B_a\sim\varepsilon\cdot\varrho$.

Lemma 2.16. *Let ε and η be idempotents in J_α and J_β, respectively, where $\alpha\leqq\beta$. Moreover, let $a\in I_\alpha$. Then ηB_a and $\varepsilon\cdot\eta$ are also idempotents in J_α.*

Proof. If $\alpha=\beta$ then the statement immediately follows from (C1) (c), Lemma 2.5 and (B5). Assume that $\alpha<\beta$ and $(\varepsilon\cdot\eta)'\in J_\gamma$. Then, by (B2) and Lemma 2.3, we have $\gamma<\beta$ whence we can see by utilizing (B6) and (C1) (d) that $\eta B_a\in J_\alpha$ provided $a\in I_\gamma$. Thus Lemma 2.15 implies that $\varepsilon\cdot\eta\sim\eta B_a$. By (C3) (b), we obtain the equality $\eta B_a=(\eta\cdot\eta)B_a=\eta B_{aA_\eta}\cdot\eta B_a$ whence it follows on the one hand, that $\eta B_{aA_\eta}\in J_\alpha$ and therefore $\varepsilon\cdot\eta\sim\eta B_{aA_\eta}$ by Lemma 2.15. On the other hand, Lemma 2.9 ensures ηB_{aA_η} to be idempotent. Thus, since $\eta B_{aA_\eta}\sim\varepsilon\cdot\eta\sim\eta B_a$ we conclude by Lemmas 2.5 and 2.8 (i) that both $\varepsilon\cdot\eta$ and ηB_a are idempotent. Hence $\alpha=\gamma$ as $(\varepsilon\cdot\eta)'\in J_\alpha=J_\gamma$. The proof of the lemma is complete.

Lemma 2.17. *If ϱ_1, ϱ_2 and $\sigma_1, \sigma_2\in J$ with the property that $\varrho_1\sim\varrho_2$ and $\sigma_1\sim\sigma_2$ and, moreover, $\varrho_1\cdot\sigma_1$ is defined then $\varrho_2\cdot\sigma_2$ is also defined and $\varrho_1\cdot\sigma_1\sim\varrho_2\cdot\sigma_2$.*

Proof. One can see immediately by (B2) and Lemma 2.8 (ii) that $\varrho_1\cdot\sigma_1$ is defined if and only if $\varrho_2\cdot\sigma_2$ is defined. Suppose $\varrho_1', \varrho_2'\in J_{\alpha'}$ and $\sigma_1', \sigma_2'\in J_{\beta'}$. Then, again by

Lemma 2.8 (ii), there exist idempotents ε and η in $J_{\alpha'}$ and $J_{\beta'}$, respectively, such that $\varrho_2=\varrho_1\cdot\varepsilon$ and $\sigma_2=\sigma_1\cdot\eta$. Hence $\varrho_2\cdot\sigma_2=(\varrho_1\cdot\varepsilon)\cdot(\sigma_1\cdot\eta)=\varrho_1\cdot(\varepsilon\cdot\sigma_1)\cdot\eta$ by (B3). Assume that $(\varepsilon\cdot\sigma_1)'\in J_\gamma$. Then, by (B2), $\gamma\leqq\beta'$ and we can see by (B6) and (C1) (c) or (d) that $\sigma_1 B_a\in J_{\alpha'}$ provided $a\in I_\gamma$. Thus Lemma 2.15 implies $\sigma_1 B_a\sim\varepsilon\cdot\sigma_1\sim\sim(\varrho_1'\cdot\varrho_1)\cdot\sigma_1$. Hence we obtain by Lemma 2.8 (ii) that $\varepsilon\cdot\sigma_1=(\varrho_1'\cdot\varrho_1)\cdot\sigma_1\cdot\bar\eta$ for some idempotent element $\bar\eta$ in $J((\varepsilon\cdot\sigma_1)')$. Therefore $\varrho_2\cdot\sigma_2=\varrho_1\cdot(\varrho_1'\cdot\varrho_1)\cdot\sigma_1\cdot\bar\eta\cdot\eta==\varrho_1\cdot\sigma_1\cdot(\bar\eta\cdot\eta)$. Here $J((\varrho_1\cdot\sigma_1)')=J(((\varrho_1'\cdot\varrho_1)\cdot\sigma_1)')$. is implied by Lemma 2.2 whence it follows that $J((\varrho_1\cdot\sigma_1)')=J((\varepsilon\cdot\sigma_1)')=J(\bar\eta)$. Since $\bar\eta\cdot\eta$ is an idempotent in $J(\bar\eta)$ by Lemma 2.16 we conclude by Lemma 2.8 (ii) that $\varrho_2\cdot\sigma_2\sim\varrho_1\cdot\sigma_1$ which was to be proved.

Now we can turn to verifying that $\mathfrak{B}(I,J;\ A,B)$ is an orthodox semigroup which is a band $\mathscr{B}(\bar I,\bar J;\ \bar A,\bar B)$ of orthodox semigroups.

Lemma 2.18. *$\mathfrak{B}(I,J;\ A,B)$ is a semigroup.*

Proof. A straightforward calculation shows that the operation defined in (2) is associative. We have to apply properties (C3) (a)—(b), (C1) (a), (B3)*, (B3) and (C2) (a)—(b).

Lemma 2.19. *In the semigroup $\mathfrak{B}(I,J;\ A,B)$ the elements (a,ϱ) and (b,σ) are inverses of each other if and only if ϱ and σ are inverses of each other in J.*

Proof. By definition, $(a,\varrho)(b,\sigma)(a,\varrho)=(a,\varrho)$ and $(b,\sigma)(a,\varrho)(b,\sigma)=(b,\sigma)$ hold if and only if the following four equalities are satisfied in I and J.

$$a\cdot bA_\varrho\cdot aA_\sigma A_{\varrho B_b}=a, \tag{4}$$

$$b\cdot aA_\sigma\cdot bA_\varrho A_{\sigma B_a}=b, \tag{4$'$}$$

$$\varrho B_{b\cdot aA_\sigma}\cdot\sigma B_a\cdot\varrho=\varrho, \tag{5}$$

$$\sigma B_{a\cdot bA_\varrho}\cdot\varrho B_b\cdot\sigma=\sigma. \tag{5$'$}$$

Suppose first that (a,ϱ) and (b,σ) satisfy the equalities (4), (4)′, (5) and (5)′. By (C1) (a), we have $I(a)\geqq I(bA_\varrho)\geqq I(aA_\sigma A_{\varrho B_b})$. The equality (4) implies by (B2)* that $I(aA_\sigma A_{\varrho B_b})=I(a)$. Hence $I(a)=I(bA_\varrho)=I(aA_\sigma A_{\varrho B_b})$. Similarly, by (4)′, we have $I(b)=I(aA_\sigma)=I(bA_\varrho A_{\sigma B_a})$. Since $I(a)$ and $I(b)$ are left zero semigroups the equalities $a\cdot bA_\varrho=a$ and $b\cdot aA_\sigma=b$ are valid. In the equality (5) we have $J(\varrho)==J(\varrho B_b)=J(\xi)$ where $\xi=\varrho B_b\cdot\sigma B_a$. On the one hand, this implies by Lemma 2.13 that $J(\varrho')\leqq J(\sigma)$ and hence, by Lemma 2.12, we conclude that $\varrho B_b\sim\varrho$. On the other hand, applying Lemma 2.9 we obtain that ξ is an idempotent element in $J(\varrho)$ and therefore $\varrho B_b=\xi\cdot\varrho B_b=\varrho B_b\cdot\sigma B_a\cdot\varrho B_b$ since $J(\varrho B_b)=J(\varrho)$ by Lemma 2.8 (ii). In the same way, we can deduce from (5)′ that $\sigma B_a\sim\sigma$ and $\sigma B_a=\sigma B_a\cdot\varrho B_b\cdot\sigma B_a$.

Thus the elements ϱB_b and σB_a are inverses of each other in J. Consequently, Lemma 2.8 (iii) ensures ϱ and σ to be inverses of each other in J which was to be proved.

Conversely, assume that $a\in I_\alpha$, $\varrho\in J_\alpha$, $b\in I_{\alpha'}$, $\sigma\in J_{\alpha'}$ and ϱ and σ are inverses of each other in J. Then we have $\varrho B_b\sim\varrho$ and $\sigma B_a\sim\sigma$ by (C1) (c). This implies by (C1) (a) that $bA_\varrho\in I_\alpha$ and $aA_\sigma\in I_{\alpha'}$ whence, in the same way, we obtain that $aA_\sigma A_{\varrho B_b}\in I_\alpha$ and $bA_\varrho A_{\sigma B_a}\in I_{\alpha'}$. Since both I_α and $I_{\alpha'}$ are left zero semigroups the equalities (4) and (4)' follow. Moreover, we have $b\cdot aA_\sigma=b$ and $a\cdot bA_\varrho=a$ in (5) and (5)', respectively. Then the equalities (5) and (5)' are implied by the relations $\varrho B_b\sim\varrho$ and $\sigma B_a\sim\sigma$ by making use of Lemma 2.11. Thus (a, ϱ) and (b, σ) are inverses of each other.

Lemma 2.20. *$\mathfrak{B}(I, J;\ A, B)$ is an orthodox semigroup with band of idempotents*

$$\mathbf{B} = \{(a, \varepsilon):\ a\in I_\alpha \ \text{ and } \varepsilon \text{ is an idempotent in } J_\alpha \text{ for some } \alpha\in Y\}. \tag{6}$$

Proof. By (B4), every element in J has an inverse which implies by Lemma 2.19 that $\mathfrak{B}(I, J;\ A, B)$ is regular. We show first that the set $\mathbf{B}$ defined in (6) is the set of all idempotents in $\mathfrak{B}(I, J;\ A, B)$. Suppose that (a, ε) is an idempotent element. Then, by definition, we have

$$a\cdot aA_\varepsilon = a \tag{7}$$

and

$$\varepsilon B_a\cdot\varepsilon = \varepsilon. \tag{8}$$

The equality (7) ensures by (B2)* that $I(aA_\varepsilon)=I(a)$. Then $J(\varepsilon B_a)=J(\varepsilon)$. On the other hand, we have $J((\varepsilon B_a)')\leqq J(\varepsilon)$ by (C1) (a). Thus Lemma 2.9 implies by (8) that εB_a is an idempotent in $J(\varepsilon)$. From $J(\varepsilon B_a)=J(\varepsilon)$ we infer by Lemma 2.13 that $J(\varepsilon')\leqq J(\varepsilon)$ which implies by Lemma 2.12 that $\varepsilon B_a\sim\varepsilon$. Hence we obtain by making use of Lemma 2.5 that ε is idempotent. Conversely, if $a\in I_\alpha$ and ε is an idempotent in J_α then we have $aA_\varepsilon\in I_\alpha$ by (C1) (b) and $\varepsilon B_a\sim\varepsilon$ by (C1) (c). The former relation implies (7) as I_α is a left zero semigroup while the latter one, taking into consideration Lemma 2.5, ensures that εB_a is an idempotent in J_α. Hence (8) follows by (B5). Thus we have verified that (a, ε) is an idempotent element in $\mathfrak{B}(I, J;\ A, B)$. Owing to Lemma 2.5, the inverses of the idempotents in J are idempotent. Therefore we obtain by Lemma 2.19 that the inverses of the elements in $\mathbf{B}$ are contained in $\mathbf{B}$. This completes the proof of the fact that $\mathfrak{B}(I, J;\ A, B)$ is orthodox.

Lemma 2.21. *The band of idempotents $\mathbf{B}$ of $\mathfrak{B}(I, J;\ A, B)$ is a semilattice Y of rectangular bands*

$$\mathbf{D}_\alpha = \{(a, \varepsilon):\ a\in I_\alpha,\ \varepsilon \ \text{ is an idempotent in } \ J_\alpha\} \quad (\alpha\in Y).$$

Proof. By applying Lemma 2.19 we see that the set of all inverses of an element (a, ε) in $\mathbf{B}$ with $a\in I_\alpha$ and $\varepsilon\in J_\alpha$ is just $\mathbf{D}_\alpha$, that is, the $\mathcal{D}$-classes in $\mathbf{B}$ are the sets $\mathbf{D}_\alpha (\alpha \in Y)$. If $(a, \varepsilon) \in \mathbf{D}_\alpha$ and $(b, \zeta) \in \mathbf{D}_\beta$ then $(a, \varepsilon)(b, \zeta) = (a \cdot bA_\varepsilon, \varepsilon B_b \cdot \zeta) \in \mathbf{D}_{\alpha\beta}$. For we have $bA_\varepsilon \in I_{\alpha\beta}$ and $\varepsilon B_b \in J_{\alpha\beta}$ by (C1) (b). This implies $a \cdot bA_\varepsilon \in I_{\alpha\beta}$ and $\varepsilon B_b \cdot \zeta \in J_{\alpha\beta}$ by (B2)* and (B2), respectively. Thus we have $\mathbf{D}_\alpha \mathbf{D}_\beta \subseteq \mathbf{D}_{\alpha\beta}$ which was to be proved.

Consider the following subset in the band $\mathbf{B}$:

(9) $$\bar{\mathbf{B}} = \{(a, \varepsilon) \colon a \in I_{\bar{\alpha}} \text{ and } \varepsilon \text{ is an idempotent in } J_{\bar{\alpha}} \text{ for some } \bar{\alpha} \in \bar{Y}\}.$$

Since $\bar{Y}$ is a subsemilattice in Y Lemma 2.21 implies $\bar{\mathbf{B}}$ to be a subband in $\mathbf{B}$ with the property that $\bar{\mathbf{B}}$ is a union of some $\mathscr{D}$-classes of $\mathbf{B}$. For every element $(\bar{i}, \bar{j})$ in the band $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$, let us define a subset in $\mathfrak{B}(I, J; A, B)$ as follows: if $\bar{i} \in \bar{I}_{\bar{\alpha}}$, $\bar{j} \in \bar{J}_{\bar{\alpha}}$ with $\bar{\alpha}$ in $\bar{Y}$ then put

(10) $$\mathbf{F}_{(\bar{i}, \bar{j})} = \{(a, \varrho) \colon a \in I_\alpha^{\bar{i}},\ \varrho \in J_\alpha^{\bar{j}} \text{ for some } \alpha \text{ in } Y_{\bar{\alpha}}\}.$$

Lemma 2.22. *The semigroup $\mathfrak{B}(I, J; A, B)$ is a band $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$ of the orthodox semigroups $\mathbf{F}_{(\bar{i}, \bar{j})}((\bar{i}, \bar{j}) \in \mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B}))$. For every $(\bar{i}, \bar{j})$ in $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$, the greatest $\mathscr{D}$-class of idempotents in $\mathbf{F}_{(\bar{i}, \bar{j})}$ is $\mathbf{F}_{(\bar{i}, \bar{j})} \cap \bar{\mathbf{B}}$.*

Proof. Let $(a, \varrho) \in \mathbf{F}_{(\bar{i}, \bar{j})}$ and $(b, \sigma) \in \mathbf{F}_{(\bar{k}, \bar{l})}$. By definition, we have $(a, \varrho)(b, \sigma) = (a \cdot bA_\varrho, \varrho B_b \cdot \sigma)$ where $bA_\varrho \in I_{\alpha_1}^{\bar{k}\bar{A}_{\bar{j}}}$ and $\varrho B_b \in J_{\alpha_1}^{\bar{j}\bar{B}_{\bar{k}}}$ by (C1) (a). Thus (B2)* and (B2) imply $a \cdot bA_\varrho \in I_{\alpha_1}^{\bar{i} \cdot \bar{k}\bar{A}_{\bar{j}}}$ and $\varrho B_b \cdot \sigma \in J_{\alpha_1}^{\bar{j}\bar{B}_{\bar{k}} \cdot \bar{l}}$, respectively, whence we infer $(a, \varrho)(b, \sigma) \in \mathbf{F}_{(\bar{i}, \bar{j})(\bar{k}, \bar{l})}$. This shows that the equivalence relation on $\mathfrak{B}(I, J; A, B)$ defined by $(a, \varrho) \varkappa (b, \sigma)$ if and only if (a, ϱ), $(b, \sigma) \in \mathbf{F}_{(\bar{i}, \bar{j})}$ for some $(\bar{i}, \bar{j})$ in $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$ is compatible. The second assertion of the lemma immediately follows from Lemma 2.21. Thus the congruence relation $\varkappa$ is subband-parcelling by Proposition 1.1 whence we conclude by Theorem 1.2 that $\mathbf{F}_{(\bar{i}, \bar{j})}$ is an orthodox semigroup for every $(\bar{i}, \bar{j})$ in $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$. The proof is complete.

The following theorem sums up the most important properties of a semidirect product of a partial left band and a right orthodox partial semigroup.

Theorem 2.23. *Let $\bar{Y}$ be a semilattice, $\bar{I}$ a lower associative semilattice $\bar{Y}$ of left zero semigroups $\bar{I}_{\bar{\alpha}}$ $(\bar{\alpha} \in \bar{Y})$ and $\bar{J}$ an upper associative semilattice $\bar{Y}$ of right zero semigroups $\bar{J}_{\bar{\alpha}}$ $(\bar{\alpha} \in \bar{Y})$. For every $\bar{\alpha}$ in $\bar{Y}$, consider a semilattice $Y_{\bar{\alpha}}$ with identity $\bar{\alpha}$. Let Y be a semilattice $\bar{Y}$ of semilattices $Y_{\bar{\alpha}}$ $(\bar{\alpha} \in \bar{Y})$ such that $\bar{Y}$ is a subsemilattice in Y. Let I be a partial left band over $\bar{I} \otimes_{\bar{Y}} Y$ and J a right orthodox partial semigroup over $\bar{J} \otimes_{\bar{Y}} Y$. Suppose $\bar{A}, \bar{B}$ is an $(\bar{I}, \bar{J})$-pair and A, B is an (I, J)-pair over $\bar{A}, \bar{B}$. Then the semidirect product $\mathfrak{B}(I, J; A, B)$ of I and J is an orthodox semigroup with band of idempotents $\mathbf{B}$ defined in (6). The subset $\bar{\mathbf{B}}$ in $\mathbf{B}$ defined in (9) is a subband in $\mathbf{B}$. Moreover, $\mathfrak{B}(I, J; A, B)$ is a band $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$ of the orthodox semigroups $\mathbf{F}_{(\bar{i}, \bar{j})}$ $((\bar{i}, \bar{j}) \in \mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B}))$ defined in (10) where the greatest $\mathscr{D}$-class of idempotents in $\mathbf{F}_{(\bar{i}, \bar{j})}$ is $\mathbf{F}_{(\bar{i}, \bar{j})} \cap \bar{\mathbf{B}}$.*

3. The construction

In this section we introduce the construction which will be applied in the next section to describe the strong subband-parcelling extensions of orthodox semigroups.

Let $\bar{Y}, \bar{I}, \bar{J}, \bar{A}, \bar{B}, Y, I, J, A$ and B have the properties required in Theorem 2.23. Suppose that

(C4) for every α in Y, idempotents i_α and j_α in I_α and J_α, respectively, are distinguished such that $aA_{j_\beta}=i_\beta\cdot a$ and $\sigma B_{i_\beta}=\sigma\cdot j_\beta$ provided $\alpha, \beta\in Y$ with $\alpha\leqq\beta$ and $a\in I_\alpha, \sigma'\in J_\alpha$.

If $\bar{\alpha}\in\bar{Y}$ then denote by $\bar{i}_{\bar{\alpha}}$ the element $\bar{i}$ in $\bar{I}$ for which $i_{\bar{\alpha}}\in I^{\bar{i}}_{\bar{\alpha}}$ holds. Similarly, by $\bar{j}_{\bar{\alpha}}$ we mean the element $\bar{j}$ in $\bar{J}$ with the property that $j_{\bar{\alpha}}\in J^{\bar{j}}_{\bar{\alpha}}$. By (C1) (a), it follows from (C4) that $\bar{i}\bar{A}_{\bar{j}_{\bar{\beta}}}=\bar{i}_{\bar{\beta}}\cdot\bar{i}$ and $\bar{j}\bar{B}_{\bar{i}_{\bar{\beta}}}=\bar{j}\cdot\bar{j}_{\bar{\beta}}$ provided $\bar{\alpha}, \bar{\beta}\in\bar{Y}$ with $\bar{\alpha}\leqq\bar{\beta}$ and $\bar{i}\in\bar{I}_{\bar{\alpha}}, \bar{j}\in\bar{J}_{\bar{\alpha}}$.

Let S be an orthodox semigroup with band of idempotents $E=\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$. The band E is a semilattice $\bar{Y}$ of the rectangular bands $E_{\bar{\alpha}}=\bar{I}_{\bar{\alpha}}\times\bar{J}_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$. For every s in S we denote by $r(s)$ $[l(s)]$ the element $\bar{\alpha}$ in $\bar{Y}$ which has the property that $E_{\bar{\alpha}}\ni ss^*$ $[E_{\bar{\alpha}}\ni s^*s]$ for some inverse s^* of s. If $s\in S$ then there exists a unique inverse s' of S such that $(\bar{i}_{r(s)}, \bar{j}_{r(s)})\mathscr{L}s'\mathscr{R}(\bar{i}_{l(s)}, \bar{j}_{l(s)})$.

For every element s in S, let τ_s be an isomorphism of $r(s)Y$ onto $l(s)Y$. Suppose that $\tau_{s^*}=\tau_s^{-1}$ provided s^* is an inverse of s and $s^*E_{\bar{\alpha}}s\subseteqq E_{\bar{\alpha}\tau_s}$ whenever s^* is an inverse of s and $\bar{\alpha}\in\bar{Y}$ with $\bar{\alpha}\leqq r(s)$. Since Y is a semilattice $\bar{Y}$ of the semilattices $Y_{\bar{\alpha}}$ $(\bar{\alpha}\in\bar{Y})$ with identity $\bar{\alpha}$ such that $\bar{Y}$ is a subsemilattice in Y it is not difficult to verify that $Y_{\bar{\alpha}}\tau_s\subseteqq Y_{\bar{\alpha}\tau_s}$ for every $\bar{\alpha}\in\bar{Y}$ with $\bar{\alpha}\leqq r(s)$.

Let us be given mappings $h_s\colon \bigcup\{I_\alpha\colon \alpha\leqq l(s)\}\to\bigcup\{I_\alpha\colon \alpha\leqq r(s)\}$ and $\chi_s\colon \bigcup\{J_\alpha\colon \alpha\leqq r(s)\}\to\bigcup\{J_\alpha\colon \alpha\leqq l(s)\}$ for each s in S and constants $\gamma_{s,\bar{s}}$ in $J_{l(s\bar{s})}$ for each pair of elements $s, \bar{s}$ in S such that the following conditions are satisfied:

(D1) (a) if $a\in I^{\bar{i}}_\alpha$ with $\bar{i}\in\bar{I}_{\bar{\alpha}}$ and $\alpha\leqq l(s)$ then $ah_s\in I^{\bar{k}}_{\alpha\tau_s^{-1}}$ where $s(\bar{i}, \bar{j}_{\bar{\alpha}})(s(\bar{i}, \bar{j}_{\bar{\alpha}}))'= =(\bar{k}, \bar{j}_{\bar{\alpha}\tau_s^{-1}})$,

(b) if $\sigma\in J^{\bar{j}}_\alpha$ with $\bar{j}\in\bar{J}_{\bar{\alpha}}$ and $\alpha\leqq r(s)$, $\sigma'\in J_{\alpha'}$ then $\sigma\chi_s\in J^{\bar{l}}_{\alpha\tau_s}$ with $((\bar{i}_{\bar{\alpha}}, \bar{j})s)'(\bar{i}_{\bar{\alpha}}, \bar{j})s=(\bar{i}_{\bar{\alpha}\tau_s}, \bar{l})$ and $(\sigma\chi_s)'\in J_{\alpha'\tau_s}$;

(D2) (a) if $a\in I_\alpha$, $b\in I_\beta$ with $l(s)\geqq\alpha\geqq\beta$ then $ah_s\cdot bh_s=(a\cdot bA_{j_{\alpha\tau_s^{-1}}\chi_s})h_s$,

(b) if $\varrho\in J_\alpha$, $\varrho'\in J_{\alpha'}$ and $\sigma\in J_\beta$ with $\alpha'\leqq\beta\leqq r(s)$ then $\varrho\chi_s\cdot\sigma\chi_s=(\varrho B_{i_{\beta\tau_s}h_s}\cdot\sigma)\chi_s$;

(D3) if $a\in I_\alpha$ with $\alpha\leqq l(s)$ and $\sigma\in J_\beta$ with $\beta\leqq r(s)$ then

(a) $aA_{\sigma\chi_s}h_s=i_{\beta\tau_s}h_s\cdot ah_sA_\sigma$,

(b) $\sigma B_{ah_s}\chi_s=\sigma\chi_sB_a\cdot j_{\alpha\tau_s^{-1}}\chi_s$;

(D4) (a) if $a\in I_\alpha$ with $\alpha\leqq l(s\bar{s})$ then $ah_{\bar{s}}h_s=c\cdot aA_{\gamma_{s,\bar{s}}}h_{s\bar{s}}$ for some c in $I_{r(s\bar{s})}$,

(b) if $\varrho\in J_\alpha$ with $\alpha\leqq r(s\bar{s})$ then $(j_{\alpha\tau_{s\bar{s}}}\cdot\gamma_{s,\bar{s}})\cdot\varrho\chi_s\chi_{\bar{s}}=\varrho B_c\chi_{s\bar{s}}\cdot\gamma_{s,\bar{s}}$ for some c in $I_{r(s\bar{s})}$;

(D5) if $\alpha \leqq l(s\bar{s})$ then $(j_\alpha \cdot \gamma_{s,\bar{s}})' \in J_{\alpha\tau_{s\bar{s}}^{-1}\tau_s\tau_{\bar{s}}}$;

(D6) if $s, \bar{s}, \bar{\bar{s}} \in S$ then $\gamma_{s,\bar{s}\bar{\bar{s}}} \cdot \gamma_{\bar{s},\bar{\bar{s}}} \sim \gamma_{s\bar{s},\bar{\bar{s}}} \cdot (j_{r(s\bar{s}\bar{\bar{s}})\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}})\chi_{\bar{\bar{s}}}$;

(D7) (a) if $e \in E$, $\alpha \leqq r(e) = l(e)$ and $a \in I_\alpha$ then $ah_e = c \cdot aA_{\gamma_{e,e}}$ for some c in $I_{r(e)}$,
(b) if $e \in E$, $\alpha \leqq r(e) = l(e)$ and $\varrho \in J_\alpha$ then $\varrho\chi_e = \varrho B_c \cdot \gamma_{e,e}$ for some c in $I_{r(e)}$;

(D8) γ_{s,s^*} is idempotent whenever $s \in S$ is idempotent or s is not an inverse of itself, and s^* is an inverse of s in S.

If h, χ and γ fulfil these conditions then we call them an *(S, I, J)-triple.*

Note that $\bar{\alpha}\tau_s^{-1}$ and $\bar{\alpha}\tau_s$ are defined in (D1) (a) and (b), respectively, as $\alpha \in Y_{\bar{\alpha}}$, $\alpha \leqq l(s)$ imply $\bar{\alpha} \leqq l(s)$ and, similarly, $\alpha \in Y_{\bar{\alpha}}$, $\alpha \leqq r(s)$ imply $\bar{\alpha} \leqq r(s)$. It is not difficult to check that it follows from conditions (C1) (a) and (D1) that both sides of the equalities in (D2), (D3), (D4) and (D7) are defined. Similarly, (D5) ensures that both sides of (D6) are also defined.

Before introducing the construction by means of which we describe the strong subband-parcelling extensions of orthodox semigroups we prove some lemmas concerning (S, I, J)-triples which make the computations easier.

Lemma 3.1. *If ε is an idempotent in J_α and $s \in S$ with $\alpha \leqq r(s)$ then $\varepsilon\chi_s$ is idempotent.*

Proof. By definition, $i_{\alpha\tau_s}h_s \in I_\alpha$ and hence, by (C1) (c), we have $\varepsilon B_{i_{\alpha\tau_s}h_s} \sim \varepsilon$. Therefore Lemma 2.5 ensures $\varepsilon B_{i_{\alpha\tau_s}h_s}$ to be also an idempotent in J_α, that is, $\varepsilon B_{i_{\alpha\tau_s}h_s} \cdot \varepsilon = \varepsilon$ by (B5). Consequently, (D2) (b) implies that $\varepsilon\chi_s \cdot \varepsilon\chi_s = \varepsilon\chi_s$ which was to be proved.

Lemma 3.2. *If $\varrho \in J_\alpha$ and ϱ^* is an inverse of ϱ contained in $J_{\alpha'}$ and, moreover, $s \in S$ with $\alpha \leqq r(s)$ then $\alpha' \leqq r(s)$ and $\varrho\chi_s$ and $\varrho^*\chi_s$ are inverses of each other in J.*

Proof. If $\alpha \in Y_{\bar{\alpha}}$ then $\alpha' \in Y_{\bar{\alpha}}$ follows from Lemma 2.4 and (B1). Since $r(s) \in \bar{Y}$ the relation $\alpha \leqq r(s)$ implies $\bar{\alpha} \leqq r(s)$. Thus $\alpha' \leqq \bar{\alpha} \leqq r(s)$ whence we obtain that both $\varrho\chi_s$ and $\varrho^*\chi_s$ are defined. By definition, $i_{\alpha\tau_s}h_s \in I_\alpha$ and $i_{\alpha'\tau_s}h_s \in I_{\alpha'}$. Moreover, we have $i_{\alpha\tau_s}h_s A_{\varrho^*} \in I_{\alpha'}$ by (C1) (a), (c) and Lemma 2.8 (ii). Since $I_{\alpha'}$ is a left zero semigroup the equality $i_{\alpha'\tau_s}h_s \cdot i_{\alpha\tau_s}h_s A_{\varrho^*} = i_{\alpha'\tau_s}h_s$ holds. On the other hand, we have $\varrho B_{i_{\alpha'\tau_s}h_s} \sim \varrho$ and $\varrho^* B_{i_{\alpha\tau_s}h_s} \sim \varrho^*$ by (C1) (c). Therefore we can see by applying the equality (D2) (b) twice and making use of properties (C3) (b), (C2) (a) and Lemma 2.11 that

$$\varrho\chi_s \cdot \varrho^*\chi_s \cdot \varrho\chi_s = (\varrho B_{i_{\alpha'\tau_s}h_s} \cdot \varrho^*)\chi_s \cdot \varrho\chi_s = ((\varrho B_{i_{\alpha'\tau_s}h_s} \cdot \varrho^*) B_{i_{\alpha\tau_s}h_s} \cdot \varrho)\chi_s =$$

$$= (\varrho B_{i_{\alpha'\tau_s}h_s \cdot i_{\alpha\tau_s}h_s A_{\varrho^*}} \cdot \varrho^* B_{i_{\alpha\tau_s}h_s} \cdot \varrho)\chi_s = (\varrho B_{i_{\alpha'\tau_s}h_s} \cdot \varrho^* B_{i_{\alpha\tau_s}h_s} \cdot \varrho)\chi_s = \varrho\chi_s.$$

Dually, one obtains $\varrho^*\chi_s \cdot \varrho\chi_s \cdot \varrho^*\chi_s = \varrho^*\chi_s$ which completes the proof.

Lemma 3.3. *If* $s, \bar{s} \in S$ *then* $\gamma'_{s,\bar{s}} \in J_{l(s\bar{s})}$.

Proof. By definition, $\gamma_{s,\bar{s}} \in J_{l(s\bar{s})}$ and therefore $j_{l(s\bar{s})} \cdot \gamma_{s,\bar{s}} = \gamma_{s,\bar{s}}$ by Lemma 2.9. Thus we have $\gamma'_{s,\bar{s}} \in J_{l(s,\bar{s})}$ by (D5) as $l(s\bar{s})\tau_{s\bar{s}}^{-1}\tau_s\tau_{\bar{s}} = l(s\bar{s})$.

Lemma 3.4. *Let* $a \in I_\alpha$, $\sigma \in J_\alpha$ *and* $s \in S$ *with* $\alpha \leqq l(s)$. *Then we have* $ah_s \cdot bA_\sigma h_s = (a \cdot bA_\sigma)h_s$ *for every* b *in* I.

Proof. By (C1) (a), we have $bA_\sigma \in I_{\alpha_1}$ where $\alpha_1 \leqq \alpha$. Thus both sides are defined. Taking into consideration (D2) (a), it suffices to verify that $bA_\sigma = bA_\sigma A_{j_{\alpha\tau_s^{-1}}\chi_s}$. In consequence of Lemma 3.1, $j_{\alpha\tau_s^{-1}}\chi_s$ is an idempotent in J_α. Hence $j_{\alpha\tau_s^{-1}}\chi_s \cdot \sigma = \sigma$ by Lemma 2.9 and therefore (C3) (a) implies the equality required.

The following lemma is dual to Lemma 3.4.

Lemma 3.5. *Let* $a \in I_\alpha$, $\sigma \in J_\alpha$ *and* $s \in S$ *with* $\alpha \leqq r(s)$. *Then we have* $\varrho B_a\chi_s \cdot \sigma\chi_s = (\varrho B_a \cdot \sigma)\chi_s$ *for every* ϱ *in* J.

Proof. Suppose that $\alpha \in Y_{\bar{\alpha}}$ and $\varrho \in J_\beta$ with $\beta \in Y_{\bar{\beta}}$. By (C1) (a), we have $\varrho B_a \in J_{\alpha_1}$ and $(\varrho B_a)' \in J_{\alpha_1'}$ where $\alpha_1 \leqq \beta$ and $\alpha_1' \leqq \alpha$. The remark after the definition of an (I, J)-pair ensures that $\alpha_1, \alpha_1' \in Y_{\bar{\alpha}\bar{\beta}}$. Since $\bar{\alpha} \leqq r(s)$ follows from $\alpha \leqq r(s)$ we have $\alpha_1, \alpha_1' \leqq r(s)$. Thus both sides of the equality are defined. It suffices to prove by (D2) (b) that $\varrho B_a B_{i_{\alpha\tau_s}h_s} = \varrho B_a$. Here $i_{\alpha\tau_s}h_s \in I_\alpha$. Since I_α is a left zero semigroup we have $a \cdot i_{\alpha\tau_s}h_s = a$ whence the equality to be proved follows immediately by making use of (C2) (a).

Lemma 3.6. *If* s *and* s^* *are inverses of each other in* S *then* $\gamma_{ss^*,s}$ *and* γ_{s,s^*s} *are idempotent elements in* $J_{l(s)}$.

Proof. By definition and Lemma 3.3, we obtain that both $\gamma_{ss^*,s}$ and $\gamma'_{ss^*,s}$ belong to $J_{l(s)}$. Moreover, by (D6), we have

$$\gamma_{ss^*,s} \cdot (j_{r(s)} \cdot \gamma_{ss^*,ss^*})\chi_s \sim \gamma_{ss^*,s} \cdot \gamma_{ss^*,s}.$$

Here $j_{r(s)} \cdot \gamma_{ss^*,ss^*} = \gamma_{ss^*,ss^*}$ is an idempotent in $J_{r(s)}$ by (D8) and hence, by Lemma 3.1, $(j_{r(s)} \cdot \gamma_{ss^*,ss^*})\chi_s \in J_{l(s)}$ is also idempotent. Thus Lemma 2.8 (ii) implies that $\gamma_{ss^*,s} \sim \gamma_{ss^*,s} \cdot \gamma_{ss^*,s}$, that is, $\gamma_{ss^*,s} \cdot \gamma_{ss^*,s} = \gamma_{ss^*,s} \cdot \varepsilon$ for some idempotent ε in $J_{l(s)}$. Multiplying this equality on the right by the idempotent element $\gamma'_{ss^*,s} \cdot \gamma_{ss^*,s}$ in $J_{l(s)}$ and applying (B5) we obtain that $\gamma_{ss^*,s} \cdot \gamma_{ss^*,s} = \gamma_{ss^*,s}$, that is, $\gamma_{ss^*,s}$ is, indeed, idempotent. A similar argument shows that γ_{s,s^*s} is also idemdontent.

Let us define a groupoid $\mathbf{S} = \mathfrak{S}(S, I, J; h, \chi, \gamma)$ in the following way. The underlying set of $\mathbf{S}$ is

$$\mathbf{S} = \{(a, s, \sigma) \colon s \in S, \quad a \in I_\alpha^1 \text{ and } \sigma \in J_{\alpha\tau_s}^1 \text{ where } \alpha \in Y_{r(s)}, \quad ss' = (1, j_{r(s)}) \text{ and } s's = (1_{l(s)}, j)\}$$

and the operation is defined by

$$(11)\qquad (a, s, \sigma)(\bar{a}, \bar{s}, \bar{\sigma}) = \left(a \cdot \bar{a}A_\sigma h_s, s\bar{s}, (j_{\beta\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}}) \cdot \sigma B_{\bar{a}}\chi_{\bar{s}} \cdot \bar{\sigma}\right)$$

where β is the element of Y with the property that $\bar{a}A_\sigma h_s \in I_\beta$.

We have to show that the products occurring in (11) are defined in I and J and, moreover, that the set S is closed under the multiplication defined in (11). Suppose that $a \in I^i_\alpha$, $\bar{a} \in I^{\bar{i}}_{\bar{\alpha}}$, $\sigma \in J^j_{\alpha\tau_s}$ and $\bar{\sigma} \in J^{\bar{j}}_{\bar{\alpha}\tau_{\bar{s}}}$ where $ss' = (i, \bar{j}_{r(s)})$, $\bar{s}\bar{s}' = (\bar{i}, \bar{j}_{r(\bar{s})})$, $s's = (\bar{i}_{l(s)}, j)$ and $\bar{s}'\bar{s} = (\bar{i}_{l(\bar{s})}, \bar{j})$. By (C1) (a) we have $\bar{a}A_\sigma \in I^{\bar{i}A_j}_{\bar{\alpha}_1}$ and $\sigma B_{\bar{a}} \in J^{jB_{\bar{i}}}_{\bar{\alpha}_1}$. As we have seen, $\bar{\alpha}_1 \in Y_{l(s)r(\bar{s})}$ since $\bar{\alpha} \in Y_{r(\bar{s})}$ and $\alpha\tau_s \in Y_{l(s)}$. Thus $\bar{\alpha}_1 \leqq l(s)$ and $\bar{\alpha}_1 \leqq r(\bar{s})$. Therefore $\bar{a}A_\sigma h_s$ and $\sigma B_{\bar{a}}\chi_{\bar{s}}$ are defined and $\bar{a}A_\sigma h_s \in I^{\bar{k}}_{\bar{\alpha}_1\tau_s^{-1}}$, $\sigma B_{\bar{a}}\chi_{\bar{s}} \in J^{\bar{l}}_{\bar{\alpha}_1\tau_{\bar{s}}}$ where, by (D1) (a), we have $s(\bar{i}A_j, \bar{j}_{l(s)r(\bar{s})})(s(\bar{i}A_j, \bar{j}_{l(s)r(\bar{s})}))' = (\bar{k}, \bar{j}_{(l(s)r(\bar{s}))\tau_s^{-1}})$ and $((\bar{i}_{l(s)r(\bar{s})}, j\bar{B}_{\bar{i}})\bar{s})'(\bar{i}_{l(s)r(\bar{s})}, j\bar{B}_{\bar{i}})\bar{s} = (\bar{i}_{(l(s)r(\bar{s}))\tau_{\bar{s}}}, \bar{l})$. Here $(l(s)r(\bar{s}))\tau_s^{-1} = r(s\bar{s})$ and $(l(s)r(\bar{s}))\tau_{\bar{s}} = l(s\bar{s})$. Hence $\beta = \bar{\alpha}_1\tau_s^{-1} \in Y_{r(s\bar{s})}$ and $\beta\tau_{s\bar{s}} \in Y_{l(s\bar{s})}$. Property (D5) implies that $(j_{\beta\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}})' \in J_{\beta\tau_s\tau_{\bar{s}}}$ where $\beta\tau_s\tau_{\bar{s}} = \bar{\alpha}_1\tau_{\bar{s}}$. Lemma 3.2 implies that $(\sigma B_{\bar{a}})'\chi_{\bar{s}}$ is also defined and it is an inverse of $(\sigma B_{\bar{a}})\chi_{\bar{s}}$. Thus we obtain by (D1) (b) and Lemma 2.4 that $(\sigma B_{\bar{a}}\chi_{\bar{s}})' \in J_{\bar{\alpha}_1'\tau_{\bar{s}}}$ provided $(\sigma B_{\bar{a}})' \in J_{\bar{\alpha}_1'}$. Since $\bar{\alpha}_1 \leqq \alpha\tau_s$ and $\bar{\alpha}_1' \leqq \bar{\alpha}$ we have $\beta = \bar{\alpha}_1\tau_s^{-1} \leqq \alpha$, $\bar{\alpha}_1\tau_{\bar{s}} = \beta\tau_s\tau_{\bar{s}}$ and $\bar{\alpha}_1'\tau_{\bar{s}} \leqq \bar{\alpha}\tau_{\bar{s}}$. Thus we see that the products $a \cdot \bar{a}A_\sigma h_s$ and $(j_{\beta\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}}) \cdot \sigma B_{\bar{a}}\chi_{\bar{s}} \cdot \bar{\sigma}$ in (11) are defined and $a \cdot \bar{a}A_\sigma h_s \in^{i \cdot \bar{k}}_\beta$, $(j_{\beta\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}}) \cdot \sigma B_{\bar{a}}\chi_{\bar{s}} \cdot \bar{\sigma} \in J^{\bar{x} \cdot \bar{l} \cdot \bar{j}}_{\beta\tau_{s\bar{s}}}$ where $\beta \in Y_{r(s\bar{s})}$ and $j_{\beta\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}} \in J^{\bar{x}}_{\beta\tau_{s\bar{s}}}$. Here $\bar{x} \in \bar{J}_{l(s\bar{s})}$ and $\bar{l} \in \bar{J}_{l(s\bar{s})}$ whence we infer that $\bar{x} \cdot \bar{l} = \bar{l}$. All we have to verify is that $s\bar{s}(s\bar{s})' = (i \cdot \bar{k}, \bar{j}_{r(s\bar{s})})$ and $(s\bar{s})'s\bar{s} = (\bar{i}_{l(s\bar{s})}, \bar{l} \cdot \bar{j})$. We will show the first equality. The second one follows dually. In the band $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$ we have

$$(s's)(\bar{s}\bar{s}') = (\bar{i}_{l(s)}, j)(\bar{i}, \bar{j}_{r(\bar{s})}) = (\bar{i}_{l(s)} \cdot \bar{i}\bar{A}_j, j\bar{B}_{\bar{i}} \cdot \bar{j}_{r(\bar{s})}).$$

Taking into consideration the remark after (C4) we obtain that

$$(s's)(\bar{s}\bar{s}') = (\bar{i}\bar{A}_j\bar{A}_{\bar{j}_{l(s)}}, j\bar{B}_{\bar{i}}\bar{B}_{\bar{i}_{r(\bar{s})}}).$$

Since $j \in \bar{J}_{l(s)}$ and $\bar{i} \in \bar{I}_{r(\bar{s})}$ where $\bar{J}_{l(s)}$ is a right zero semigroup and $\bar{I}_{r(\bar{s})}$ a left zero semigroup we infer by (W1) (a) and (W2) (a) that

$$(s's)(\bar{s}\bar{s}') = (\bar{i}\bar{A}_j, j\bar{B}_{\bar{i}})\mathscr{R}(\bar{i}\bar{A}_j, \bar{j}_{l(s)r(\bar{s})}).$$

Similarly, by applying the remark after (C4) and the fact that $\bar{I}_{r(\bar{s})}$ is a left zero semigroup one sees that

$$(i, \bar{j}_{r(s)})(\bar{k}, \bar{j}_{r(s\bar{s})}) = (i \cdot \bar{k}\bar{A}_{\bar{j}_{r(s)}}, \bar{j}_{r(s)}\bar{B}_{\bar{k}} \cdot \bar{j}_{r(s\bar{s})}) =$$

$$= (i \cdot \bar{i}_{r(s)} \cdot \bar{k}, \bar{j}_{r(s)}\bar{B}_{\bar{k}} \cdot \bar{j}_{r(s\bar{s})}) = (i \cdot \bar{k}, \bar{j}_{r(s)}\bar{B}_{\bar{k}} \cdot \bar{j}_{r(s\bar{s})}) = (i \cdot \bar{k}, \bar{j}_{r(s\bar{s})}).$$

In the last step we have utilized that $J_{r(s)}\bar{B}_{\bar{k}}\in\bar{J}_{r(s\bar{s})}$ and $\bar{J}_{r(s\bar{s})}$ is a right zero semigroup. Now we can easily check that

$$(i\cdot\bar{k}, j_{r(s\bar{s})}) = (ss')(s(\bar{l}\bar{A}_j, j_{l(s)r(\bar{s})}))(s(\bar{l}\bar{A}_j, j_{l(s)r(\bar{s})}))' =$$
$$= s(\bar{l}\bar{A}_j, j_{l(s)r(\bar{s})})(s(\bar{l}\bar{A}_j, j_{l(s)r(\bar{s})}))'\mathscr{R}s(\bar{l}\bar{A}_j, j_{l(s)r(\bar{s})})s'\mathscr{R}s(s's)(\bar{s}\bar{s}')s' = s\bar{s}\bar{s}'s'\mathscr{R}s\bar{s}(s\bar{s})'$$

hold in the semigroup S. On the other hand, $s\bar{s}(s\bar{s})'\mathscr{L}(i\cdot\bar{k}, j_{r(s\bar{s})})$ by the definition of $(s\bar{s})'$. This completes the proof of the fact that $(s\bar{s})(s\bar{s})'=(i\cdot\bar{k}, j_{r(s\bar{s})})$.

By applying the technique used in Lemmas 3.1, 3.2, 3.4, 3.5 and 3.6 one can prove the following lemmas.

Lemma 3.7. *$\mathbf{S}$ is a semigroup.*

Lemma 3.8. *Let $(a, s, \sigma)\in\mathbf{S}$ with $\sigma\in J_\alpha$, $\sigma'\in J_{\alpha'}$. Then $(a', s', \sigma^*\chi_{s'}\cdot\varepsilon)\in\mathbf{S}$ and the equality*

(12) $$(a, s, \sigma)(a', s', \sigma^*\chi_{s'}\cdot(j_{\alpha\tau_s^{-1}}\cdot\gamma'_{s,s'})\cdot\varepsilon)(a, s, \sigma) = (a, s, \sigma)$$

holds for every inverse σ^* of σ, for every a' in $I_\alpha^{\bar{1}l(s)}$, and for every idempotent ε in $J_{\alpha\tau_s^{-1}}^{\bar{1}r(s)}$. Consequently, $\mathbf{S}$ is regular.

Since the proofs need rather long and complicated calculations we left them to the reader.

Observe that the relation $\mathfrak{C}$ defined on $\mathbf{S}$ by

(13) $$(a, s, \varrho)\mathfrak{C}(\bar{a}, \bar{s}, \bar{\varrho}) \quad \text{if and only if} \quad s=\bar{s}$$

is a congruence relation. The idempotent $\mathfrak{C}$-classes are

$$\mathbf{C}_e = \{(a, e, \varrho) : (a, e, \varrho)\in\mathbf{S}\}, \quad (e\in E).$$

Lemma 3.9. *The mapping $\varphi\colon \bigcup\{\mathbf{C}_e : e\in E\}\to\mathfrak{B}(I, J; A, B)$ defined by $(a, e, \varrho)\varphi=(a, \varrho)$ is an onto isomorphism.*

Proof. φ is one-to-one and onto since if $(a, \varrho)\in\mathfrak{B}(I, J; A, B)$ with $a\in I_\alpha^{\bar{1}}$, $\varrho\in J_\alpha^{\bar{j}}$ then $e=(\bar{1}, \bar{j})$ is the unique idempotent element in S such that $(a, e, \varrho)\in\mathbf{S}$ and, obviously, we have $(a, e, \varrho)\varphi=(a, \varrho)$. A straightforward calculation shows that $((a, e, \sigma)(\bar{a}, \bar{e}, \bar{\sigma}))\varphi=(a, e, \sigma)\varphi$, $(\bar{a}, \bar{e}, \bar{\sigma})\varphi$, that is, φ is an isomorphism.

Lemma 2.20 shows that $\mathfrak{B}(I, J; A, B)$ is an orthodox semigroup. Hence $\bigcup\{\mathbf{C}_e : e\in E\}$ is also an orthodox semigroup which implies that the idempotents in $\mathbf{S}$ form a subsemigroup. Since $\mathbf{S}$ is regular by Lemma 3.8 $\mathbf{S}$ is also orthodox. Moreover, Lemma 2.22 ensures by the proof of Proposition 1.1 that the congruence relation $\mathfrak{C}$ defined in (13) is $(\bar{\mathbf{C}}, \mathfrak{C}')$-parcelling where, by using the notations of (6), (9) and Lemma 2.22 we define $\bar{\mathbf{C}}=\bar{\mathbf{B}}\varphi^{-1}$ and $\mathfrak{C}'$ to be the congruence relation on the band of idempotents of $\mathbf{S}$ corresponding under φ^{-1} to the congruence $\mathscr{D}\cap\varkappa|\mathbf{B}$ on $\mathbf{B}$.

Theorem 3.10. *Let $\bar{Y}, \bar{I}, \bar{J}, \bar{A}, \bar{B}, Y, I, J, A$ and B have the properties required in Theorem 2.23. Suppose that* (C4) *holds for A and B. Let S be an orthodox semigroup with band of idempotents $\mathscr{B}(\bar{I}, \bar{J}; \bar{A}, \bar{B})$. Moreover, let h, χ, γ be an (S, I, J)-triple. Then $\mathbf{S} = \mathfrak{S}(S, I, J; h, \chi, \gamma)$ is an orthodox semigroup and the relation $\mathfrak{C}$ defined in* (13) *is a strong subband-parcelling congruence on $\mathbf{S}$ such that the factor semigroup $\mathbf{S}/\mathfrak{C}$ is isomorphic to S.*

Proof. The reasoning carried out before stating the theorem shows that all we have to prove is that $\mathfrak{C}$ is strong. We verify that, for every s in S, there exists $(a, s, \sigma) \in \mathbf{S}$ with $a \in I_{r(s)}$ and $\sigma \in J_{l(s)}$ idempotent such that (a, s, σ) is $\mathscr{L}$- and $\mathscr{R}$-equivalent to idempotents in $\bar{\mathbf{C}}$. By Lemma 1.3 and Theorem 1.4, we can restrict ourselves to elements s with $(\bar{\imath}_{r(s)}, \bar{\jmath}_{r(s)}) \mathscr{R} s \mathscr{L} (\bar{\imath}_{l(s)}, \bar{\jmath}_{l(s)})$. These are precisely those elements for which $s = s''$ holds. Suppose that s fulfils this property and $a \in I_{r(s)}^{\bar{\imath}_{r(s)}}$. Then $(a, s, j_{l(s)}) \in \mathbf{S}$ and we have seen in Lemma 3.8 that $(a', s', j_{l(s)} \chi_{s'} \cdot (j_{r(s)} \cdot \gamma'_{s,s'})) \in \mathbf{S}$ and (16) holds with $\sigma = j_{l(s)}$ and $\varepsilon = \gamma_{s,s'} \cdot \gamma'_{s,s'}$ for any $a' \in I_{l(s)}^{\bar{\imath}_{l(s)}}$. Since $j_{l(s)} \chi_{s'}$ is an idempotent in $J_{r(s)}$ by Lemma 3.1 and the idempotents in $J_{r(s)}$ form a right zero semigroup we have $j_{l(s)} \chi_{s'} \cdot (j_{r(s)} \cdot \gamma'_{s,s'}) = \gamma'_{s,s'}$. Thus

$$(a, s, j_{l(s)})(a', s', \gamma'_{s,s'})(a, s, j_{l(s)}) = (a, s, j_{l(s)})$$

where one can easily check that $(a, s, j_{l(s)})(a', s', \gamma'_{s,s'}) = (a, ss', \gamma_{s,s'} \cdot \gamma'_{s,s'}) \in \bar{\mathbf{C}}$. Since $s = s''$ a similar argument shows that $(a, s, \zeta) \in \mathbf{S}$ with $\zeta = \gamma_{s,s'} \chi_s \cdot \gamma'_{s',s} \cdot j_{l(s)}$ and

$$(a', s', \gamma'_{s,s'})(a, s, \zeta)(a', s', \gamma'_{s,s'}) = (a', s', \gamma'_{s,s'}).$$

Here $\gamma_{s,s'} \chi_s = \gamma_{ss',s} \cdot (j_{r(s)} \cdot \gamma_{s,s'}) \chi_s \sim \gamma_{s,s's} \cdot \gamma_{s',s}$ by Lemma 3.6 and (D6). Thus we have $\zeta \sim \gamma_{s,s's} \cdot (\gamma_{s',s} \cdot \gamma'_{s',s}) \cdot j_{l(s)}$ by Lemma 2.17. Since $\gamma_{s,s's} \in J_{l(s)}$ is idempotent by Lemma 3.6 and, clearly, both $(\gamma_{s',s} \cdot \gamma'_{s',s})$ and $j_{l(s)}$ are idempotents in $J_{l(s)}$ we obtain that $\zeta \sim j_{l(s)}$. This implies that ζ is also idempotent and, since $\zeta \cdot j_{l(s)} = \zeta$ we infer that $\zeta = j_{l(s)}$.
Since $(a', s', \gamma'_{s,s'})(a, s, j_{l(s)}) = (a', s's, j_{l(s)}) \in \bar{\mathbf{C}}$ we conclude that $(a, s, j_{l(s)})$ and $(a', s', \gamma'_{s,s'})$ are inverses of each other in $\mathbf{S}$ and therefore $(a, ss', \gamma_{s,s'} \cdot \gamma'_{s,s'}) \mathscr{R} (a, s, j_{l(s)}) \mathscr{L}$ $\mathscr{L} (a', s's, j_{l(s)})$. Thus we have proved the theorem.

4. The main result

In this section we prove that any strong subband-parcelling extension of an orthodox semigroup S is isomorphic to some semigroup $\mathfrak{S}(S, I, J; h, \chi, \gamma)$.

An orthodox semigroup T is said to be a *strong subband-parcelling extension* of the orthodox semigroup S if S is isomorphic to $T/\varkappa$ for some strong subband-parcelling congruence $\varkappa$ on T.

Before drawing up our main result we verify two lemmas which make the proof of the theorem easier.

Lemma 4.1. *Let T be an orthodox semigroup whose band of idempotents is a semilattice Y of rectangular bands E_α ($\alpha \in Y$). Let t and u be elements in T such that, for some inverses t' and u' of t and u, respectively, we have $E(tt') \leqq E(u'u)$. If $x \in E(tt')$ and $z \in E(uxu')$ then $zuxt = zut$.*

Proof. By assumption, we have $E(z) = E(uxu') = E(uxtt'u') = E(utt'u')$. Moreover, $tt'u'utt' = tt'$ as $E(tt') \leqq E(u'u)$. Thus we obtain that

$$zuxt = zux(tt')t = zux(tt')u'u(tt')t = z(uxtt'u')(utt'u')ut = zutt'u'ut = zut.$$

Lemma 4.2. *Let T be an orthodox semigroup with band of idempotents B. Suppose $\bar{B}, \delta$ is an associated pair in B. Let $\varkappa$ be a strong $(\bar{B}, \delta)$-parcelling congruence on T. Let $T/\varkappa$ be denoted by S. Then there exists a cross-section $\{u_s : s \in S, u_s\varkappa = s\}$ of the $\varkappa$-classes contained in $S_{\bar{B}}$ such that $u_e \in \bar{B}$ whenever e is idempotent and, furthermore, u_s and u_{s^*} are inverses of each other and $u_s u_{s^*} = u_{ss^*}$ provided s is idempotent or s is not an inverse of itself and s^* is an inverse of s in S.*

Proof. Let the band of idempotents in S be E which is a semilattice Y of rectangular bands E_α ($\alpha \in Y$). Let us choose and fix an element e_α in E_α for every α in Y. Moreover, choose and fix an element i_α of $\bar{B}$ in each $\varkappa$-class e_α. If $e\mathcal{R}e_\alpha$ in S then we have $i_\alpha j \varkappa j$ and $i_\alpha j \mathcal{R} i_\alpha$ for every element j of the $\varkappa$-class e contained in $\bar{B}$. For $e_\alpha e = e$ implies $i_\alpha j \varkappa j$, the equality $i_\alpha(i_\alpha j) = i_\alpha j$ trivially holds and $(i_\alpha j) i_\alpha = i_\alpha$ follows from the fact that $\varkappa | \bar{B} = \delta | \bar{B} \subseteq \mathcal{D}$ whence we infer $(i_\alpha j) i_\alpha \mathcal{D} i_\alpha$ by $((i_\alpha j) i_\alpha)\varkappa = e_\alpha$. Thus we have seen that every $\varkappa$-class e with $e\mathcal{R}e_\alpha$ contains an element $i \in \bar{B}$ such that $i\mathcal{R}i_\alpha$. The dual assertion holds for the $\varkappa$-classes e with $e\mathcal{L}e_\alpha$. Now let us choose and fix an element $i_e^* \in \bar{B}$ in every $\varkappa$-class e with $e\mathcal{R}e_\alpha$ or $e\mathcal{L}e_\alpha$ such that $i_e^* \mathcal{R} i_\alpha$ and $i_e^* \mathcal{L} i_\alpha$, respectively. In particular, it is clear that $i_{e_\alpha}^* = i_\alpha$. If $f\mathcal{D}e_\alpha$ then there exist uniquely determined elements e_1 and e_2 in E such that $e_1 \mathcal{R} f \mathcal{L} e_2$ and $e_1 \mathcal{L} e_\alpha \mathcal{R} e_2$. Then $e_1 e_2 = f$. Define i_f^* to be $i_{e_1}^* \cdot i_{e_2}^*$. Since, for each α in Y, $\bar{B} \cap \{j \in B : j\varkappa = f$ for some $f \in E$ with $f\mathcal{D}e_\alpha\}$ is a rectangular band the set $\{i_f^* : f\mathcal{D}e_\alpha\}$ forms a subband in it and therefore $\{i_f^* : f\mathcal{D}e_\alpha\}$ is also a rectangular band. Let us define u_e to be i_e^* for every e in E.

Now let s be a non-idempotent element in S such that $e_\alpha \mathcal{R} s \mathcal{L} e_\beta$ for some α, β in Y. Let s' be the inverse of s with $e_\alpha \mathcal{L} s' \mathcal{R} e_\beta$. Proposition 1.5 ensures the existence of elements t, t' in $S_{\bar{B}}$ which are inverses of each other and $t\varkappa = s$, $t'\varkappa = s'$. If $s = s'$ define $u_s = i_\alpha t i_\beta$. Now consider the case when $s \neq s'$. Since both tt' and $t't$ belong to $\bar{B}$ and $(tt')\varkappa = e_\alpha$, $(t't)\varkappa = e_\beta$ the elements $u_s = i_\alpha t i_\beta$ and $u_{s'} = i_\beta t' i_\alpha$ are also in $S_{\bar{B}}$, they are inverses of each other and $u_s u_{s'} = i_\alpha$, $u_{s'} u_s = i_\beta$. Clearly, we have $u_s \varkappa = s$ and $u_{s'}\varkappa = s'$. Thus we have defined u_s for those s in S for which $e_\alpha \mathcal{R} s \mathcal{L} e_\beta$ for some $\alpha, \beta \in Y$. Finally, let $\bar{s}$ be any element in S. Assume that $e\mathcal{R}\bar{s}\mathcal{L}f$ where $e \in E_\alpha$ and $f \in E_\beta$. Then $s = e_\alpha \bar{s} e_\beta$ satisfies $e_\alpha \mathcal{R} s \mathcal{L} e_\beta$. Define $u_{\bar{s}}$ to be $u_{\bar{s}} = i_e^* u_s i_f^*$. Obviously, we have $u_{\bar{s}}\varkappa = \bar{s}$. The idempotents i_e^* $(e \in E)$ are chosen such that the definition of $u_{\bar{s}}$ is independent of the choice of e and f. If $\bar{s}$ and $\bar{s}'$ are inverses of each other in S such that s is not an inverse of itself and $\bar{s}\bar{s}' = e \in E_\alpha$, $\bar{s}'\bar{s} = f \in E_\beta$ then $s = e_\alpha \bar{s} e_\beta \neq$ $\neq e_\beta \bar{s}' e_\alpha = s'$. Thus $u_{\bar{s}}$ and $u_{\bar{s}'}$ are inverses of each other in S_B and we have $u_{\bar{s}} u_{\bar{s}'} = i_e^*$, $u_{\bar{s}'} u_{\bar{s}} = i_f^*$. Thus the required conditions are fufilled by the cross-section $\{u_s : s \in S, u_s \varkappa = s\}$ which completes the proof of the lemma.

Now we turn to the main theorem of the paper.

Theorem 4.3. *Suppose T is an orthodox semigroup and $\varkappa$ is a strong subband-parcelling congruence on T. Denote $T/\varkappa$ by S. Then there exist $\bar{Y}, \bar{I}, \bar{J}, \bar{A}, \bar{B}, Y, I, J, A, B$ satisfying the conditions of Theorem* 2.23 *and* (C4) *and there exists an (S, I, J)-triple h, χ, γ such that T is isomorphic to $\mathfrak{S}(S, I, J; h, \chi, \gamma)$.*

Proof. Assume that $\varkappa$ is a $(\bar{B}_0, \delta)$-parcelling congruence on T where $\bar{B}_0, \delta$ is an associated pair on the band of idempotents B_0 in T. Denote the semilattice $B_0/\mathcal{D}$ by Y. One can easily see by Theorem 1.2 that $\bar{Y} = \bar{B}_0/\mathcal{D}$ is a subsemilattice in Y and Y is a semilattice $\bar{Y}$ of semilattices $Y_{\bar{\alpha}}$ with identities $\bar{\alpha}$ in $\bar{Y}$. Since $\varkappa$ is a strong $(\bar{B}_0, \delta)$-parcelling congruence the band of idempotents E in S is isomorphic to $\bar{B}_0/\delta$. Thus E is a semilattice $\bar{Y}$ of rectangular bands $E_{\bar{\alpha}} (\bar{\alpha} \in \bar{Y})$. Let us choose and fix an element $e_{\bar{\alpha}}$ in every $\mathcal{D}$-class $E_{\bar{\alpha}}$. Moreover, for every α in Y, select an element i_α in the $\mathcal{D}$-class α such that $i_\alpha \varkappa = e_{\bar{\alpha}}$ provided $\alpha \in Y_{\bar{\alpha}}$. This can be done by Lemma 1.3. If $\bar{\alpha} \in \bar{Y}$ then let $\bar{I}_{\bar{\alpha}}$ and $\bar{J}_{\bar{\alpha}}$ stand for the $\mathcal{L}$-class and $\mathcal{R}$-class, respectively, in $E_{\bar{\alpha}}$ containing $e_{\bar{\alpha}}$. If $\alpha \in Y$ then denote by I_α the $\mathcal{L}$-class in B_0 containing i_α and let J_α be the set of all elements σ in T for which $\sigma\varkappa$ is idempotent and $i_\alpha \mathcal{R} \sigma$. Suppose the transformation "$'$" on $J = \bigcup \{J_\alpha : \alpha \in Y\}$ assigns an inverse to each element. Clearly, such a transformation exists on J. Define a partial operation "$\cdot$" on $\bar{I} = \bigcup \{\bar{I}_{\bar{\alpha}} : \bar{\alpha} \in \bar{Y}\}$ as follows: if $a \in \bar{I}_{\bar{\alpha}}$, $b \in \bar{I}_{\bar{\beta}}$ then $a \cdot b$ is defined if and only if $\bar{\alpha} \geqq \bar{\beta}$ and if this is the case then $a \cdot b$ means their product in E. It is clear that $a \cdot b \mathcal{L} e_{\bar{\beta}}$, that is, $a \cdot b \in \bar{I}_{\bar{\beta}}$. With respect to this partial operation $\bar{I}$ is a lower associative semilattice $\bar{Y}$ of the left zero semigroups $\bar{I}_{\bar{\alpha}}$ $(\bar{\alpha} \in \bar{Y})$. Analogously, one can define a multiplication on the set $I = \bigcup \{I_\alpha : \alpha \in Y\}$ with respect to which I becomes a lower associative semilattice Y of the left zero semigroups I_α $(\alpha \in Y)$. For every $\bar{\imath}$ in $\bar{I}$, denote by $I^{\bar{\imath}}$ the set $\{i \in I : i\varkappa = \bar{\imath}\}$. The elements i_α $(\alpha \in Y)$ are chosen such that I is a disjoint union of the subsets $I^{\bar{\imath}}$ $(\bar{\imath} \in \bar{I})$. Let $I_\alpha^{\bar{\imath}} = I^{\bar{\imath}} \cap I_\alpha$ provided $\alpha \in Y_{\bar{\alpha}}$ and $\bar{\imath} \in \bar{I}_{\bar{\alpha}}$. By Lemma 1.3, these subsets are non-void and, since $\varkappa$ is a congruence, one can immediately see that $I = \bigcup \{I_\alpha^{\bar{\imath}} : (\bar{\imath}, \alpha) \in \bar{I} \otimes_{\bar{Y}} Y\}$ and I is a partial left band over $\bar{I} \otimes_{\bar{Y}} Y$. Let us define $\bar{J}$ dually to $\bar{I}$. Obviously, $\bar{J}$ is an upper associative semilattice $\bar{Y}$ of the right zero semigroups $\bar{J}_{\bar{\alpha}}$ $(\bar{\alpha} \in Y)$. Finally, define a partial operation on J in the following way: if $\varrho \in J_\alpha$, $\varrho' \in J_{\alpha'}$, $\sigma \in J_\beta$, $\sigma' \in J_{\beta'}$ and $\alpha' \leqq \beta$ then let $\varrho \cdot \sigma$ mean the product of ϱ

and σ in T and, in the opposite case, let $\varrho\cdot\sigma$ be undefined. Clearly, we have $\varrho\cdot\sigma\mathcal{R}i_\alpha$ and hence $\varrho\cdot\sigma\in J_\alpha$. Denote by $J^{\bar{J}}$ the set $\{j\in J: j\varkappa=\bar{j}\}$. Similarly to the case of I one can see that $J=\bigcup\{J^{\bar{j}}:\bar{j}\in\bar{J}\}$. Moreover, if $(\bar{j},\alpha)\in J\otimes_Y Y$ then $J^{\bar{j}}_\alpha=J^{\bar{j}}\cap J_\alpha$ is non-void and, since $\varkappa$ is a congruence we have $\varrho\cdot\sigma\in J^{\bar{j}\cdot\bar{k}}$ provided $\varrho\in J^{\bar{j}}$, $\sigma\in J^{\bar{k}}$ and their product is defined in J. Since $(\varrho\cdot\sigma)'\sim\sigma'\varrho'$ in T where $\sim$ is used to mean the least inverse semigroup congruence on T and $(\sigma'\sigma)\sigma'\varrho'=\sigma'\varrho'$ we have $J((\varrho\cdot\sigma)')\leqq J(\sigma'\cdot\sigma)=J(\sigma')$. This proves (B2). Properties (B1), (B3), (B4) and (B5) trivially hold in J. As far as (B6) is concerned, if ϱ and σ are idempotents in J_α and J_β, respectively, and $\tau\in J_\gamma$, $\tau'\in J_{\gamma'}$ with $\alpha\leqq\gamma$, $\beta\leqq\gamma'$ then $\tau'\varrho\tau\mathcal{D}\sigma$ holds in B_0 if and only if $\tau\sigma\tau'\mathcal{D}\varrho$. Thus we have shown that J is a right orthodox semigroup over $\bar{J}\otimes_Y Y$.

For every element s in S, let s' stand for the inverse of s satisfying $e_{r(s)}\mathcal{L}s'\mathcal{R}e_{l(s)}$.

Let us choose a cross-section $\{u_s: s\in S,\ u_s\varkappa=s\}$ of the $\varkappa$-classes possessing the properties required in Lemma 4.2. The proof of Lemma 4.2 ensures that this cross-section can be chosen such that $u_{e_{\bar{\alpha}}}=i_{\bar{\alpha}}$ for every $\bar{\alpha}$ in $\bar{Y}$. Denote by u'_s the inverse of u_s fulfilling $i_{r(s)}\mathcal{L}u'_s\mathcal{R}i_{l(s)}$. Clearly, we have $u'_s\varkappa=s$.

Let τ_s be the isomorphism of $r(s)Y$ onto $l(s)Y$ which corresponds the element β with $u'_si_\alpha u_s\mathcal{D}i_\beta$ to every α in $r(s)Y$. Clearly, τ_e is the identity automorphism of $r(e)Y=$ $=l(e)Y$ provided $e\in E$. Moreover, if s and s^* are inverses of each other in S then $\tau_{s^*}=\tau_s^{-1}$. If $\bar{\alpha}\in\bar{Y}$ and $\bar{\alpha}\leqq r(s)$ then $s'E_{\bar{\alpha}}s\subseteq E_{\bar{\alpha}\tau_s}$.

Now we verify that every element t in T is uniquely represented in the form $t=au_s\sigma$ where $s=t\varkappa\in S$, $a\in I^{ss'}_\alpha$ and $\sigma\in J^{s's}_{\alpha\tau_s}$ for some α in $Y_{r(s)}$. Let $t\in T$ and denote $t\varkappa$ by s. Let a be an element in I which is $\mathcal{R}$-related to t. Obviously, such an a exists and is unique. Suppose that $a\in I_\alpha$ and $\alpha\in Y_{\bar{\alpha}}$. Then $s=t\varkappa\mathcal{R}a\varkappa\mathcal{L}e_{\bar{\alpha}}$, that is, $\bar{\alpha}=r(s)$ and $a\in I^{ss'}_\alpha$. On the other hand, we have $\sigma=i_{\alpha\tau_s}u'_st\mathcal{R}i_{\alpha\tau_s}u'_sau_s\mathcal{R}i_{\alpha\tau_s}$ as $u'_sau_s\in\alpha\tau_s$. Since $\sigma\varkappa=e_{r(s)\tau_s}s's=s's$ we obtain that $\sigma\in J^{s's}_{\alpha\tau_s}$. We can easily see that $au_s\sigma=$ $=au_si_{\alpha\tau_s}u'_st=a(u_si_{\alpha\tau_s}u'_s)at=at=t$ as $a\mathcal{R}t$ and $a,\ u_si_{\alpha\tau_s}u'_s\in\alpha$. As far as the uniqueness of this representation is concerned, observe that if $s\in S$ and $a\in I^{ss'}_\alpha$, $\sigma\in J^{s's}_{\alpha\tau_s}$ with $\alpha\in Y_{r(s)}$ then $(au_s\sigma)\varkappa=ss'ss's=s$ and $au_s\sigma\mathcal{R}au_s\sigma\sigma'u'_sa=a$. Therefore if an element t is represented in both of the forms $au_s\sigma$ and $\bar{a}u_{\bar{s}}\bar{\sigma}$ then $s=\bar{s}$ and $a=\bar{a}$. From the latter equality $\alpha=\bar{\alpha}$ follows where $a\in I_\alpha$ and $\bar{a}\in I_{\bar{\alpha}}$. Multiplying the equality $au_s\sigma=au_s\bar{\sigma}$ by $i_{\alpha\tau_s}u'_s$ on the left we conclude $\sigma=\bar{\sigma}$. Put

$$\mathrm{S}=\{(a,s,\sigma):\ s\in S,\ a\in I^{ss'}_\alpha \text{ and } \sigma\in J^{s's}_{\alpha\tau_s} \text{ for some } \alpha \text{ in } Y_{r(s)}\}.$$

We have shown in this paragraph that the mapping $\Phi\colon T\to\mathrm{S}$ which assigns (a,s,σ) to $t=au_s\sigma$ is one-to-one and onto. In the sequel we give an (S,I,J)-triple h,χ,γ such that Φ becomes an isomorphism of T onto $\mathfrak{S}(S,I,J;h,\chi,\gamma)$.

First we deal with the idempotent $\varkappa$-classes. If $t\varkappa=e\in E$ then $t=au_e\sigma=$ $=ai_\alpha u_ei_\alpha\sigma=ai_\alpha\sigma=a\sigma$ where $a\in I_\alpha$. For $\alpha\leqq r(e)=l(e)$ and u_e is an idempotent element in T with $u_e \mathscr{D} i_{l(e)}$. If $a \in I_\alpha^{\bar{l}}$ and $\sigma \in J_\alpha^{\bar{j}}$ then $(a\sigma)\varkappa = \bar{l}\bar{j}$. Therefore every element t in T with the property that $t\varkappa$ is idempotent is uniquely representable in the form $a\sigma$ where $a \in I_\alpha$ and $\sigma \in J_\alpha$ for some α in Y. Now we will use this representation for the elements of the idempotent $\varkappa$-classes. Let $\bar{l} \in \bar{I}_{\bar{\alpha}}, \bar{j} \in \bar{J}_{\bar{\beta}}$ and $\alpha \in Y_{\bar{\alpha}}$, $\beta \in Y_{\bar{\beta}}$. Assume that $a \in I_\alpha^{\bar{l}}$ and $\sigma \in J_\beta^{\bar{j}}, \sigma' \in J_{\beta'}$. Then $(\sigma a)\varkappa = \bar{j}\bar{l}$ and thus σa is uniquely written in the form $a_1\sigma_1$ where $a_1 \in I_{\alpha_1}^{(\bar{j}\bar{l})(\bar{j}\bar{l})'}$ and $\sigma_1 \in J_{\alpha_1}^{(\bar{j}\bar{l})'(\bar{j}\bar{l})}$. Let us denote a_1 by aA_σ and σ_1 by σB_a and, moreover, $(\bar{j}\bar{l})(\bar{j}\bar{l})'$ by $\bar{l}\bar{A}_{\bar{j}}$ and $(\bar{j}\bar{l})'(\bar{j}\bar{l})$ by $\bar{j}\bar{B}_{\bar{l}}$. Clearly, $\bar{l}\bar{A}_{\bar{j}} \in \bar{I}_{\bar{\alpha}\bar{\beta}}$ and $\bar{j}\bar{B}_{\bar{l}} \in \bar{J}_{\bar{\alpha}\bar{\beta}}$. Observe that $\sigma a\sigma' \in \alpha_1$. Moreover, $(\sigma B_a)' \in J_{\alpha_1'}$ if and only if $a\sigma'\sigma a \in \alpha_1'$. Thus $\alpha_1 \leqq \beta$ and $\alpha_1' \leqq \alpha$ immediately follow. If σ is idempotent then $\alpha_1 = \alpha_1' = \alpha\beta$. Suppose $\alpha = \beta'$. Then $\sigma a\sigma' = \sigma\sigma'$ whence $\alpha_1 = \beta$. Furthermore, $\sigma B_a = i_{\alpha_1}\sigma a = i_\beta \sigma a = \sigma a = \sigma i_\alpha \sim \sigma$. However, if $\alpha < \beta'$ then we clearly have $(i_\alpha\sigma')' \in J_{\alpha_1}$ as $\sigma i_\alpha \sigma' \mathscr{D} \sigma a\sigma'$. Thus we have verified that the families $A = \{A_\sigma : \sigma \in J\}$ and $B = \{B_a : a \in I\}$ of transformations of I and J, respectively, satisfy (C1).

Now let $a \in I_\alpha$, $b \in I_\beta$ with $\alpha \geqq \beta$ and $\sigma \in J$. By definition, we have

$$\sigma ab = \sigma(a \cdot b) = (a \cdot b)A_\sigma \cdot \sigma B_{a \cdot b}, \tag{14}$$

where $(a \cdot b)A_\sigma \in I_{\alpha_1}$ and $\sigma B_{a \cdot b} \in J_{\alpha_1}$. On the other hand,

$$\sigma ab = (\sigma a)b = (aA_\sigma \cdot \sigma B_a)b = aA_\sigma(\sigma B_a \cdot b) = (aA_\sigma \cdot bA_{\sigma B_a}) \cdot \sigma B_a B_b, \tag{14'}$$

where the product $aA_\sigma \cdot bA_{\sigma B_a}$ is defined in I as it was noted after the definition of an (I, J)-pair. Since σab is uniquely representable in the form $a_0\sigma_0$ with $a_0 \in I_{\alpha_0}$, $\sigma_0 \in J_{\alpha_0}$ for some $\alpha_0 \in Y$ we infer that (14) and (14′) imply (C2) to be valid. Dually, one can prove that (C3) also holds. Since $\varkappa$ is a congruence relation (C2) and (C3) show by the definition of $\bar{A}$ and $\bar{B}$ that (W1) and (W2) are fulfilled by $\bar{A}, \bar{B}$. This completes the proof of the facts that $\bar{A}, \bar{B}$ is an $(\bar{I}, \bar{J})$-pair and A, B is an (I, J)-pair over $\bar{A}, \bar{B}$. (C4) trivially holds for A, B with $i_\alpha = j_\alpha$, $\alpha \in Y$.

Let s be an element in S. Define the mapping $h_s \colon \cup \{I_\alpha : \alpha \leqq l(s)\} \to \cup \{I_\alpha : \alpha \leqq r(s)\}$ by

$$ah_s = u_s a u_s' i_{\alpha\tau_s^{-1}}$$

provided $a \in I_\alpha^{\bar{l}}$ and $\alpha \leqq l(s)$. Similarly, let $\chi_s \colon \cup \{J_\alpha : \alpha \leqq r(s)\} \to \cup \{J_\alpha : \alpha \leqq l(s)\}$ be the mapping for which

$$\sigma\chi_s = i_{\beta\tau_s} u_s' \sigma u_s$$

whenever $\sigma \in J_\beta^{\bar{j}}$ with $\beta \leqq r(s)$. Clearly, $ah_s \in I_{\alpha\tau_s^{-1}}^{s\bar{l}(s\bar{l})'}$ and $\sigma\chi_s \in J_{\beta\tau_s}^{(\bar{j}s)'\bar{j}s}$ as $u_s a u_s' \in \alpha\tau_s^{-1}$, $(ah_s)\varkappa = s\bar{l}s' e_{\bar{\alpha}\tau_s^{-1}} = s\bar{l}(s\bar{l})'$ and $u_s' \sigma u_s \mathscr{R} u_s' i_\beta u_s \in \beta\tau_s$, $(\sigma\chi_s)\varkappa = e_{\bar{\beta}\tau_s} s'\bar{j}s = (\bar{j}s)'\bar{j}s$, respectively. Here $\alpha \in Y_{\bar{\alpha}}$ and $\beta \in Y_{\bar{\beta}}$. It is obvious that $\bar{\alpha} \leqq l(s)$ and $\bar{\beta} \leqq r(s)$. Therefore (D1) (a) and (b) are satisfied by h and χ, respectively. After the properties (D2)—(D8) we have noted that both sides of the respective equalities are defined. Thus we must check only that the equalities are valid.

In order to prove (D2) (a) assume that $a\in I_\alpha$, $b\in I_\beta$ and $l(s)\geqq\alpha\geqq\beta$. By definition, we have

$$ah_s\cdot bh_s\mathcal{R}u_sa(u_s' i_{\alpha\tau_s^{-1}}u_s)bu_s' = u_sa(i_\alpha u_s' i_{\alpha\tau_s^{-1}}u_s)bu_s' =$$

$$=u_sa(i_{\alpha\tau_s^{-1}}\chi_s)bu_s' = u_s(a\cdot bA_{i_{\alpha\tau_s^{-1}}\chi_s})\cdot i_{\alpha\tau_s^{-1}}\chi_sB_bu_s'.$$

Here $bA_{i_{\alpha\tau_s^{-1}}\chi_s}\mathcal{D}i_{\alpha\tau_s^{-1}}\chi_sB_b$ in the band B_0 whence we obtain that

$$a\cdot bA_{i_{\alpha\tau_s^{-1}}\chi_s}\mathcal{R}a\cdot bA_{i_{\alpha\tau_s^{-1}}\chi_s}\cdot i_{\alpha\tau_s^{-1}}\chi_sB_b.$$

Thus

$$ah_s\cdot bh_s\mathcal{R}u_sa\cdot bA_{i_{\alpha\tau_s^{-1}}\chi_s}u_s'\mathcal{R}(a\cdot bA_{i_{\alpha\tau_s^{-1}}\chi_s})h_s$$

and, since both $ah_s\cdot bh_s$ and $(a\cdot bA_{i_{\alpha\tau_s^{-1}}\chi_s})h_s$ belong to $I_{\beta\tau_s^{-1}}$, they are equal. For (D2) (b), suppose that $\varrho\in J_\alpha$, $\varrho'\in J_{\alpha'}$, $\sigma\in J_\beta$ and $\alpha'\leqq\beta\leqq r(s)$. Then we have

$$\varrho\chi_s\cdot\sigma\chi_s = i_{\alpha\tau_s}u_s'\varrho u_si_{\beta\tau_s}u_s'\sigma u_s = i_{\alpha\tau_s}u_s'\varrho(u_si_{\beta\tau_s}u_s'i_\beta)\sigma u_s =$$

$$= i_{\alpha\tau_s}u_s'\varrho(i_{\beta\tau_s}h_s)\sigma u_s = i_{\alpha\tau_s}u_s'(i_{\beta\tau_s}h_sA_\varrho)(\varrho B_{i_{\beta\tau_s}h_s})\sigma u_s.$$

Here $i_{\beta\tau_s}h_sA_\varrho\mathcal{L}i_\alpha\mathcal{R}\varrho B_{i_{\beta\tau_s}h_s}$ whence it follows by Lemma 4.1 that

$$\varrho\chi_s\cdot\sigma\chi_s = i_{\alpha\tau_s}u_s'(\varrho B_{i_{\beta\tau_s}h_s}\cdot\sigma)u_s = (\varrho B_{i_{\beta\tau_s}h_s}\cdot\sigma)\chi_s.$$

Now we check that property (D3) is satisfied. Assume that $a\in I_\alpha$ with $\alpha\leqq l(s)$ and $\sigma\in J_\beta$ with $\beta\leqq r(s)$. Suppose that $aA_{\sigma\chi_s}\in I_{\alpha_1}$ and $\sigma B_{ah_s}\in J_{\beta_1}$. Then, by definition, we have

$$aA_{\sigma\chi_s}h_s = u_s(aA_{\sigma\chi_s})u_s'i_{\alpha_1\tau_s^{-1}} = u_s(\sigma\chi_s)a(\sigma\chi_s)^*i_{\alpha_1}u_s'i_{\alpha_1\tau_s^{-1}} = \tag{15}$$

$$= u_si_{\beta\tau_s}u_s'\sigma u_sau_s'\sigma^*u_si_{\alpha_1}u_s'i_{\alpha_1\tau_s^{-1}},$$

where σ^* and $(\sigma\chi_s)^*$ are arbitrary inverses of σ and $\sigma\chi_s$, respectively. Since $u_si_{\beta\tau_s}u_{s'}\in\beta$ and $\sigma\in J_\beta$ we have

$$u_si_{\beta\tau_s}u_s'\sigma u_s = u_si_{\beta\tau_s}u_s'i_\beta\sigma u_s = i_{\beta\tau_s}h_s\cdot\sigma u_s. \tag{16}$$

On the other hand, $aA_{\sigma\chi_s}\in\alpha_1$ which implies $\sigma u_sau_s'\sigma^*\in\alpha_1\tau_s^{-1}$. Thus

$$(\sigma u_sau_s'\sigma^*)(u_si_{\alpha_1}u_s')i_{\alpha_1\tau_s^{-1}} = (\sigma u_sau_s'\sigma^*)i_{\alpha_1\tau_s^{-1}}. \tag{17}$$

Since $\sigma u_sau_s'\sigma^*\mathcal{R}\sigma\cdot ah_s\cdot\sigma^*$ the equality

$$(\sigma u_sau_s'\sigma^*)i_{\alpha_1\tau_s^{-1}} = \sigma\cdot ah_s\cdot\sigma^*i_{\alpha_1\tau_s^{-1}} = ah_sA_\sigma \tag{18}$$

yields. The equality in (D3) (a) follows from (15) by applying (16), (17) and (18). Moreover, observe that this equality ensures $\beta_1=\alpha_1\tau_s^{-1}$. As far as the dual property

(D3) (b) is concerned, one can see by definition that

$$\sigma B_{ah_s}\chi_s = i_{\beta_1\tau_s}u_s'(\sigma B_{ah_s})u_s = i_{\beta_1\tau_s}u_s' i_{\beta_1}\sigma(ah_s)u_s.$$

Since $\sigma B_{ah_s}\in J_{\beta_1}$ we have $\sigma(ah_s)\sigma^*\in\beta_1$. Therefore Lemma 4.1 implies

$$\sigma B_{ah_s}\chi_s = i_{\beta_1\tau_s}u_s'\sigma(ah_s)u_s.$$

Furthermore, $u_s'\sigma(ah_s)\sigma^*u_s\in\beta_1\tau_s=\alpha_1$ whence we obtain that $\alpha_1\leqq\beta\tau_s$ and hence

$$\sigma B_{ah_s}\chi_s = i_{\alpha_1}i_{\beta\tau_s}u_s'\sigma(ah_s)u_s = i_{\alpha_1}i_{\beta\tau_s}u_s'\,\sigma u_s a u_s' i_{\alpha\tau_s^{-1}}u_s = i_{\alpha_1}\cdot\sigma\chi_s\cdot a u_s'\, i_{\alpha\tau_s^{-1}}u_s.$$

Utilizing that both a and $u_s' i_{\alpha\tau_s^{-1}}u_s$ are contained in α it follows that

$$\sigma B_{ah_s}\chi_s = i_{\alpha_1}\cdot\sigma\chi_s\cdot a i_\alpha u_s'\, i_{\alpha\tau_s^{-1}}u_s = i_{\alpha_1}\cdot\sigma\chi_s\cdot a\cdot i_{\alpha\tau_s^{-1}}\chi_s = \sigma\chi_s B_a\cdot i_{\alpha\tau_s^{-1}}\chi_s$$

which was to be proved.

Now we define constants $c_{s,\bar s}$ and $\gamma_{s,\bar s}$ for each pair of elements $s,\bar s$ in S. Since $(u_s u_{\bar s})\varkappa = s\bar s = u_{s\bar s}\varkappa$ there exist uniquely determined elements $c_{s,\bar s}$ in $I_\alpha^{s\bar s(s\bar s)'}$ and $\gamma_{s,\bar s}$ in $J_{\alpha\tau_{s\bar s}}^{(s\bar s)'s\bar s}$ such that $\alpha\in Y_{r(s\bar s)}$ and

$$u_s u_{\bar s} = c_{s,\bar s}u_{s\bar s}\gamma_{s,\bar s}.$$

Here $u_s u_{\bar s}\in S_{B_0}$ whence we infer that $\alpha=r(s\bar s)$ and $\alpha\tau_{s\bar s}=l(s\bar s)$. This implies $\gamma_{s,\bar s}= =u_{s\bar s}'u_s u_{\bar s}$ as $u_{s\bar s}'\mathscr{R}i_{l(s\bar s)}$. Thus $\gamma_{s,\bar s}\in S_{B_0}$, that is, $\gamma_{s,\bar s}$ is also contained in $J_{l(s\bar s)}$. If $\alpha\leqq l(s\bar s)$ then

$$\gamma_{s,\bar s}' i_\alpha\gamma_{s,\bar s}\sim u_{\bar s}'\,u_s'\,u_{s\bar s}i_\alpha u_{s\bar s}' u_s u_{\bar s}\in\alpha\tau_{s\bar s}^{-1}\tau_s\tau_{\bar s}$$

which proves (D5). If $s,\bar s,\bar{\bar s}\in S$ then we have

$$\xi = \gamma_{s\bar s,\bar{\bar s}}\cdot(i_{r(s\bar s\bar{\bar s})\tau_{s\bar s}}\cdot\gamma_{s,\bar s})\chi_{\bar{\bar s}} = \gamma_{s\bar s,\bar{\bar s}}i_{r(s\bar s\bar{\bar s})\tau_{s\bar s}\tau_{\bar{\bar s}}}u_{\bar{\bar s}}'(i_{r(s\bar s\bar{\bar s})\tau_{s\bar s}}\cdot\gamma_{s,\bar s})u_{\bar{\bar s}} =$$

$$= u_{s\bar s\bar{\bar s}}'\,u_{s\bar s}u_{\bar{\bar s}}i_{l(s\bar s\bar{\bar s})}u_{\bar{\bar s}}' i_{r(s\bar s\bar{\bar s})\tau_{s\bar s}}u_{s\bar s}'u_s u_{\bar s}u_{\bar{\bar s}}.$$

Here $u_{\bar{\bar s}}i_{l(s\bar s\bar{\bar s})}u_{\bar{\bar s}}'\in r(s\bar s\bar{\bar s})\tau_{s\bar s}$ and hence $u_{s\bar s}(u_{\bar{\bar s}}i_{l(s\bar s\bar{\bar s})}u_{\bar{\bar s}}')\cdot i_{r(s\bar s\bar{\bar s})\tau_{s\bar s}}u_{s\bar s}'\in r(s\bar s\bar{\bar s})$. Therefore we obtain that $\xi=u_{s\bar s\bar{\bar s}}'u_s u_{\bar s}u_{\bar{\bar s}}$. On the other hand, $\gamma_{s,\bar s\bar{\bar s}}\cdot\gamma_{\bar s,\bar{\bar s}}=u_{s\bar s\bar{\bar s}}'u_s u_{\bar s\bar{\bar s}}u_{\bar s\bar{\bar s}}'u_{\bar s}u_{\bar{\bar s}}\sim\xi$ in the semigroup T and thus in J, too. This verifies (D6). If s is idempotent or s is not an inverse of itself and s^* is an inverse of s in S then, by definition, $\gamma_{s,s^*}= =u_{ss^*}'u_s u_{s^*}=u_{ss^*}'u_{ss^*}$. Hence γ_{s,s^*} is idempotent which shows (D8).

In order to verify (D7) suppose that $e\in E$, $\alpha\leqq r(e)$ and $a\in I_\alpha$, $\varrho\in J_\alpha$. Then $u_e=c_{e,e}\gamma_{e,e}$ as u_e, $c_{e,e}$ and $\gamma_{e,e}$ belong to $r(e)=l(e)$. Thus, on the one hand, we have

$$ah_e = u_e a u_e' i_{\alpha\tau_e^{-1}} = u_e a i_\alpha = u_e a\gamma_{e,e}' i_\alpha = c_{e,e}\gamma_{e,e}a\gamma_{e,e}' i_\alpha = c_{e,e}aA_{\gamma_{e,e}}.$$

On the other hand,

$$\varrho\chi_e = i_{\alpha\tau_e}u_e'\varrho u_e = i_\alpha\varrho u_e = i_\alpha\varrho c_{e,e}\cdot\gamma_{e,e} = \varrho B_{c_{e,e}}\cdot\gamma_{e,e}$$

which shows that (D7) also holds.

Finally, we prove (D4). Let $s, \bar{s} \in S$ and $a \in I_\alpha$ with $\alpha \leqq l(s\bar{s})$. Utilizing that both $u_s u_{\bar{s}} a u'_{\bar{s}} u'_s$ and $u_s i_{\alpha\tau_{\bar{s}}^{-1}} u'_s$ are contained in $\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}$ one sees that

$$ah_{\bar{s}} h_s = u_s\left(u_{\bar{s}} a u'_{\bar{s}} i_{\alpha\tau_{\bar{s}}^{-1}}\right) u'_s\, i_{\alpha\tau_{\bar{s}}^{-1}\tau_{\bar{s}}^{-1}} = u_s u_{\bar{s}} a u'_{\bar{s}}\, u'_s\, u_s i_{\alpha\tau_{\bar{s}}^{-1}} u'_s\, i_{\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}} =$$

$$= u_s u_{\bar{s}} a u'_{\bar{s}} u'_s\, i_{\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}} \mathscr{R}\, c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} a \gamma'_{s,\bar{s}} u'_{s\bar{s}} i_{\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}} \mathscr{R} c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} a \gamma'_{s,\bar{s}} i_\beta u'_{s\bar{s}} i_{\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}}$$

where $\gamma_{s,\bar{s}} a \gamma'_{s,\bar{s}} \in \beta$. However, the latter element is $\mathscr{L}$-related to $i_{\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}}$ as well as $ah_{\bar{s}}h_s$. Thus we obtain that

$$ah_{\bar{s}} h_s = c_{s,\bar{s}} u_{s\bar{s}}\left(\gamma_{s,\bar{s}} a \gamma'_{s,\bar{s}} i_\beta\right) u'_{s\bar{s}} i_{\alpha\tau_{\bar{s}}^{-1}\tau_s^{-1}} = c_{s,\bar{s}} \cdot aA_{\gamma_{s,\bar{s}}} h_{s\bar{s}},$$

that is, (D4) (a) is fulfilled. Now let $s, \bar{s} \in S$ and $\varrho \in J_\alpha$ with $\alpha \leqq r(s\bar{s})$. Applying Lemma 4.1 we infer that

$$\varrho\chi_s\chi_{\bar{s}} = i_{\alpha\tau_s\tau_{\bar{s}}} u'_{\bar{s}} i_{\alpha\tau_s} u'_s \varrho u_s u_{\bar{s}} = i_{\alpha\tau_s\tau_{\bar{s}}} u'_{\bar{s}} u'_s \varrho u_s u_{\bar{s}}.$$

Here $u_s u_{\bar{s}} = c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}}$ and $u'_{\bar{s}} u'_s \sim \gamma'_{s,\bar{s}} u'_{s\bar{s}}$ whence it follows that $u'_{\bar{s}} u'_s \varrho u_s u_{\bar{s}} \mathscr{R}$ $\mathscr{R} u'_{\bar{s}} u'_s i_\alpha u_s u_{\bar{s}} \mathscr{L} \gamma'_{s,\bar{s}} u'_{s\bar{s}} i_\alpha c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} \mathscr{R} \gamma'_{s,\bar{s}} u'_{s\bar{s}} \varrho c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}}$. Consequently, we have

$$\varrho\chi_s\chi_{\bar{s}} = i_{\alpha\tau_s\tau_{\bar{s}}} \gamma'_{s,\bar{s}} u'_{s\bar{s}} \varrho c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}}.$$

(D5) ensures $(i_{\alpha\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}})' \in J_{\alpha\tau_s\tau_{\bar{s}}}$ and therefore (B6) implies that $(i_{\alpha\tau_s\tau_{\bar{s}}} \cdot \gamma'_{s,\bar{s}})' \in J_{\alpha\tau_{s\bar{s}}}$. Thus we can deduce from the last equality that

$$(i_{\alpha\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}}) \cdot \varrho\chi_s\chi_{\bar{s}} = i_{\alpha\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}} i_{\alpha\tau_s\tau_{\bar{s}}} \gamma'_{s,\bar{s}} u'_{s\bar{s}} \varrho c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} =$$

$$= i_{\alpha\tau_{s\bar{s}}} u'_{s\bar{s}} i_\alpha \varrho c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} = \varrho B_{c_{s,\bar{s}}} \chi_{s\bar{s}} \cdot \gamma_{s,\bar{s}}$$

as $\varrho c_{s,\bar{s}} \mathscr{R} \varrho \mathscr{R} i_\alpha$. This shows (D4) (b) which completes the proof of the fact that h, χ, γ is an (S, I, J)-triple.

All that remained to be proved is that the one-to-one and onto mapping Φ defined above is an isomorphism of T onto $\mathfrak{S}(S, I, J; h, \chi, \gamma)$. Let t and $\bar{t}$ be elements in T with $t\Phi = (a, s, \sigma)$ and $\bar{t}\Phi = (\bar{a}, \bar{s}, \bar{\sigma})$, respectively. This means that $t = au_s\sigma$ and $\bar{t} = \bar{a}u_{\bar{s}}\bar{\sigma}$ where $a \in I_\alpha^{ss'}$, $\sigma \in J_{\alpha\tau_s}^{s's}$ and $\bar{a} \in I_{\bar{\alpha}}^{\bar{s}\bar{s}'}$, $\bar{\sigma} \in J_{\bar{\alpha}\tau_s}^{\bar{s}'\bar{s}}$, for some α in $Y_{r(s)}$ and $\bar{\alpha}$ in $Y_{r(\bar{s})}$, respectively. Suppose that $\bar{a}A_\sigma \in I_{\alpha_1}$ and $\sigma B_{\bar{a}} \in J_{\alpha_1}$. Here $\alpha_1 \in Y_{r(\bar{s})l(s)}$ by (C1) (a). Thus we have $\alpha_1\tau_s^{-1} \in Y_{r(s\bar{s})}$ and $\alpha_1\tau_{\bar{s}} \in Y_{l(s\bar{s})}$. (D5) and (B6) ensure that $\gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}} \gamma'_{s,\bar{s}} \in \alpha_1\tau_s^{-1}\tau_{s\bar{s}}$ and $u_{s\bar{s}} \gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}} \gamma'_{s,\bar{s}} u_{(s\bar{s})'} \in \alpha_1\tau_s^{-1}$. Hence it follows that

$$i_{\alpha_1\tau_s^{-1}} c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}} = i_{\alpha_1\tau_s^{-1}} u_{s\bar{s}} \gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}} = i_{\alpha_1\tau_s^{-1}} u_{s\bar{s}} i_{\alpha_1\tau_s^{-1}\tau_{s\bar{s}}} \gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}}.$$

Applying this equality one can see that

$$(au_s\sigma)(\bar{a}u_{\bar{s}}\bar{\sigma}) = au_s \cdot \bar{a}A_\sigma \cdot \sigma B_{\bar{a}} \cdot u_{\bar{s}}\bar{\sigma} = au_s \cdot \bar{a}A_\sigma \cdot i_{\alpha_1} \cdot \sigma B_{\bar{a}} \cdot u_{\bar{s}}\bar{\sigma} =$$

$$= au_s \cdot \bar{a}A_\sigma \cdot i_{\alpha_1}(u'_s\, i_{\alpha_1\tau_s^{-1}} u_s)(u_{\bar{s}} i_{\alpha_1\tau_{\bar{s}}} u'_{\bar{s}})\, i_{\alpha_1} \cdot \sigma B_{\bar{a}} \cdot u_{\bar{s}}\bar{\sigma} =$$

$$= a(u_s \cdot \bar{a}A_\sigma u'_s\, i_{\alpha_1\tau_{\bar{s}}^{-1}})\, u_s u_{\bar{s}} (i_{\alpha_1\tau_{\bar{s}}} u'_{\bar{s}} \cdot \sigma B_{\bar{a}} \cdot u_{\bar{s}})\bar{\sigma} =$$

$$= a \cdot \bar{a}A_\sigma h_s \cdot c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} \cdot \sigma B_{\bar{a}} \chi_{\bar{s}} \cdot \bar{\sigma} = \tag{19}$$

$$= a \cdot \bar{a}A_\sigma h_s (i_{\alpha_1\tau_s^{-1}} c_{s,\bar{s}} u_{s\bar{s}} \gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}}) \sigma B_{\bar{a}} \chi_{\bar{s}} \cdot \bar{\sigma} =$$

$$= a \cdot \bar{a}A_\sigma h_s (i_{\alpha_1\tau_s^{-1}} u_{s\bar{s}} i_{\alpha_1\tau_s^{-1}\tau_{s\bar{s}}} \gamma_{s,\bar{s}} i_{\alpha_1\tau_{\bar{s}}}) \sigma B_{\bar{a}} \chi_{\bar{s}} \cdot \bar{\sigma} =$$

$$= (a \cdot \bar{a}A_\sigma h_s) u_{s\bar{s}} ((i_{\alpha_1\tau_s^{-1}\tau_{s\bar{s}}} \cdot \gamma_{s,\bar{s}}) \cdot \sigma B_{\bar{a}} \chi_{\bar{s}} \cdot \bar{\sigma}).$$

This proves that Φ is a homomorphism and therefore an isomorphism. The proof of the theorem is complete.

As an application of Theorem 3.10 and 4.3 we describe the structure of orthodox semigroups by means of their bands of idempotents and greatest inverse semigroup homomorphic images. An alternative structure theorem was given by YAMADA [7]. However, our construction is more economic as it makes use of structure mappings in a single variable only.

Let S be an inverse semigroup with semilattice of idempotents Y. For every α in Y, let I_α and J_α be a left zero semigroup and a right zero semigroup with distinguished elements i_α and j_α, respectively. Let I be a lower associative semilattice Y of the left zero semigroups I_α $(\alpha \in Y)$ and J an upper associative semilattice Y of the right zero semigroups J_α $(\alpha \in Y)$. Assume that A, B is an (I, J)-pair satisfying the property that

(C4)' $aA_{j_\beta} = i_\beta \cdot a$ and $\sigma B_{i_\beta} = \sigma \cdot j_\beta$ provided $\alpha, \beta \in Y$ with $\alpha \leqq \beta$ and $a \in I_\alpha$, $\sigma \in J_\alpha$.

Let h_s: $\bigcup \{I_\alpha : \alpha \leqq s^{-1}s\} \to \bigcup \{I_\alpha : \alpha \leqq ss^{-1}\}$ and χ_s: $\bigcup \{J_\alpha : \alpha \leqq ss^{-1}\} \to \bigcup \{J_\alpha : \alpha \leqq s^{-1}s\}$ be mappings such that $I_\alpha h_s \subseteq I_{s\alpha s^{-1}}$ for $\alpha \leqq s^{-1}s$, $J_\alpha \chi_s \subseteq J_{s^{-1}\alpha s}$ for $\alpha \leqq ss^{-1}$ and the following conditions are fulfilled:

(D2)' (a) if $a \in I_\alpha$, $b \in I_\beta$ with $s^{-1}s \geqq \alpha \geqq \beta$ then $ah_s \cdot bh_s = (a \cdot bA_{j_{s\alpha s^{-1}}\chi_s})h_s$,

(b) if $\varrho \in J_\alpha$, $\sigma \in J_\beta$ with $\alpha \leqq \beta \leqq ss^{-1}$ then $\varrho\chi_s \cdot \sigma\chi_s = (\varrho B_{i_{s^{-1}\beta s}h_s} \cdot \sigma)\chi_s$;

(D3)' if $a \in I_\alpha$ with $\alpha \leqq s^{-1}s$ and $\sigma \in J_\beta$ with $\beta \leqq ss^{-1}$ then

(a) $aA_{\sigma\chi_s}h_s = i_{s^{-1}\beta s}h_s \cdot ah_s A_\sigma$,

(b) $\sigma B_{ah_s}\chi_s = \sigma\chi_s B_a \cdot j_{s\alpha s^{-1}}\chi_s$;

(D4)' (a) if $a \in I_\alpha$ with $\alpha \leqq (s\bar{s})^{-1}s\bar{s}$ then $ah_{\bar{s}}h_s = c \cdot aA_\gamma h_{s\bar{s}}$ for some c in $I_{s\bar{s}(s\bar{s})^{-1}}$ and γ in $J_{(s\bar{s})^{-1}s\bar{s}}$,

(b) if $\varrho \in J_\alpha$ with $\alpha \leqq s\bar{s}(s\bar{s})^{-1}$ then $\varrho\chi_s\chi_{\bar{s}} = \varrho B_c \chi_{s\bar{s}} \cdot \gamma$ for some c in $I_{s\bar{s}(s\bar{s})^{-1}}$ and γ in $J_{(s\bar{s})^{-1}s\bar{s}}$;

(D7)' (a) if $a \in I_\alpha$ and $\alpha \leqq \beta$ then $ah_\beta = c \cdot aA_\gamma$ for some c in I_β and γ in J_β,
(b) if $\varrho \in J_\alpha$ and $\alpha \leqq \beta$ then $\varrho\chi_\beta = \varrho B_c \cdot \gamma$ for some c in I_β and γ in J_β.

A pair of mappings h, χ possessing these properties is called a *reduced* (S, I, J)-*pair*.

Let us define a multiplication on the set

$$\mathbf{S} = \{(a, s, \sigma) \colon s \in S, a \in I_{ss^{-1}}, \sigma \in J_{s^{-1}s}\}$$

as follows:

$$(a, s, \sigma)(\bar{a}, \bar{s}, \bar{\sigma}) = (a \cdot \bar{a}A_\sigma h_s, s\bar{s}, \sigma B_{\bar{a}} \chi_{\bar{s}}).$$

If the constant γ in (D4)' can be chosen such that it depends on s and $\bar{s}$ only and, moreover, (D7)' holds with the γ corresponding to the pair (β, β) then the reduced (S, I, J)-pair h, χ can be easily extended to an (S, I, J)-triple. Then Theorem 3.10 immediately implies that $\mathbf{S}$ forms an orthodox semigroup with respect to this multiplication whose band of idempotents is isomorphic to $\mathscr{B}(I, J; A, B)$ and whose greatest inverse semigroup homomorphic image is isomorphic to S. However, as c is not needed to be independent of the choice of a and ϱ in properties (D4) and (D7) since I_α is a left zero semigroup for every α in Y, we need not assume this property for γ if J_α is a right zero semigroup for each α in Y. Therefore the conclusion for $\mathbf{S}$ drawn up above holds for any reduced (S, I, J)-pair. The orthodox semigroup obtained in this way will be denoted by $\mathfrak{S}(S, I, J; h, \chi)$.

Conversely, if T is an orthodox semigroup with band of idempotents E then its least inverse semigroup congruence is clearly a strong $(E, \mathscr{D})$-parcelling congruence. The second part of the following theorem immediately follows from Theorem 4.3, one has to observe only that, in this special case, the constants $\gamma_{s,\bar{s}}$ can be eliminated in (D4) (b) and (19) as well as the contants $c_{s,\bar{s}}$ can in (D4) (a) and (19).

Theorem 4.4. *Let S be an inverse semigroup with semilattice of idempotents Y. For every α in Y, let I_α be a left zero semigroup and J_α a right zero semigroup with distinguished elements i_α and j_α, respectively. Suppose I to be a lower associative semilattice Y of the left zero semigroups I_α $(\alpha \in Y)$ and J an upper associative semilattice Y of the right zero semigroups J_α $(\alpha \in Y)$. Let A, B be an (I, J)-pair satisfying* (C4)'. *Assume that h, χ is a reduced (S, I, J)-pair. Then $\mathfrak{S}(S, I, J; h, \chi)$ is an orthodox semigroup with band of idempotents isomorphic to $\mathscr{B}(I, J; A, B)$ and with greatest inverse semigroup homomorphic image isomorphic to S.*

Conversely, if T is an orthodox semigroup with band of idempotents E which is a semilattice Y of rectangular bands then E is isomoprhic to $\mathscr{B}(I, J; A, B)$ for some I, J, A and B which fulfil the conditions required above. Moreover, denoting by S the greatest inverse semigroup homomorphic image of T, there exists a reduced (S, I, J)-pair h, χ such that T is isomorphic to $\mathfrak{S}(S, I, J; h, \chi)$.

References

[1] A. H. CLIFFORD and G. B. PRESTON, *The Algebraic Theory of Semigroups. I*, Amer. Math. Soc. (Providence, R. I., 1961).

[2] J. M. HOWIE, *An Introduction to Semigroup Theory*, Academic Press (London, New York, San Francisco, 1976).

[3] N. KIMURA, The structure of idempotent semigroups (I), *Pacific J. Math.*, **8** (1958), 257—275.

[4] M. B. SZENDREI, On subband-parcelling congruences of orthodox semigroups, *Colloq. Math. Soc. János Bolyai*, **20**. Algebraic Theory of Semigroups (Szeged, 1976), 613—627.

[5] M. B. SZENDREI, Strong subband-separating extensions of orthodox semigroups, *Acta Math. Acad. Sci. Hungar.*, **35** (1980), 437—450.

[6] R. J. WARNE, On the structure of semigroups which are unions of groups, *Trans. Amer. Math. Soc.*, **186** (1973), 385—401.

[7] M. YAMADA, On a regular semigroup in which the idempotents form a band, *Pacific J. Math.*, **33** (1970), 261—272.

JÓZSEF ATTILA UNIVERSITY
BOLYAI UNSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Inner injective transextensions of semigroups

M. DEMLOVÁ, P. GORALČÍK and V. KOUBEK

1. Introduction

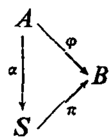
Consider an extension A of a semigroup B, that is to say, a surjective homomorphism $\varphi\colon A\to B$. The multiplication in A naturally appears as determined by B up to the extension congruence $\operatorname{Ker}\varphi$. If we succeed in localizing elements in each of the blocks of $\operatorname{Ker}\varphi$ (e.g. by an assignment of some kind of coordinates to them) then we can refine the multiplication modulo $\operatorname{Ker}\varphi$ in A up to the multiplication of individual elements. These may be represented as couples $(\varphi(a), \lambda(a))$, where $\varphi(a)$ is the label of the block containing a and $\lambda(a)$ is the secondary label (coordinate) by which a can be located within its block.

The secondary labels are taken from some suitable auxiliary set X and assigned to the elements of A by a function $\lambda\colon A\to X$, which we call a *localizer* and require (for this purpose) to be injective on each block of $\operatorname{Ker}\varphi$. Identifying A with a subset $S\subseteq B\times X$ by $a\mapsto(\varphi(a), \lambda(a))$, we can determine a unique function $f\colon S\times S\to X$ such that

$$(b, m)(c, n) = \big(bc, f(b, m, c, n)\big) \tag{1}$$

for every (b, m), $(c, n)\in S$.

The set S together with the multiplication given by (1) is thus a semigroup isomorphic to A by the isomorphism $\alpha\colon A\to S$ with $\alpha(a)=(\varphi(a), \lambda(a))$, the first projection $\pi\colon S\to B\colon(b, m)\mapsto b$ is a surjective homomorphism onto B and the diagram



commutes. For this reason, we say that $\pi\colon S\to B$ is a *semiproduct* right equivalent to the extension $\varphi\colon A\to B$, or a semiproduct representing $\varphi\colon A\to B$.

Received February 13, 1981, and in revised form December 10, 1981.

It is clear that every extension can be represented by a right equivalent semiproduct, thus we can construct various extensions of a given semigroup B also in this form. The general problem of finding all extensions of B can thus be reduced, from this point of view, to special kinds of completions of partial semigroups according to the following

Extension scheme: Let a partial semigroup P be given together with a homomorphism $\varphi: P \to B$ onto a semigroup B. Complete P to a semigroup A with the same underlying set by turning it into a semiproduct in such a way that φ: $P \to B$ becomes the projection φ: $A \to B$ of the semiproduct.

In the special case when no product is defined in P one has to choose, according to this extension scheme, a localizer λ: $P \to X$ and to find a function f making the multiplication defined by (1) associative.

The classical group extensions easily come under the above extension scheme: P appears here as a partial group divided into blocks of equal size by a homomorphism φ of P onto a group B, the block which is mapped onto the identity of B is a subgroup C of P, and all products ca, ac for $c \in C$, $a \in P$, are defined in such a way that the action of C on each block is simply transitive.

Namely, the simply transitive action of C on the blocks of Ker φ was used by Schreier to build up a most natural localizer: if we choose in a block a reference point x_0 then we get a bijection λ between the block and the group C by setting $\lambda(x)=c$ iff $cx_0=x$.

In this paper we want to carry over Schreier's idea to semigroup extension schemes, in which in the partial semigroup P to be completed we have only the products ax for a single left cancellable element $a \in P$ and for all $x \in P$ defined, and the surjective homomorphism $\varphi: P \to B$ takes a to a left identity $\varphi(a)=e$ of B. In this case the blocks of Ker φ are just the connected components of the graph with edges (x, ax), $x \in P$, and since the action of a is injective on each block it is only natural to use integers Z (or integers modulo some m) to coordinatize the blocks.

Although the idea is very simple, the detailed elaboration which follows is far from being so. We would like to acknowledge our thanks to L. Márki who helped us to put right a number of technical items.

2. $\mathcal{S}$-transextension scheme

Our basic category will be the category $\mathcal{G}^\circ$ of *pointed groupoids* (G, a), $a \in G$, with morphisms h: $(G, a) \to (H, b)$ respecting both multiplication and points (considered as nullary operations). The subcategory of $\mathcal{G}^\circ$ of *pointed semigroups* will be denoted by $\mathcal{S}^\circ$.

Given a pointed groupoid (G, a), we can define its *left connectedness* to be the equivalence $\sim_a$ on G defined by

$$x \sim_a y \Leftrightarrow \exists m, n \quad (a^m \cdot x = a^n \cdot y)$$

where $a^m = a \cdot a^{m-1}$ for $m \geqq 2$.

2.1. Definition. An extension $\varphi\colon (T, a) \to (S, e)$ in $\mathscr{G}^\circ$ is called an *inner injective transextension* or shortly an *$\mathscr{I}$-extension,* if a is left cancellative and the left connectedness $\sim_a$ is a congruence on T such that $\sim_a = \operatorname{Ker} \varphi$. Then e is a left identity in S called the left identity of the extension.

Each $\mathscr{I}$-extension $\varphi\colon (T, a) \to (S, e)$ determines an assignment $x \mapsto \mathscr{U}_x = (T_x, f_x)$, $T_x = \varphi^{-1}(x)$, of unary algebras $\mathscr{U}_x$ to elements x of S, with one injective connected operation f_x coinciding with the left inner translation by a restricted to $\varphi^{-1}(x)$.

Let $\mathscr{N}$ denote the semiring $N \cup \{\infty\}$ of non-negative integers completed by a greatest element, where $m \cdot \infty = \infty \cdot m = \infty$ for all $m \neq 0$, $0 \cdot \infty = \infty \cdot 0 = 0$. We assign to every injective connected unar (X, f) an element of $\mathscr{N}$, denoted Type(f), as follows:

Type(f) = $\min \{n;\ 0 < n < \infty$ and $f^n(x) = x$ for all $x \in X\}$ if such n exists,
Type$(f) = 0$ iff $f^{-1}(x) = \emptyset$ for some $x \in X$,
Type$(f) = \infty$ otherwise.

The semiring $\mathscr{N}$ is lattice ordered by the divisibility relation. We denote by $\vee$ and $\wedge$ the lattice operations of the least common multiple and the greatest common divisor, respectively.

We have the following easy statement readily obtained from the results of Novotný [5] on commuting transformations.

2.2. Statement. *Let (X, f) and (Y, g) be injective connected unars. There exists a homomorphism $h\colon (X, f) \to (Y, g)$, $hf = gh$, iff* Type(g) *divides* Type(f) *in $\mathscr{N}$. If* Type$(f) \neq 0$, *then h must be surjective. The unars (X, f) and (Y, g) are isomorphic iff* Type(f) = Type(g).

Returning to the $\mathscr{I}$-extension $\varphi\colon (T, a) \to (S, e)$, we can describe the assignment $x \mapsto \mathscr{U}_x$ up to isomorphism by a *type function* $r\colon S \to \mathscr{N}\colon x \mapsto$ Type(f_x). On the other hand, we can start with (S, e), e a left identity of S, and a function $r\colon S \to \mathscr{N}\colon x \mapsto r(x)$ as a sort of "plot" for the construction of an $\mathscr{I}$-extension of (S, e) in the form of an S-semiproduct, using the ring Z of integers as an auxiliary algebra. We form $S \times Z$ and identify $(x, m) \equiv (y, n)$ iff $x = y$ and $m \equiv n \pmod{r(x)}$ for $0 < r(x) < \infty$, $m = n$ otherwise. With the aid of an *initialization function* $i\colon S \to Z\colon x \mapsto i(x)$ we cut out of $S \times Z/\equiv$ a unar (P, f) with $P = \{(m, n) \in S \times Z \mid m \geqq i(x)$ if $r(x) = 0\}/\equiv$

and $f(x, m) \equiv (x, m+1)$. To turn P into a groupoid we introduce two additional functions $k: S \to Z$ and $l: S \times S \to Z$ and prescribe a *multiplication formula*

$$\text{(M)} \qquad (x, m)(y, n) \equiv (xy, m+k(x)n+l(x\mu, y))$$

where $\mu=\mu(y, n)=e$ if $r(y) \neq 0$, or $r(y)=0$ and $n \neq i(y)$, μ is the empty symbol if $r(y)=0$ and $n=i(y)$. The sixtuple (S, e, r, i, k, l) sets up an *$\mathcal{I}$-extension construction scheme,* or shortly an *$\mathcal{I}$-scheme.* If (M) correctly defines a multiplication as a function $P \times P \to P$, we say that the $\mathcal{I}$-scheme is *$\mathcal{G}$-correct.* A $\mathcal{G}$-correct $\mathcal{I}$-scheme turns the unar (P, f) into a groupoid satisfying $f(uv)=f(u)v$ for all $u, v \in P$. If moreover there exists an $a \in P$ of the form $a=(e, m)$, for some $m \in Z$, and such that $f(t)=at$ for all $t \in P$, then we call the $\mathcal{I}$-scheme *$\mathcal{G}^\circ$-correct.* A $\mathcal{G}^\circ$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l) determines a unique semiproduct

$$\pi: (P, a) \to (S, e): (x, m) \mapsto x$$

in $\mathcal{G}^\circ$ and we write in this case $(P, a)=\mathcal{I}(S, e, r, i, k, l)$.

It will be our immediate task to find the conditions for an $\mathcal{I}$-extension scheme (S, e, r, i, k, l) to be $\mathcal{G}^\circ$-correct. This done, we shall next be most interested in the "associative" $\mathcal{I}$-schemes determining $\mathcal{I}$-extensions in $\mathcal{S}^\circ$ — the *$\mathcal{S}^\circ$-correct $\mathcal{I}$-extension* schemes. These found, we investigate how large a class of $\mathcal{I}$-extensions in $\mathcal{S}^\circ$ can be optained by the class of all $\mathcal{S}^\circ$-correct $\mathcal{I}$-schemes. We shall prove that all $\mathcal{I}$-extensions in $\mathcal{S}^\circ$ can thus be obtained. Then we shall clear up a technical point when two $\mathcal{I}$-schemes determine right equivalent S-semiproducts, in order to get possibly simple semiproduct representatives of $\mathcal{I}$-extensions in $\mathcal{S}^\circ$. We shall also state, in a number of statements, conditions under which an $\mathcal{S}^\circ$-correct $\mathcal{I}$-extension scheme determines an $\mathcal{I}$-extension in the category of pointed

— semigroups with identity (or "monoids"),
— commutative (=abelian) semigroups,
— right cancellative semigroups,
— left cancellative semigroups,
— right reductive semigroups,
— groups.

In particular, in the case of group $\mathcal{I}$-extensions our theory comes to a strong resemblance with the theory of extensions of P. A. Grillet [2].

Our final point will be to show the role of $\mathcal{I}$-extensions in $\mathcal{S}^\circ$ in a larger class of transextensions.

3. $\mathcal{S}^\circ$-correctness of $\mathcal{I}$-schemes

3.1. Statement. *An $\mathcal{I}$-scheme (S, e, r, i, k, l) is $\mathcal{G}$-correct iff the following three "correctness conditions" hold for any $x, y \in S$:*

(C1) $r(xy)=0 \Rightarrow k(x) \geqq 0,\ k(x)r(y)=0,$ *and* $i(x)+k(x)i(y)+\min\{k(x)+l(xe, y), l(x, y)\} \geqq i(xy),\ (r(y)\neq 0 \Rightarrow i(x)+l(xe, y) \geqq i(xy)),$

(C2) $r(xy)$ *divides* $r(x)$ *in* $\mathcal{N}$,

(C3) $r(y)\neq\infty \Rightarrow k(x)r(y)\equiv 0 \pmod{r(xy)}$.

Proof. Assume that $(x, m)(y, n)\in P$ for all $(x, m), (y, n)\in P$. If $r(xy)=0$, then $r(x)=0$. Indeed, if $r(x)\neq 0$ we can, for a given $(y, n)\in P$, choose $(x, m)\in P$ for which $m+k(x)n+\max\{l(x, y), l(xe, y)\}<i(xy)$, thus $(x, m)(y, n)\notin P$. Likewise, if $r(xy)=0$ and at the same time $k(x)<0$ or $k(x)r(y)\neq 0$, then for some $(y, n)\in P$, where $(y, n)\not\equiv(y, i(y))$ or $r(y)\neq 0$, we have $i(x)+k(x)n+l(xe, y)< <i(xy)$, thus $(x, i(x))(y, n)\notin P$, a contradiction. So if $r(xy)=0$, then $k(x)\geqq 0$ and $k(x)r(y)=0$. Now if $r(xy)=0=r(y)$, we must have both

$$i(x)+k(x)\big(i(y)+1\big)+l(xe, y) \geqq i(xy)$$

and

$$i(x)+k(x)i(y)+l(x, y) \geqq i(xy).$$

If $r(xy)=0\neq r(y)$, then $k(x)=0$ and it must be $i(x)+l(xe, y)\geqq i(xy)$. We have proved (C1) under the assumption that $(x, m)(y, n)\in P$ for any $(x, m), (y, n)\in P$.

Assume now (C1) and that $r(xy)=0$ implies $r(x)=0$, and let $(x, m), (y, n)\in P$. If $r(xy)\neq 0$ then clearly $(x, m)(y, n)\in P$. Let $r(xy)=0$. Then $r(x)=0$, hence $m\geqq i(x)$. If further $r(y)=0$, then also $n\geqq i(y)$, and since $k(x)\geqq 0, k(x)n\geqq k(x)i(y)$. Therefore, for $n>i(y)$,

$$m+k(x)n+l(xe, y) \geqq i(x)+k(x)\big(i(y)+1\big)+l(xe, y) \geqq i(xy);$$

for $n=i(y)$,

$$m+k(x)n+l(x, y) \geqq i(x)+k(x)i(y)+l(x, y) \geqq i(xy),$$

hence $(x, m)(y, n)\in P$. If $r(y)\neq 0$, then $k(x)=0$,

$$m+k(x)n+l(xe, y) \geqq i(x)+l(xe, y) \geqq i(xy),$$

hence again $(x, m)(y, n)\in P$.

We conclude that (C1) holds and $r(xy)=0$ implies $r(x)=0$ iff P is closed under the (multivalued) multiplication given by (M) for a given $\mathcal{I}$-scheme. Notice that (C2) and $r(xy)=0$ imply $r(x)=0$.

Assume next that

(1) $$(x, m)(y, n) \equiv (x, m')(y, n) \quad \text{whenever} \quad (x, m) \equiv (x, m').$$

Then for (x, m) with $r(x) \neq \infty$ and any (y, n),

$$(x, m+r(x))(y, n) \equiv (x, m)(y, n),$$

hence $r(x) \equiv 0 \pmod{r(xy)}$, which means that $r(xy)$ divides $r(x)$ in $\mathcal{N}$. If $r(x)=\infty$ then $r(xy) \neq 0$ by the above, hence again $r(xy)$ divides $r(x)$ in $\mathcal{N}$. We have proved (C2) under the assumption (1). Conversely, assume (C2) and let $(x, m) \equiv (x, m')$. Then for any $(y, n) \in P$, $(x, m)(y, n) \equiv (x, m')(y, n)$. Indeed, if $r(x)=0$ or ∞ then $m=m'$. If $0<r(x)<\infty$, then $m \equiv m' \pmod{r(x)}$, thus $m \equiv m' \pmod{r(xy)}$ by (C2), therefore

$$m+k(x)n+l(x\mu, y) \equiv m'+k(x)n+l(x\mu, y) \pmod{r(xy)}.$$

Assume finally that $(x, m)(y, n) \equiv (x, m)(y, n')$ whenever $(y, n) \equiv (y, n')$. Then for (y, n) with $r(y) \neq \infty$ and any $(x, m) \in P$, $(x, m)(y, n+r(y)) \equiv (x, m)(y, n)$, hence $k(x)r(y) \equiv 0 \pmod{r(xy)}$. We have proved (C3) under the assumption. Conversely, assume (C3) and let $(y, n) \equiv (y, n')$. Then for any $(x, m) \in P$, $(x, m)(y, n) \equiv (x, m)(y, n')$. This is clear for $r(y)=0$ or ∞, since then $n=n'$. For $0<r(y)<\infty$ we have by (C3)

$$m+k(x)n+l(xe, y) \equiv m+k(x)n'+l(xe, y) \pmod{r(xy)}$$

since $n \equiv n' \pmod{r(y)}$.

We conclude that $(x, m)(y, n) \equiv (x, m')(y, n')$ whenever $(x, m) \equiv (x, m')$ and $(y, n) \equiv (y, n')$ iff both (C2) and (C3) hold.

3.2. Statement. *A $\mathcal{G}$-correct $\mathcal{S}$-scheme (S, e, r, i, k, l) is $\mathcal{G}^\circ$-correct iff the following three „inner translativity" conditions hold for any $x \in S$:*

(IT1) $k(e) \equiv 1 \pmod{r(e)}$,
(IT2) $l(e, x) \equiv l(e, e) \pmod{r(x)}$,
(IT3) $r(e) = 0 \Rightarrow 1-l(e, e) \geqq i(e)$.

The unique a in P for which $\pi\colon (P, a) \to (S, e)\colon (x, m) \mapsto x$ is an $\mathcal{S}$-extension in $\mathcal{G}^\circ$ is then $a \equiv (e, 1-l(e, e))$.

Proof. Call $a \in P$ an *admissible point* if the $\mathcal{S}$-scheme yields an $\mathcal{S}$-extension $\pi\colon (P, a) \to (S, e)\colon (x, m) \mapsto x$ in $\mathcal{G}^\circ$ with $a(x, m) \equiv f(x, m) \equiv (x, m+1)$. Assume that $a=(e, p) \in P$ is admissible. Then $(e, p+2) \equiv (e, p)(e, p+1) \equiv (e, p+k(e)+k(e)p+ +l(e, e))$, $(e, p+1) \equiv (e, p)(e, p) \equiv (e, p+k(e)p+l(e, e))$, hence

$$p+k(e)p+k(e)+l(e, e) \equiv p+2 \pmod{r(e)},$$

$$p+k(e)p \qquad +l(e, e) \equiv p+1 \pmod{r(e)}.$$

Subtracting the two congruences we get (IT1). Since $r(x)$ divides $r(e)$ by (C2), (IT1) is equivalent to $\forall x\ (k(e)\equiv 1 \pmod{r(x)})$, therefore for any $(x, m)\in P$,

$$(x, m+1) \equiv (e, p)(x, m) \equiv (x, p+m+l(e, x)),$$

hence

$$p+l(e, x) \equiv 1 \pmod{r(x)}.$$

In particular,

$$p+l(e, e) \equiv 1 \pmod{r(e)}$$

whence the uniqueness of an admissible $a\in P$. By (C2) again, the latest congruence implies

$$p+l(e, e) \equiv 1 \pmod{r(x)}$$

hence we get (IT2).

If $r(e)=0$ then $p=1-l(e, e)\geqq i(e)$ gives (IT3).

Conversely, if the three conditions hold, then $(e, 1-l(e, e))\in P$ and

$$(e, 1-l(e, e))(x, m) \equiv (x, 1-l(e, e)+k(e)m+l(e, x)) \equiv (x, m+1)$$

for any $(x, m)\in P$.

3.3. Statement. *Let (S, e, r, i, k, l) be a $\mathcal{G}$-correct $\mathcal{I}$-scheme with S a semigroup. Then P is a semigroup iff the following "associativity conditions" hold:*

(A0) $k(x)k(y) \equiv k(xy) \pmod{r(xye)}$,
(A1) $l(xe, y)+l(xye, z) \equiv k(x)l(ye, z)+l(xe, yz) \pmod{r(xyz)}$,
(A2) *if* $r(y) = 0$ *and* $(r(yz) = 0 = k(y) \Rightarrow i(y)+l(ye, z) > i(yz))$, *then* $l(xe, y) \equiv l(x, y) \pmod{r(xyz)}$,
(A3) *if* $r(y) = 0 = r(yz)$ *and* $i(y)+k(y)(i(z)+1)+l(ye, z) = i(yz)$, *then* $l(xe, y)-l(x, y) \equiv l(xe, yz)-l(x, yz) \pmod{r(xyz)}$,
(A4) *if* $r(z) = 0$, *then* $l(xye, z)-l(xy, z) \equiv k(x)(l(ye, z)-l(y, z)) \pmod{r(xyz)}$,
(A5) *if* $r(y) = r(z) = r(yz) = 0$ *and* $i(y)+k(y)i(z)+l(y, z)=i(yz)$, *then* $l(x, y)+l(xy, z) \equiv k(x)l(y, z)+l(x, yz) \pmod{r(xyz)}$,
(A6) *if* $r(y) = r(yz) = 0 = k(y)$ *and* $i(y)+l(ye, z) = i(yz)$ *then* $l(xe, yz) \equiv l(x, yz) \pmod{r(xyz)}$.

Proof. For arbitrary $(x, m), (y, n), (z, p)\in P$,

(1) $$[(x, m)(y, n)](z, p) \equiv$$
$$\equiv (xyz, m+k(x)n+k(xy)p+l(x\mu(y, n), y)+l(xy\mu(z, p), z),$$

(2) $$(x, m)[(y, n)(z, p)] \equiv$$
$$\equiv (xyz, m+k(x)n+k(x)k(y)p+k(x)l(y\mu(z, p), z)+l(x\mu((y, n)(z, p)), yz)).$$

We see from these expressions that the consideration of the equality

$$[(x, m)(y, n)](z, p) \equiv (x, m)[(y, n)(z, p)] \tag{3}$$

will depend on the triple

$$P(y, n, z, p) \equiv \big(\mu(y, n), \mu(z, p), \mu((y, n)(z, p))\big)$$

which will be referred to as the *pattern* of the couple $((y, n), (z, p)) \in P \times P$.

From the correctness conditions (C1) and (C2) of 3.1 it follows immediately

$$\mu(y, n) = e \Rightarrow \mu\big((y, n)(z, p)\big) = e. \tag{4}$$

Indeed, if $r(y) \neq 0$ then by (C2) also $r(yz) \neq 0$. If $r(y)=0$ and $n > i(y)$, then

$$(y, n)(z, p) \equiv \big(yz, n+k(y)p+l(y\mu(z, p), z)\big).$$

If $r(yz) \neq 0$ then there is nothing to prove. Supposing $r(yz)=0$ we have by (C1)

$$n+k(y)p+l\big(y\mu(z, p), z\big) > i(y)+k(y)i(z)+\min\{k(y)+l(ye, z), l(y, z)\} \geq i(yz).$$

By (4) the number of possible patterns is reduced from eight to six listed as

$$P_1 = (e, e, e), \quad P_2 = (1, e, e), \quad P_3 = (1, e, 1),$$

$$P_4 = (e, 1, e), \quad P_5 = (1, 1, 1), \quad P_6 = (1, 1, e).$$

To each triple $((x, m), (y, n), (z, p))$ of elements of P we associate an equation (modulo $\equiv$)

$$C_j(x, y, z)\colon\ l(x\mu(y, n), y)+l(xy\mu(z, p), z) \equiv$$

$$\equiv k(x)l\big(y\mu(z, p), z\big)+l\big(x\mu((y, n)(z, p)), yz\big) \pmod{r(xyz)}$$

if the corresponding pattern is $P(y, n, z, p)=P_j$, $j=1, \dots, 6$. Written in full, the six equations are

$$C_1\colon\ l(xe, y)+l(xye, z) \equiv k(x)l(ye, z)+l(xe, yz) \pmod{r(xyz)}$$

$$C_2\colon\ l(x, y)+l(xye, z) \equiv k(x)l(ye, z)+l(xe, yz) \pmod{r(xyz)}$$

$$C_3\colon\ l(x, y)+l(xye, z) \equiv k(x)l(ye, z)+l(x, yz) \pmod{r(xyz)}$$

$$C_4\colon\ l(xe, y)+l(xy, z) \equiv k(x)l(y, z)+l(xe, yz) \pmod{r(xyz)}$$

$$C_5\colon\ l(x, y)+l(xy, z) \equiv k(x)l(y, z)+l(x, yz) \pmod{r(xyz)}$$

$$C_6\colon\ l(x, y)+l(xy, z) \equiv k(x)l(y, z)+l(xe, yz) \pmod{r(xyz)}$$

Assume that P fulfils (A0). Then we see from (1) and (2) that (3) holds iff the equation corresponding to $P(y, n, z, p)$ is true. Since $ez=z$ we get by (C2) that $r(z)$ divides $r(e)$ and (A0) implies $k(x)k(y) \equiv k(xy) \pmod{r(xyz)}$ for any $z \in S$.

Thus if one shows that the condition (A) of associativity of P implies (A0) then we can express it as

(5) $\quad$ (A) $\Leftrightarrow$ (A0) and $\forall y, z, j\big(\exists n, p(P(y, n, z, p) = P_j\big) \Rightarrow \forall x\big(C_j(x, y, z)$ holds$\big)\big)$.

To show that (A) implies (A0), note that we can always choose $n, p \in Z$ to $y, z \in S$ so that $P(y, n, z, p) = P_1$, and that then also $P(y, n, z, p') = P_1$ for every $p' \geqq p$. Comparing (1) and (2) for p and p replaced by $p+1$ we get immediately (A0), as well as (A1). It remains to consider the situations under which the patterns $P_2, \ldots, P_6$ occur, in order to get (A2)—(A6). The scheme for (Aj), $j = 2, \ldots, 5$, is

$$\exists n, p\big(P(y, n, z, p) = P_j\big) \Rightarrow \forall x\big(C'_j(x, y, z) \text{ holds}\big)$$

where $C'_5 = C_5$ and $C'_j = C_1 - C_j$ (this is meant to symbolize that C'_j is obtained by subtracting C_j from the always true equation C_1, hence C'_j is equivalent to C_j, but somewhat simpler) for $j = 2, 3, 4$.

($j=2$): We have $P(y, n, z, p) = P_2$ iff $\big(r(y) = 0$ and $n = i(y)\big)$ and $\big(r(z) = 0 \Rightarrow p > i(z)\big)$ and $\big(r(yz) = 0 \Rightarrow i(y) + k(y)p + l(ye, z) > i(yz)\big)$, therefore it follows that $\exists n, p\,\big(P(y, n, z, p) = P_2\big) \Leftrightarrow r(y) = 0$ and $\big(r(yz) = 0 = k(y) \Rightarrow i(y) + l(ye, z) > i(yz)\big)$ since for $r(yz) = 0 \neq k(y)$ we have by (C1) that $k(y) > 0$ and $r(z) = 0$, hence for some $p, p > i(z)$, it is $i(y) + k(y)p + l(ye, z) > i(yz)$.

($j=3$): $P(y, n, z, p) = P_3 \Leftrightarrow \big(r(y) = 0$ and $n = i(y)\big)$ and $\big(r(z) = 0 \Rightarrow p > i(z)\big)$ and $\big(r(yz) = 0$ and $i(y) + k(y)p + l(ye, z) = i(yz)\big)$, therefore $\exists n, p\,\big(P(y, n, z, p) = P_3\big) \Leftrightarrow \big(r(y) = 0 = r(yz)\big)$ and $\big(k(y) = 0 \Rightarrow i(y) + l(ye, z) = i(yz)\big)$ and $\big(k(y) > 0 \Rightarrow i(y) + k(y)\big(i(z) + 1\big) + l(ye, z) = i(yz)\big)$ since for $k(y) \neq 0$ it is $r(z) = 0$ by (C1) and $p = i(z) + 1$ is the least possible choice for p to get the pattern. Putting together, $\exists n, p\,\big(P(y, n, z, p) = P_3\big) \Leftrightarrow r(y) = 0 = r(yz)$ and $i(y) + k(y)\big(i(z) + 1\big) + l(ye, z) = i(yz)$.

($j=4$): $P(y, n, z, p) = P_4 \Leftrightarrow \big(r(y) = 0 \Rightarrow n > i(y)\big)$ and $\big(r(z) = 0$ and $p = i(z)\big)$ and $\big(r(yz) = 0 \Rightarrow n + k(y)i(z) + l(y, z) > i(yz)$, therefore $\exists n, p\big(P(y, n, z, p) = P_4\big) \Leftrightarrow r(z) = 0$.

($j=5$): $P(y, n, z, p) = P_5 \Leftrightarrow \big(r(y) = 0$ and $n = i(y)\big)$ and $\big(r(z) = 0$ and $p = i(z)\big)$ and $\big(r(yz) = 0$ and $i(y) + k(y)i(z) + l(y, z) = i(yz)\big)$, therefore $\exists n, p\,\big(P(y, n, z, p) = P_5\big) \Leftrightarrow r(y) = r(z) = r(yz) = 0$ and $i(y) + k(y)i(z) + l(y, z) = i(yz)$.

To conclude the proof we show that (A1)—(A6) and $P(y, n, z, p) = P_6$ imply C_6, and that C_6 and (A1)—(A5) imply (A6). We have

$$P(y, n, z, p) = P_6 \Leftrightarrow \big(r(y) = 0 \text{ and } n = i(y)\big) \text{ and } \big(r(z) = 0 \text{ and }$$
$$p = i(z)\big) \text{ and } \big(r(yz) = 0 \Rightarrow i(y) + k(y)i(z) + l(y, z) > i(yz)\big).$$

If $r(yz) = 0 \neq k(y)$ then $i(y) + l(ye, z) \geqq i(yz)$, thus by (C1), $P(y, n, z, p+2) = P_2$. Since $P(y, n+1, z, p) = P_4$, C_2 and C_4 hold by ($j=2$) and ($j=4$). Subtracting C_1 from $C_2 + C_4$ we get C_6. If $r(yz) = 0 = k(y)$ and $i(y) + l(ye, z) = i(yz)$ then

$P(y, n, z, p+1) = P_3$, hence we get C_3 and C_4 by $(j=3)$ and $(j=4)$. Subtracting C_1 from C_3+C_4 we get

$$C_6': \qquad l(x, y)+l(xy, z) \equiv k(x)l(y, z)+l(x, yz) \pmod{r(xyz)}$$

Now C_6 holds iff (A6) holds.

3.4. Corollary. *If (S, e, r, i, k, l) is a $\mathcal{G}$-correct $\mathcal{I}$-scheme and e is a two-sided identity of S then P is associative iff*

(i) $k(x)k(y) \equiv k(xy) \pmod{r(xy)}$,

(ii) $l(x, y)+l(xy, z) \equiv k(x)l(y, z)+l(x, yz) \pmod{r(xyz)}$.

Proof. Straightforward.

3.5. Remark. Statements 3.1, 3.2 and 3.3 jointly characterize $\mathcal{S}^{\circ}$-correct $\mathcal{I}$-schemes by conditions (C1)—(C3), (IT1)—(IT3), (A0)—(A5). It is easy to satisfy all these conditions by a particular choice, e.g., $k(x) \equiv 1 \pmod{r(e)}$ for all $x \in S$, $l(x, y) = 0 = i(x)$ for all $x, y \in S$ and $r(x) = r \in \mathcal{N}$ for all $x \in S$, however, we are far from being able to describe all $\mathcal{S}^{\circ}$-correct $\mathcal{I}$-schemes, even under the tremendous simplification indicated by 3.4.

4. Universality of $\mathcal{I}$-schemes for $\mathcal{S}^{\circ}$

Let a be a left cancellable element of a semigroup T. In order to facilitate some of the further calculations, we introduce partial injections $T \to T$: $u \mapsto a^m \circ u$, $m \in Z$, as follows

$$a^m \circ u = \begin{cases} a^m u & \text{if } m > 0, \\ u & \text{if } m = 0, \\ v & \text{if } m < 0 \text{ and } a^{-m} v = u, \\ \emptyset & \text{if } m < 0 \text{ and there is no } v \text{ in } T \text{ with } a^{-m} v = u. \end{cases}$$

The following lemma states some easy calculation rules.

4.1. Lemma. *If $a^n \circ u \neq \emptyset$, then for any $v \in T$ and $m \in Z$ it holds*

(a) $a^n \circ (uv) = (a^n \circ u)v$,

(b) $a^m \circ (a^n \circ u) = a^{m+n} \circ u$.

Proof. (a) is clear for $n \geqq 0$. Assume $n < 0$ and denote $w = a^n \circ u$. Then $a^{-n} w = u$, hence $a^{-n} wv = uv$, hence $a^n \circ (uv) = wv = (a^n \circ u)v$.

(b) clearly holds if $m \geqq 0$ and $n \geqq 0$, as well as in the case $m = 0$ or $n = 0$. We shall consider in detail the three remaining cases.

Assume $m > 0$ and $n < 0$, denote $w = a^n \circ u$. If $m+n > 0$, then $a^m \circ (a^n \circ u) = a^m w = a^{m+n} a^{-n} w = a^{m+n} u = a^{m+n} \circ u$. If $m+n < 0$, then $u = a^{-n} w = a^{-n-m} a^m w$, hence $a^{m+n} \circ u = a^m w = a^m \circ (a^n \circ u)$. If $m+n=0$, then $a^m \circ (a^n \circ u) = a^m w = a^{-n} w = u = a^{m+n} \circ u$.

Assume $m<0$ and $n>0$. If $m+n>0$, then $a^{-m} a^{m+n} u = a^n u$, therefore $a^m \circ (a^n \circ u) = a^m \circ (a^n u) = a^{m+n} u = a^{m+n} \circ u$. If $m+n<0$ and $a^{m+n} \circ u \neq \emptyset$, say $a^{m+n} \circ u = v$, then $a^{-m-n} v = u$, hence $a^{-m} v = a^n u$, hence $a^{m+n} \circ u = v = a^m \circ (a^n u) = a^m \circ (a^n \circ u)$. If $m+n<0$ and $a^m \circ (a^n \circ u) \neq \emptyset$, say $a^m \circ (a^n \circ u) = a^m \circ (a^n u) = w$, then $a^{-m} w = a^n u$, hence $a^{-m-n} w = u$ by the n-fold cancellation of a on the left, therefore $a^{m+n} \circ u = w = a^m \circ (a^n \circ u)$. If $m+n=0$, then $a^{-m} u = a^n u$, therefore $a^m \circ (a^n \circ u) = a^m \circ (a^n u) = u = a^{m+n} \circ u$.

Assume $m<0$ and $n<0$. If $a^{m+n} \circ u \neq \emptyset$, say $a^{m+n} \circ u = v$, then $a^{-m-n} v = a^{-n} a^{-m} v = u$, hence $a^n \circ u = a^{-m} v$ and $a^m \circ (a^n \circ u) = v = a^{m+n} \circ u$. If $a^m \circ (a^n \circ u) \neq \emptyset$, say $a^m \circ (a^n \circ u) = w$, then $a^{-m} w = a^n \circ u$, hence $a^{-n} a^{-m} w = u = a^{-n-m} w$ and $a^{m+n} \circ u = w = a^m \circ (a^n \circ u)$.

4.2. Theorem. *Every $\mathcal{I}$-extension φ: $(T, a) \to (S, e)$ in $\mathcal{S}^\circ$ is right equivalent with some semiproduct π: $(P, a) \to (S, e)$: $(x, m) \mapsto x$ determined by a suitable $\mathcal{S}^\circ$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l).*

Proof. Let φ: $(T, a) \to (S, e)$ be an $\mathcal{I}$-extension in $\mathcal{S}^\circ$. Then we have, for any $u, v \in T$,

$$\varphi(u) = \varphi(v) \Leftrightarrow \exists m \in Z \quad (a^m \circ u = v). \tag{1}$$

The type function $r: S \to \mathcal{N}: x \mapsto \mathrm{Type}(f_x)$, f_x: $\varphi^{-1}(x) \to \varphi^{-1}(x)$: $u \mapsto au$, associated with this $\mathcal{I}$-extension satisfies by 2.2,

$$r(xy) \text{ divides } r(x) \text{ in } \mathcal{N} \text{ for any } x, y \in S \tag{2}$$

since $\mathcal{U}_x$ is taken homomorphically to $\mathcal{U}_{xy}$ by the multiplication on the right by any element $w \in \varphi^{-1}(y)$.

Further we construct a mapping ψ: $S \to T$ (not necessarily a homomorphism) selecting one point from each a-component, i.e., $\varphi\psi = 1_S$, as follows:

$$\text{if } r(x) = 0, \text{ then } a^{-1} \circ \psi(x) = \emptyset, \tag{3}$$

$$\text{if } r(x) \neq 0 \text{ and } xe = x, \text{ then } \psi(x) \in \varphi^{-1}(x) \text{ is arbitrary,} \tag{4}$$

$$\text{if } r(x) \neq 0 \text{ and } xe \neq x, \text{ then } \psi(x)a = \psi(xe)a, \tag{5}$$

(this is possible since multiplication by a on the right takes $\varphi^{-1}(x)$ onto $\varphi^{-1}(xe)$, as it follows, e.g., from 2.2).

Define a function i: $S \to Z$ by

$$a^{i(x)} \circ \psi(xe)a = \psi(x)a \quad \text{and} \quad \big(r(xe) \neq 0, \infty \Rightarrow 0 \leqq i(x) < r(xe)\big). \tag{6}$$

It follows from (4) and (5) that

(7) $$i(x)=0 \quad \text{if} \quad r(x)\neq 0, \quad \text{or} \quad r(x)=0 \quad \text{and} \quad x=xe.$$

A localizer $\lambda\colon T\to Z$ is defined by

(8) $$a^{\lambda(u)-i(\varphi(u))}\circ\psi\varphi(u)=u \quad \text{and} \quad \big(r(\varphi(u))\neq 0,\ \infty \Rightarrow \Rightarrow 0\leqq \lambda(u)-i(\varphi(u))<r(\varphi(u))\big).$$

A function $k\colon S\to Z$ is defined by

(9) $$a^{k(x)}\circ\psi(xe)=\psi(xe)a \quad \text{and} \quad \big(r(xe)\neq 0,\ \infty\Rightarrow 0\leqq k(x)<r(xe)\big).$$

Replacing x by xe in (9), we get $a^{k(xe)}\circ\psi(xe)=\psi(xe)a$, therefore

(10) $$k(xe)\equiv k(x)\ \big(\mathrm{mod}\, r(xe)\big).$$

Finally, a function $l\colon S\times S\to Z$ is defined by

(11) $$l(x,y)=\lambda\big(\psi(x)\psi(y)\big)-i(x)-k(x)i(y).$$

We shall show that

(12) $$\text{if} \quad a^{-1}\circ\psi(y)\neq\emptyset, \quad \text{then} \quad l(xe,y)\equiv l(x,y)\ \big(\mathrm{mod}\ r(xy)\big).$$

Indeed, we have then $\psi(y)=au$, for some $u\in T$, furthermore, by (a) and (6)

$$a^{i(x)}\circ\psi(xe)\psi(y)=a^{i(x)}\circ\psi(xe)au=\big(a^{i(x)}\circ\psi(xe)a\big)u=\psi(x)au=\psi(x)\psi(y),$$

therefore

$$\lambda\big(\psi(x)\psi(y)\big)\equiv\lambda\big(\psi(xe)\psi(y)\big)+i(x)\ \big(\mathrm{mod}\ r(xy)\big),$$

hence by (11)

$$i(x)+k(x)i(y)+l(x,y)\equiv i(xe)+k(xe)i(y)+l(xe,y)+i(x)\ \big(\mathrm{mod}\ r(xy)\big).$$

By (2), $r(xy)=r(xey)$ divides $r(xe)$, hence by (10), $k(xe)\equiv k(x)\ \big(\mathrm{mod}\ r(xy)\big)$. By (7), $i(xe)=0$, hence we get (12).

Let u, v be arbitrary elements of T with $\varphi(u)=x$, $\varphi(v)=y$. We shall show that

(13) $$\lambda(uv)=\begin{cases}\lambda(u)+k(x)\lambda(v)+l(xe,y) & \text{if} \quad a^{-1}\circ v\neq\emptyset,\\ \lambda(u)+k(x)\lambda(v)+l(x,y) & \text{if} \quad a^{-1}\circ v=\emptyset.\end{cases}$$

By (8),

(14) $$uv=\big(a^{\lambda(u)-i(x)}\circ\psi(x)\big)\big(a^{\lambda(v)-i(y)}\circ\psi(y)\big).$$

We split the consideration of this expression into three cases :

I. Assume $\lambda(v)>i(y)$. Then, taking into account (6) and (9) we have

$$uv=a^{\lambda(u)-i(x)}\circ\psi(x)a^{\lambda(v)-i(y)}\psi(y)=a^{\lambda(u)-i(x)}\circ\big(a^{i(x)}\circ\psi(xe)a^{\lambda(v)-i(y)}\psi(y)\big)=$$
$$=a^{\lambda(u)+k(x)(\lambda(v)-i(y))}\circ\psi(xe)\psi(y).$$

By (8), $\psi(xe)\psi(y) = a^{\lambda(\psi(xe)\psi(y))-i(xy)} \circ \psi(xy)$, hence

$$uv = a^{\lambda(u)+k(x)(\lambda(v)-i(y))+\lambda(\psi(xe)\psi(y))-i(xy)} \circ \psi(xy).$$

From this we get by (8),

$$\lambda(uv) \equiv \lambda(u)+k(x)\lambda(v)-k(x)i(y)+\lambda(\psi(xe)\psi(y)) \pmod{r(xy)},$$

hence by (11),

$$\lambda(uv) \equiv \lambda(u)+k(x)\lambda(v)-k(x)i(y)+i(xe)+k(xe)i(y)+l(xe, y) \pmod{r(xy)},$$

hence by (7) and (10) we get finally

$$\lambda(uv) \equiv \lambda(u)+k(x)\lambda(v)+l(xe, y) \pmod{r(xy)}.$$

Since $\lambda(v) > i(y)$ means that $a^{-1} \circ v \neq \emptyset$, we have proved (13) under the assumption.

II. Assume $\lambda(v)=i(y)$. Then (14) becomes

$$uv = a^{\lambda(u)-i(x)} \circ \psi(x)\psi(y) = a^{\lambda(u)-i(x)+\lambda(\psi(x)\psi(y))-i(xy)} \circ \psi(xy),$$

hence by (8) and (11),

$$\lambda(uv) \equiv \lambda(u)-i(x)+i(x)+k(x)i(y)+l(x, y) \pmod{r(xy)}, \tag{15}$$

which proves (13) in case $a^{-1} \circ v = \emptyset$. However, if $a^{-1} \circ v \neq \emptyset$, then by (12), $l(x, y) \equiv$ $\equiv l(xe, y) \pmod{r(xy)}$ and (15) gives (13) also in this case.

III. Assume $\lambda(v) < i(y)$. Then by (3) and (8), $r(y)=\infty$, hence by (7), $i(y)=0$, and by (12), $l(x, y) \equiv l(xe, y) \pmod{r(xy)}$. By (9) we get

$$a^{k(x)}ua = a^{k(x)} a^{\lambda(ua)-i(xe)} \circ \psi(\varphi(ua)) =$$

$$= a^{\lambda(ua)-i(xe)} a^{k(x)} \circ \psi(xe) = a^{\lambda(ua)-i(xe)} \circ \psi(xe)a = uaa,$$

hence for all $n \in Z$, $ua^n = a^{-k(x)\lambda(v)} ua^{n+\lambda(v)}$. Now an easy calculation yields

$$a^{-k(x)\lambda(v)} \circ uv = a^{-k(x)\lambda(v)} \circ u(a \circ (a^{-1} \circ v)) = ua^{-\lambda(v)} \circ v = u\psi(y) =$$

$$= a^{\lambda(u)-i(x)} \circ \psi(x)\psi(y) = a^{\lambda(u)-i(x)+\lambda(\psi(x)\psi(y))-i(xy)} \circ \psi(xy),$$

hence

$$\lambda(uv) \equiv \lambda(u)-i(x)+k(x)\lambda(v)+i(x)+k(x)i(y)+l(x, y) \equiv$$

$$\equiv \lambda(u)+k(x)\lambda(v)+l(xe, y) \pmod{r(xy)},$$

which proves (13) under $\lambda(v) < i(y)$.

We have $\mathrm{Type}(f_x)=r(x)$, hence by (3), (4), (5), (6), and (8), the assignment $u \mapsto (\varphi(u), \lambda(u))$ is a bijection establishing by (13) a right equivalence between $\varphi: (T, a) \to (S, e)$ and the semiproduct $\pi: (P, a) \to (S, e)$ determined by (S, e, r, i, k, l), whence the latter must be in $\mathcal{S}^{\circ}$.

5. Morphisms of $\mathscr{I}$-extensions

Let π: $(P, a)\to(S, e)$ and π': $(P', a')\to(S', e')$ be two $\mathscr{I}$-extensions in $\mathscr{S}^\circ$ determined by two $\mathscr{I}$-schemes (S, e, r, i, k, l) and (S', e', r', i', k', l'), respectively. By a morphism of π into π' we understand a couple (h, χ) of $\mathscr{S}^\circ$-homomorphisms h: $(P, a)\to(P', a')$, χ: $(S, e)\to(S', e')$ making the diagram

$$\begin{array}{ccc} (P,a) & \xrightarrow{\pi} & (S,e) \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle \chi} \\ (P',a') & \xrightarrow[\pi']{} & (S',e') \end{array}$$

commutative. A morphism (h, χ): $\pi\to\pi'$ is injective (surjective, bijective) iff both h and χ are so.

For a given $\mathscr{S}^\circ$-homomorphism χ: $(S, e)\to(S', e')$ we shall try to find "companion" $\mathscr{S}^\circ$-homomorphisms h: $(P, a)\to(P', a')$ such that (h, χ) is a morphism of π into π'. If such an h exists then it can always be expressed in the form $h=h_p$,

(1) $$h_p(x, m) \equiv (\chi(x), m+p(x)), \quad (x, m)\in P,$$

with the aid of a suitable "parameter" function p: $S\to Z$. The next theorem relates the properties of possible parameter functions to χ and the two $\mathscr{I}$-schemes.

5.1. Theorem. *Let* χ: $(S, e)\to(S', e')$ *be an* $\mathscr{S}^\circ$-*morphism. Then* p: $S\to Z$ *determines a mapping* h_p: $P\to P'$ *by* (1), *such that* (h_p, χ) *is a morphism of the* $\mathscr{I}$-*extension* π: $(P, a)\to(S, e)$ *determined by* (S, e, r, i, k, l) *into the extension* π': $(P', a')\to(S', e')$ *determined by* (S', e', r', i', k', l'), *iff the following conditions are satisfied for all* $x, y\in S$:

(H1) $r'(\chi(x))$ *divides* $r(x)$ *in* $\mathscr{N}$,
(H2) $r'(\chi(x))\equiv 0\Rightarrow i(x)+p(x)\geqq i'(\chi(x))$,
(H3) $k(x)\equiv k'(\chi(x)) \pmod{r'(\chi(x))}$,
(H4) $l(xe, y)+p(xy)\equiv p(x)+k(x)p(y)+l'(\chi(xe), \chi(y)) \pmod{r'(\chi(xy))}$,
(H5) *if* $r(y)=0=r'(\chi(y))$ *and* $i(y)+p(y)=i'(\chi(y))$, *then* $l(xe, y)-l(x, y)\equiv$ $\equiv l'(\chi(xe), \chi(y))-l'(\chi(x), \chi(y)) \pmod{r'(\chi(xy))}$,
(H6) *if* $r(y)=0$ *and* $(r'(\chi(y))\neq 0$ *or* $i(y)+p(y)\neq i'(\chi(y)))$, *then* $l(xe, y)\equiv$ $\equiv l(x, y) \pmod{r'(\chi(xy))}$.

Moreover, (h_p, χ) *is injective iff* χ *is so and*

(H7) $r(x)=r'(\chi(x))$ *for* $r(x)\neq 0$, $r'(\chi(x))=0$ *or* ∞ *for* $r(x)=0$,

and (h_p, χ) *is surjective iff* χ *is so and*

(H8) *if* $r(x)=0$ *then* $r'(\chi(x))=0$ *and* $i(x)+p(x)=i'(\chi(x))$.

Proof. Clearly, the formula (1) defines a function h_p from P to P' iff both (H1) and (H2) hold. Hence we shall further assume that $h_p\colon P\to P'$ is a mapping.

It is also clear from (1) that the couple (h_p, χ) is a morphism of π into π' iff $h_p\colon P\to P'$ is an $\mathcal{S}$-homomorphism, that is

$$h_p((x,m)(y,n)) \equiv h_p(x,m)h_p(y,n) \quad \text{for all} \quad (x,m),(y,n)\in P. \tag{2}$$

So (to prove the theorem) we have to prove (2) is equivalent with (H3)—(H6).

Assume (2). If $n>i(y)$, $n+p(y)>i'(\chi(y))$, then (2) becomes

$$\begin{aligned}(\chi(xy), m+k(x)n+l(xe,y)+p(xy)) \equiv\\ \equiv (\chi(x)\chi(y), m+p(x)+k'(\chi(x))(n+p(y))+l'(\chi(x)e',\chi(y))).\end{aligned} \tag{3}$$

If we compare (3) as it is and (3) with n replaced by $n+1$, we get (H3); (H3) and (3) imply (H4). If $r(y)=0=r'(\chi(y))$ and $i(y)+p(y)=i'(\chi(y))$, then (2) becomes, for $n=i(y)$,

$$\begin{aligned}(\chi(xy), m+k(x)n+l(x,y)+p(xy)) \equiv\\ \equiv (\chi(x)\chi(y), m+p(x)+k'(\chi(x))(n+p(y))+l'(\chi(x),\chi(y))),\end{aligned} \tag{4}$$

hence by (H3),

$$l(x,y)+p(xy) \equiv p(x)+k(x)p(y)+l'(\chi(x),\chi(y)) \pmod{r'(\chi(xy))}.$$

Subtracting this equality from (H4) we get the equality of (H5). If $r(y)=0$, but $r'(\chi(y))\neq 0$ or $i(y)+p(y)\neq i'(\chi(y))$, then for $n=i(y)$, (2) becomes

$$\begin{aligned}(\chi(xy), m+k(x)n+l(x,y)+p(xy)) \equiv\\ \equiv (\chi(x)\chi(y), m+p(x)+k'(\chi(x))(n+p(y))+l'(\chi(x)e',\chi(y))),\end{aligned} \tag{5}$$

hence

$$l(x,y)+p(xy) \equiv p(x)+k(x)p(y)+l'(\chi(x)e',\chi(y)) \pmod{r'(\chi(xy))},$$

which together with (H4) yields the equation of (H6).

To prove (2) from (H3)—(H6), we have to consider four cases:

$$(r(y)\neq 0 \quad\text{or}\quad n>i(y)) \quad\text{and}\quad (r'(\chi(y))\neq 0 \quad\text{or}\quad n+p(y)>i'(\chi(y))), \tag{6}$$

$$(r(y)=0 \quad\text{and}\quad n=i(y)) \quad\text{and}\quad (r'(\chi(y))=0 \quad\text{and}\quad n+p(y)=i'(\chi(y))), \tag{7}$$

$$(r(y)=0 \quad\text{and}\quad n=i(y)) \quad\text{and}\quad (r'(\chi(y))\neq 0 \quad\text{or}\quad n+p(y)>i'(\chi(y))), \tag{8}$$

$$(r(y)\neq 0 \quad\text{or}\quad n>i(y)) \quad\text{and}\quad (r'(\chi(y))=0 \quad\text{and}\quad n+p(y)=i'(\chi(y))). \tag{9}$$

Now, under (6), (2) is equivalent to (3) and the latter follows from (H3) and (H4). Under (7), (2) is equivalent to (4), while this follows from (H3), (H4), and (H5). Under (8), (2) is equivalent to (5), and this follows from (H3), (H4), and (H6). Case

(9) cannot occur since by (H1), $r'(\chi(y))=0$ implies $r(y)=0$, hence $n>i(y)$, while by (H2), $n+p(y)>i(y)+p(y)\geqq i'(\chi(y))$.

The rest of the theorem concerning injectiveness and surjectiveness of (h_p, χ) is obvious.

Let us call two $\mathscr{I}$-schemes (S, e, r, i, k, l) and (S', e', r', i', k', l') *$\mathscr{S}^\circ$-equivalent* if they yield equivalent $\mathscr{I}$-extensions in $\mathscr{S}^\circ$. The above theorem has the following straightforward

5.2. Corollary. *Two $\mathscr{I}$-schemes (S, e, r, i, k, l) and (S', e', r', i', k', l') are $\mathscr{S}^\circ$-equivalent iff there exist an $\mathscr{S}^\circ$-isomorphism $\chi: (S, e)\to(S', e')$ and a function $p: S\to Z$, such that*

(E1) $r(x) = r'(\chi(x))$,
(E2) $r(x) = 0 \Rightarrow i(x)+p(x) = i'(\chi(x))$,
(E3) $k(x) \equiv k'(\chi(x)) \pmod{r(x)}$,
(E4) $l(xe, y)+p(xy) \equiv p(x)+k(x)p(y)+l'(\chi(xe), \chi(y)) \pmod{r(xy)}$,
(E5) $r(y) = 0 \Rightarrow l(xe, y)-l(x, y) \equiv l'(\chi(xe), \chi(y))-l'(\chi(x), \chi(y)) \pmod{r(xy)}$.

Proof. We get (E1) and (E2) by replacing (H1) and (H2) by the stronger (H7) and (H8), (E3)—(E5) are obvious modifications of (H3)—(H5), and a version of (H6) is omitted since its assumption cannot occur here.

6. Special properties of $\mathscr{I}$-extensions

6.1. Statement. *An $\mathscr{I}$-scheme (S, e, r, i, k, l) determines an $\mathscr{I}$-extension $\pi: (P, a)\to(S, e)$ in the category of semigroups with identity iff*

(M1) *S is a semigroup with identity e,*
(M2) *if $r(xy)=0$ then $k(x)\geqq 0$ and $k(x)r(y)=0$ and $i(x)+k(x)i(y)+l(x, y) \geqq i(xy)$,*
(M3) *$r(xy)$ divides $r(x)$ in $\mathscr{N}$,*
(M4) *if $r(y) \neq \infty$ then $k(x)r(y) \equiv 0 \pmod{r(xy)}$,*
(M5) $k(e) \equiv 1 \pmod{r(e)}$,
(M6) $l(e, x) \equiv l(e, e) \pmod{r(x)}$,
(M7) *if $r(e) = 0$, then $i(e)=-l(e, e)$,*
(M8) $k(x)k(y) \equiv k(xy) \pmod{r(xy)}$,
(M9) $l(x, y)+l(xy, z) \equiv k(x)l(y, z)+l(x, yz) \pmod{r(xyz)}$.

Proof. The condition (M1) is clearly necessary. If it is satisfied, then (M2)—(M4) restates (C1)—(C3) of 3.1 (slightly simplified by (M1)), (M5), (M6) are exactly (IT1), (IT2) of 3.2. The condition (M7) is stronger than (IT3) of 3.2, its necessity follows from the fact that the identity (e, m) of P must be $(e, i(e))$ in the case $r(e)=0$. Indeed, we have $a(e, m)\equiv a$, where $a\equiv(e, 1-l(e, e))$ by 3.2, whence $m=-l(e, e)$, thus $i(e)\leqq -l(e, e)$. Now, if $(e, -l(e, e))$ is not the "endpoint" $(e, i(e))$, then $i(e)<-l(e, e)$ and there is an element $b=(e, -l(e, e)-1)$ in P. But then also $b^n=(e, -l(e, e)-n)$ is in P for all $n\geqq 1$, which is a contradiction with $r(e)=0$.

The conditions (M8), (M9) are exactly (i), (ii) of 3.4. We have proved the necessity of (M1)—(M9).

Assume now (M1)—(M9) fulfilled. Our $\mathcal{I}$-scheme is then clearly $\mathcal{S}^{\circ}$-correct and the only thing we have to show is that $(e, -l(e, e))$ is the identity of P:

$$(e, -l(e, e))(x, m) \equiv (x, -l(e, e)+k(e)m+l(e, x)) \equiv (x, m)$$

by (M5) and (M6),

$$(x, m)\,(e, -l(e, e)) \equiv (x, m-k(x)l(e, e)+l(x, e)) \equiv (x, m)$$

since $l(x, e)\equiv k(x)l(e, e) \pmod{r(x)}$ by (M9).

6.2. Remark. An $\mathcal{S}^{\circ}$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l) determines $\pi\colon (P, a)\to(S, e)$ with (P, a) finite iff S is finite and $0<r(x)<\infty$ for all $x\in S$. If P has an identity, then each a-component $\mathcal{U}_x$ of type $r(x)$, $x\in S$, decomposes into cycles of length $q(x)=\dfrac{r(x)}{k(x)\wedge r(x)}$ of the right inner translation of P by a. The integer $r(x)$ divides $k(x)q(x)$, hence $p(x)=\dfrac{k(x)q(x)}{r(x)}=\dfrac{k(x)}{k(x)\wedge r(x)}$ is an integer and $p(x)\wedge q(x)=\dfrac{k(x)}{k(x)\wedge r(x)}\wedge\dfrac{r(x)}{k(x)\wedge r(x)}=1$. If, for some $y\in S$, we have $q(y)=q(x)$ and $xy=y$ or $yx=y$, then $k(x)k(y)\equiv k(y) \pmod{r(y)}$, and using $k(x)=\dfrac{p(x)r(x)}{q(x)}$, we get $k(x)k(y)-k(y)=\dfrac{p(x)r(x)}{q(x)}\cdot\dfrac{p(y)r(y)}{q(y)}-\dfrac{p(y)r(y)}{q(y)}=p(y)r(y)\times\left[\dfrac{p(x)r(x)-q(x)}{q^2(x)}\right]\equiv 0 \pmod{r(y)}$. It follows that $p(y)\left[\dfrac{p(x)r(x)-q(x)}{q(x)}\right]$ is an integer, and since $p(y)\wedge q(y)=1$, $\dfrac{p(x)r(x)-q(x)}{q^2(x)}$ is an integer, too.

Conversely, if $\dfrac{p(x)r(x)-q(x)}{q^2(x)}$ is an integer, then for arbitrary $r(y)$, $p(y)$ with $p(y)\wedge q(y)=1$, and $k(y)=\dfrac{p(y)r(y)}{q(x)}$ it holds

$$k(x)k(y)-k(y) = p(y)r(y)\left[\frac{p(x)r(x)-q(x)}{q^2(x)}\right]\equiv 0 \pmod{r(y)}.$$

This property has been used in [3] and [4] for a special kind of constructions of monoids, starting with S a left zero semigroup with an identity 1 adjoined, $q \in N^+$ and $r, p \colon S \to N^+$ (N^+ means the positive integers) such that $r(x)$ divides $r(1)$, q divides $r(x)$, $q \wedge p(x) = 1$, and $\dfrac{p(x)r(x)-q}{q^2}$ is an integer, for all $x \in S$. These constructions amount to those of $\mathcal{I}$-extensions determined by (S, e, r, i, k, l) with $e=1$, $i(x)=0$, $k(x)=\dfrac{p(x)r(x)}{q}$ and $l(x, y)=0$, for all $x, y \in S$.

6.3. Statement. *An $\mathcal{S}^\circ$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l) determines an $\mathcal{I}$-extension $\pi \colon (P, a) \to (S, e)$ with (P, a) commutative iff*

(AB1) *S is commutative,*
(AB2) $l(x, y) \equiv l(y, x) \pmod{r(xy)}$,
(AB3) $k(x) \equiv 1 \pmod{r(x)}$.

Proof. If (AB1)—(AB3) hold, then e is an identity in S and the two products

$$(x, m)(y, n) \equiv \big(xy, m + k(x)n + l(x, y)\big) \tag{1}$$

$$(y, n)(x, m) \equiv \big(yx, n + k(y)m + l(y, x)\big) \tag{2}$$

are equal for any $(x, m), (y, n) \in P$.

Conversely, if (1) and (2) are equal for any $(x, m), (y, n) \in P$, then (AB1) holds,

$$m + k(x)n + l(x, y) \equiv n + k(y)m + l(y, x) \pmod{r(xy)}, \tag{3}$$

and also, replacing n by $n+1$,

$$m + k(x)n + k(x) + l(x, y) \equiv n + 1 + k(y)m + l(y, x) \pmod{r(xy)}, \tag{4}$$

hence subtracting (3) from (4) we get $k(x) \equiv 1 \pmod{r(xy)}$, for any $y \in S$, which is equivalent to (AB3). By (3) and (AB3) we get (AB2).

6.4. Statement. *Let (S, e, r, i, k, l) be an $\mathcal{S}^\circ$-correct $\mathcal{I}$-scheme. The semigroup $(P, a) = \mathcal{I}(S, e, r, i, k, l)$ is right cancellative iff*

(RC1) *S is right cancellative,*
(RC2) $\big(r(x) \neq 0 \Rightarrow r(xy) = r(x)\big)$ *and* $\big(r(x) = 0 \Rightarrow r(xy) = 0$ *or* $\infty\big)$.

Proof. (RC2) means that right inner translations of P take each a-component $\mathcal{U}_x$ into another component injectively, (RC1) ensures that distinct a-components are taken to distinct a-components.

6.5. Statement. *For an $\mathcal{S}^\circ$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l), the semigroup $(P, a)=\mathcal{I}(S, e, r, i, k, l)$ is left cancellative iff*

(LC1) *S is left cancellative,*
(LC2) *if some $r(y) \neq 1$, then $k(x) \neq 0$ for all $x \in S$,*
(LC3) *if $0 < r(xy) < \infty$, then $k(x)r(y) = k(x) \vee r(xy)$,*
(LC4) *if $r(y) = 0$, then for every $n \neq i(y)$,*

$$k(x)i(y)+l(x, y) \not\equiv k(x)n+l(xe, y) \pmod{r(xy)}.$$

Proof. If P is left cancellative, then so is its quotient S. Under (LC1), P is left cancellative iff

(5) $$(x, m)(y, n_1) \equiv (x, m)(y, n_2) \Rightarrow (y, n_1) \equiv (y, n_2).$$

Under the assumption that

(6) $$r(y) \neq 0 \quad \text{or} \quad n_1 \neq i(y) \neq n_2,$$

(5) is equivalent to

(7) $$k(x)(n_1-n_2) \equiv 0 \pmod{r(xy)} \Rightarrow (n_1-n_2) \equiv 0 \pmod{r(y)}.$$

Since the difference n_1-n_2 ranges over the whole Z, (5) is equivalent to

(8) $$k(x)n \equiv 0 \pmod{r(xy)} \Rightarrow n \equiv 0 \pmod{r(y)},$$

for all $n \in Z$.

We shall prove (8) to be equivalent to the conjunction of (LC2) and (LC3).

Assume (8). If $k(x)=0$ for some $x \in S$, then from (8) it follows that $r(y)=1$ for all $y \in S$, thus (LC2) holds.

Of course, if $r(y)=1$ for all $y \in S$, the $\mathcal{I}$-extension is improper, $P \cong S$, so we further exclude this case from our consideration. If $0<r(xy)<\infty$, then also $0<r(y)<\infty$, since under $r(y)=0$ or ∞ it would be $n\equiv 0 \pmod{r(y)}$ iff $n=0$, while $k(x)n \equiv 0 \pmod{r(xy)}$ for $n=r(xy)\neq 0$, a contradiction to (8). By (C3) of 3.1, $r(xy)$ divides $k(x)r(y)$, hence $k(x)\vee r(xy)$ divides $k(x)r(y)$. If $\frac{k(x)r(y)}{k(x)\vee r(xy)}=p\neq 1$, then $r(y)$ does not divide $n=\frac{r(y)}{p}=\frac{r(xy)}{k(x)\wedge r(xy)}$, while $r(xy)$ divides $k(x)n=\frac{k(x)r(xy)}{k(x)\wedge r(xy)}$, again a contradiction to (8). Thus (8) implies (LC3).

Conversely, assume (LC2) and (LC3). If $k(x)=0$ then by (LC2) we have $r(y)=1$ for every $y \in S$ and (P, a) is isomorphic to (S, e), and by (LC1), P is left-cancellative. Let us therefore assume that $k(x)\neq 0$ for every $x \in S$. If $r(xy)=0$ or ∞ then $k(x)n\equiv 0 \pmod{r(xy)}$ iff $n=0$, hence we get $n\equiv 0 \pmod{r(y)}$ trivially. If $0<r(xy)<\infty$ then by (LC3), $k(x)r(y)=\frac{k(x)r(xy)}{k(x)\wedge r(xy)}$, whence $r(xy)=$

$=r(y)(k(x)\wedge r(xy))$. Thus $r(xy)$ divides $k(x)n$ iff $r(y)$ divides $\frac{k(x)}{k(x)\wedge r(xy)}n$. Since $r(y)\wedge\frac{k(x)}{k(x)\wedge r(xy)}=\frac{r(xy)}{k(x)\wedge r(xy)}\wedge\frac{k(x)}{k(x)\wedge r(xy)}=1$, $r(y)$ dividing $\frac{k(x)n}{k(x)r(xy)}$ must divide n. Assuming $r(y)=0$ and $i(y)=n_1$ or n_2, we have that (5) is equivalent to $n\neq i(y)\Rightarrow(x, m)(y, n)\not\equiv(x, m)(y, i(y))$, which is equivalent to (LC4).

Recall that a semigroup P is called *right reductive* if, for any $x, y\in P$,

$$\text{(RR)}\qquad \forall z\in P\ (xz = yz) \Rightarrow x = y,$$

i.e., the family of the right inner translations of P separates points.

6.6. Statement. *For an $\mathcal{S}^\circ$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l), the semigroup $(P, a)=\mathcal{I}(S, e, r, i, k, l)$ is right reductive iff*

(RR1) $(r(x)\neq 0\Rightarrow r(xe) = r(x))$ *and* $(r(x)=0\Rightarrow r(xe) = 0$ or $\infty)$,

(RR2) *if* $x\neq y$, $xe=ye$, *and* $k(x)\equiv k(y)\ (\mathrm{mod}\ r(x))$ *for* $x, y\in S$, *then for every* $q\in Z$ *there exists* $z\in S$ *such that either*

$$l(x, z)-l(y, z)\not\equiv q\ (\mathrm{mod}\ r(xz))\quad \textit{and}\quad r(z) = 0,$$

or

$$l(xe, z)-l(ye, z)\not\equiv q\ (\mathrm{mod}\ r(xz)).$$

Proof. Assume P right reductive. If $(x, m_1)\not\equiv(x, m_2)$ then, for some $(z, p)\in P$, $(x, m_1)(z, p)\not\equiv(x, m_2)(z, p)$. Hence if $m_1\not\equiv m_2\ (\mathrm{mod}\ r(x))$, then also $m_1\not\equiv m_2\ (\mathrm{mod}\ r(xz))$, and since $r(xz)$ divides $r(xe)$, $m_1\not\equiv m_2\ (\mathrm{mod}\ r(xe))$. The condition (RR1) follows.

Let now $x\neq y$, $xe=ye$, and $k(x)\equiv k(y)\ (\mathrm{mod}\ r(x))$. It is $r(x)=r(y)$, by $xe=ye$ and (RR1). For any $q\in Z$, there are (x, m), $(y, n)\in P$ with $n-m=q$. Since $(x, m)\not\equiv(y, n)$, there exists $(z,p)\in P$ such that $(x, m)(z, p)\not\equiv(y, n)(z, p)$. This means that either

$$q = n-m\not\equiv l(x, z)-l(y, z)\ (\mathrm{mod}\ r(xz))$$

or

$$q = n-m\not\equiv l(xe, z)-l(ye, z)\ (\mathrm{mod}\ r(xz)),$$

according to whether or not, $r(z)=0$ and $p=i(z)$. This proves (RR2) necessary.

Assume now (RR1) and (RR2). Let $(x, m)\not\equiv(y, n)$. If $x=y$, then by (RR1), $(x, m)a\not\equiv(y, n)a$, where $a\equiv(e, 1-l(e, e))$. If $x\neq y$ and $xe\neq ye$, then again $(x, m)a\not\equiv(y, n)a$. Let $x\neq y$, $xe=ye$, and assume that $(x, m)(z, p)\equiv(y, n)(z, p)$ for all $(z, p)\in P$. It follows that $n-m\equiv(k(x)-k(y))p+l(xe, z)-l(ye, z)\ (\mathrm{mod}\ r(xz))$ for all $p>i(z)$, hence $k(x)\equiv k(y)\ (\mathrm{mod}\ r(xz))$ for all $z\in S$. Therefore

$$n-m\equiv l(x, z)-l(y, z)\ (\mathrm{mod}\ r(xz)),\quad \text{for}\quad r(z) = 0\quad \text{and}\quad p = i(z),$$

and

$$n-m \equiv l(xe, z)-l(ye, z) \pmod{r(xz)} \quad \text{otherwise},$$

a contradiction to (RR2).

6.7. Statement. *For an $\mathcal{S}^{\circ}$-correct $\mathcal{I}$-scheme (S, e, r, i, k, l), the semigroup $(P, a)=\mathcal{I}(S, e, r, i, k, l)$ is a group iff*

(G1) *S is group,*
(G2) $r(e)\neq 0$.

Proof. The necessity of the conditions is obvious. Assume next (G1) and (G2). We show that $(e, -l(e, e))$ is an identity of P in the same way as in the proof of 6.1. By (G1) and (C2) of 3.1, $r(x)=r(e)$ for all $x\in S$. The proof will be completed by showing that

$$(x, m)^{-1} \equiv (x^{-1}, -l(e, e)-k(x^{-1})m-l(x^{-1}, x)).$$

First, since by (G2) and (C2) we have $r(x)\neq 0$, there exists an element of this form in P. By (A0) and (IT1), $k(x)k(x^{-1})\equiv k(e)\equiv 1 \pmod{r(e)}$, and

$$(x, m)(x, m)^{-1} \equiv (e, m-k(x)l(e, e)-k(x)k(x^{-1})m-k(x)l(x^{-1}, x)+l(x, x^{-1})) \equiv$$
$$\equiv (e, -k(x)l(e, e)-k(x)l(x^{-1}, x)+l(x, x^{-1})) \equiv (e, -l(e, e)),$$

since by (A1) of 3.3,

$$l(x, e)\equiv k(x)l(e, e), \quad l(x, x^{-1})+l(e, x) \equiv k(x)l(x^{-1}, x)+l(x, e),$$

and thus by (IT2) of 3.2,

$$-l(x, e)-k(x)l(x^{-1}, x)+l(x, x^{-1}) \equiv -l(e, e).$$

Finally,

$$(x, m)^{-1}(x, m) \equiv (e, -l(e, e)-k(x^{-1})m-l(x^{-1}, x)+k(x^{-1})m+l(x^{-1}, x)) \equiv$$
$$\equiv (e, -l(e, e)).$$

7. Transextension in monoids factorize through $\mathcal{I}$-extensions

The aim of this concluding section is to show that every transextension in the category $\mathcal{M}^{\circ}$ of pointed semigroups with identity factorizes through some $\mathcal{I}$-extension and that there is a biggest one among such $\mathcal{I}$-extensions. Thus $\mathcal{I}$-extensions form, in this category, an important intermediary step in the constructions of transextensions, to be followed by an extension of a different kind using the decomposition of connected components of a translation into levels.

Unfortunately, we cannot do the same in $\mathcal{S}^{\circ}$, the assumption of the presence of an identity in T cannot be relaxed as is shown by the example following the statement.

7.1. Statement. *Let T be a semigroup with identity 1. If for some $a\in T$, the a-connectedness*

$$x \sim y \Leftrightarrow \exists m, n\in\mathcal{N}(a^m x = a^n y)$$

is a congruence on T, then so is its refinement

$$x \approx y \Leftrightarrow \exists m\in\mathcal{N}(a^m x = a^m y),$$

and the quotient $(T/\approx, a/\approx)$ is the biggest $\mathcal{I}$-extension of $(T/\sim, a/\sim)$ extendable to T.

Proof. Let $\sim$ be a congruence on T, let $x\in T$. From $1\sim a$ we get $x\sim xa$, hence for some $k=k(x)\geqq 0$, we have either $a^k x\approx xa$ or $x\approx a^k xa$. We next show that if $a^k x\approx xa$ and $x\sim y$, then also $a^k y\approx ya$. Indeed, let us have, say, $a^p a^k x = a^p xa$ and $a^m x=a^n y$. Then $a^p a^n ya=a^p a^m xa=a^m a^p xa=a^m a^p a^k x=a^p a^k a^m x= a^p a^k a^n y$, hence $a^k y\approx ya$. Similarly we show that if $x\approx a^k xa$, then also $y\approx a^k ya$ for any $y\sim x$. For, if $a^p x=a^p a^k xa$ and $a^m x=a^n y$, then $a^p a^n y=a^p a^m x=a^m a^p x= a^m a^p a^k xa=a^p a^k a^n ya=a^p a^n a^k ya$, hence $y\approx a^k ya$.

The equivalence $\approx$, as well as $\sim$, is clearly a right congruence. We prove $\approx$ to be a left congruence as follows. Take some $z\in T$ and assume first $a^k z\approx za$, for some $k\geqq 0$. Then we have $a^{pk}z\approx za^p$ for all $p\geqq 0$. If $x\approx y$, say $a^p x=a^p y$, then $a^{pk}zx\approx za^p x=za^p y\approx a^{pk}zy$, hence $zx\approx zy$. Assume now $z\approx a^k za$, for some $k\geqq 0$. Then $z\approx a^{pk}za^p$ for all $p\geqq 0$. Again, if $x\approx y$, say $a^p x=a^p y$, then $zx\approx a^{pk}za^p x= a^{pk}za^p y\approx zy$.

The last assertion of the statement follows from the fact that $\approx$ is obviously the least congruence on T making $a/\approx$ left cancellable in $T/\approx$.

7.2. Example. Consider the semigroup S

	a	b	c	d	f	g	h	k	m	n	p	q
a	b	a	d	c	g	f	k	m	k	p	n	q
b	a	b	c	d	f	g	m	k	m	n	p	q
c	g	f	q	q	q	q	n	n	p	q	q	q
d	f	g	q	q	q	q	p	p	n	q	q	q
f	g	f	q	q	q	q	p	n	p	q	q	q
g	f	g	q	q	q	q	n	p	n	q	q	q
h	q	q	q	q	q	q	q	q	q	q	q	q
k	q	q	q	q	q	q	q	q	q	q	q	q
m	q	q	q	q	q	q	q	q	q	q	q	q
n	q	q	q	q	q	q	q	q	q	q	q	q
p	q	q	q	q	q	q	q	q	q	q	q	q
q	q	q	q	q	q	q	q	q	q	q	q	q

An easy computation gives that a-connectedness is a congruence on S. On the other hand, $x \approx y$ iff either $x=y$ or $\{x, y\}=\{h, m\}$. Further $c \cdot h=n$, $c \cdot m=p$ and $n \not\approx p$. Thus $\approx$ is not a congruence on S.

References

[1] P. Goralčík, V. Koubek, Translational extensions of semigroups, in: *Algebraic Theory of Semigroups* (Proc. Conf. Szeged 1976), Colloq. Math. Soc. J. Bolyai, vol 20, North-Holland (Amsterdam, 1979); 173—218.

[2] P. A. Grillet, Building semigroups from groups (and reduced semigroups), *Semigroup Forum,* **4** (1972), 327—334.

[3] M. Münzová, Transformations determining uniquely a monoid, *Comment. Math. Univ. Carolinae,* **11** (1970), 595—618.

[4] M. Münzová, Transformations determining uniquely a monoid. II, *Comment. Math. Univ. Carolinae,* **15** (1974), 311—333.

[5] M. Novotný, Über Abbildungen von Mengen, *Pacific J. Math.,* **13** (1963), 1359—1369.

(M. D.)
ELECTROENGINEERING FACULTY
TECHNICAL UNIVERSITY OF PRAGUE
PRAGUE, CZECHOSLOVAKIA

(P. G. and V. K.)
FACULTY OF MATHEMATICS AND PHYSICS
CHARLES UNIVERSITY
PRAGUE, CZECHOSLOVAKIA

Acta Sci. Math. 44 (1982), 239—253

Idempotent distributive semirings. I

F. PASTIJN and A. ROMANOWSKA

0. Introduction

A *semiring* $(S, +, \cdot)$ is an algebra of type $\langle 2,2 \rangle$ where

(0.1) $$(S, +) \text{ is a semigroup}$$

(0.2) $$(S, \cdot) \text{ is a semigroup}$$

(0.3) $$(a+b)c = ac+bc \quad \text{and} \quad a(b+c) = ab+ac \quad \text{for all} \quad a, b, c \in S.$$

We shall always write ab instead of $a \cdot b$, and we suppose that $\cdot$ links stronger than $+$; with this assumption we can omit brackets.

The semiring $(S, +, \cdot)$ is *idempotent* if

(0.4) $$a+a = a = aa \quad \text{for all} \quad a \in S,$$

i.e. the reducts $(S, +)$ and $(S, \cdot)$ are bands. We say that the semiring $(S, +, \cdot)$ is *distributive* if

(0.5) $$ab+c = (a+c)(b+c) \quad \text{and} \quad a+bc = (a+b)(a+c) \quad \text{for all} \quad a, b, c \in S.$$

The purpose of this paper is to clarify the structure of idempotent distributive semirings which henceforth will be called ID-semirings.

We refer the reader to [6], [7] for a construction and a classification of bands. We shall assume that the reader is acquainted with the definition of the Płonka sum of a semilattice ordered system of algebras [8].

We shall now list diverse examples which supply the motivation for our investigations.

Result 1 ([1], Corollary 8). *Let $(I, \cdot)$ and $(\Lambda, \cdot)$ be any semigroups. On $I \times \Lambda = S$ we define an addition and a multiplication in the following way. For all $(i, \lambda), (j, \mu) \in S$, let*

(0.6) $$(i, \lambda)+(j, \mu) = (i, \mu), \quad \textit{and} \quad (i, \lambda)(j, \mu) = (ij, \lambda\mu).$$

Received October 21, 1980.

Then $(S, +, \cdot)$ *is a semiring for which the additive reduct is a rectangular band; conversely, every semiring for which the additive reduct is a rectangular band is isomorphic to a semiring constructed in this fashion.*

We give a direct proof for the sake of completeness.

Proof. Let S be a semiring for which the additive reduct is a rectangular band, and let us introduce the relations $\mathcal{R}$ and $\mathcal{L}$ on S in the following way:

$$a\mathcal{R}b \text{ if and only if } a+b=b \text{ and } b+a=a,$$

$$a\mathcal{L}b \text{ if and only if } a+b=a \text{ and } b+a=b.$$

It should be evident that $\mathcal{R}$ and $\mathcal{L}$ are equivalence relations and that $\mathcal{L}\cap\mathcal{R}$ is the equality. Let $S/\mathcal{L}=I$ and $S/\mathcal{R}=\Lambda$, and for any $a\in S$, let $i_a\in S/\mathcal{L}=I$ denote the $\mathcal{L}$-class of a, and $\lambda_a\in S/\mathcal{R}=\Lambda$ the $\mathcal{R}$-class containing a. By the foregoing the mapping $\varphi\colon S\to I\times\Lambda$, $a\to(i_a, \lambda_a)$ is an injective mapping. Let (i, λ) be any element of $I\times\Lambda$. Then $i=i_a$ and $\lambda=\lambda_b$ for some $a, b\in S$; since $(S, +)$ is a rectangular band it follows that $i_a=i_{b+a}$ and $\lambda_b=\lambda_{b+a}$. We conclude that

$$(i, \lambda) = (i_a, \lambda_b) = (i_{b+a}, \lambda_{b+a}) = (b+a)\varphi.$$

Thus the injection φ is in fact a bijection onto $I\times\Lambda$.

From (0.3) it follows that $\mathcal{L}$ and $\mathcal{R}$ are congruence relations. From the above we then have $S\cong S/\mathcal{L}\times S/\mathcal{R}=I\times\Lambda$ by φ.

We consider the semiring $I=S/\mathcal{L}$. Let i_a, i_b be any elements of I. Then

$$i_a+i_b = i_{a+b} = i_b$$

since $(a+b)+b=a+b$ and $b+(a+b)=b$. Hence the additive reduct of I is a right-zero semigroup. Analogously, the additive reduct of Λ is a left-zero semigroup. We conclude that every semiring for which the additive reduct is a rectangular band may be constructed as stated above.

The direct part is obvious.

Corollary 1. *Let* $(S, +, \cdot)$ *be a semiring for which the additive reduct is a rectangular band. Then* $(S, +, \cdot)$ *is distributive if and only if* $(S, +, \cdot)$ *is idempotent.*

Obviously the semiring $(S, +, \cdot)=(I\times\Lambda, +, \cdot)$ of Result 1 is an ID-semiring if and only if the semigroups $(I, \cdot)$ and $(\Lambda, \cdot)$ are bands.

We shall say that an ID-semiring is a *rectangular* [*normal, left-zero, ...*] *semiring* if and only if both the reducts are rectangular [normal, left-zero, ...] bands. A *nest* $(S, +, \cdot)$ is an algebra of type $\langle 2,2\rangle$ which satisfies

(0.7) $$a+b = b$$

and

$$ab = a \tag{0.8}$$

for all $a, b \in S$ [2] [18]. It is easy to see that a nest is necessarily an ID-semiring where the additive reduct is a right-zero band and the multiplicative reduct a left-zero band. Using the notations of Result 1 we can state that any nest is of the form $(I \times \Lambda, +, \cdot)$ where $|I|=1$ and where $(\Lambda, \cdot)$ is a left-zero band. A *dual nest* $(S, +, \cdot)$ is an ID-semiring where

$$a+b = a \tag{0.9}$$

and

$$ab = b \tag{0.10}$$

holds for all $a, b \in S$. With the notations of Result 1 we have that a dual nest is of the form $(I \times \Lambda, +, \cdot)$ where $|\Lambda|=1$ and where $(I, \cdot)$ is a right-zero band. A left-zero semiring is of the form $(I \times \Lambda, +, \cdot)$ where $|\Lambda|=1$ and where $(I, \cdot)$ is a left-zero band, whereas a right-zero semiring is of the form $(I \times \Lambda, +, \cdot)$ where $|I|=1$ and where $(\Lambda, \cdot)$ is a right-zero band.

Corollary 2. *A semiring is rectangular if and only if it is the direct product of a left-zero semiring, a right-zero semiring, a nest and a dual nest.*

Proof. Let $(S, +, \cdot)$ be a rectangular semiring. Since the additive reduct is a rectangular band, $(S, +, \cdot)=(I \times \Lambda, +, \cdot)$ can be constructed as in Result 1, where $(I, \cdot)$ and $(\Lambda, \cdot)$ are rectangular bands. It follows that $(I, \cdot)$ is of the form $(I_1 \times I_2, \cdot)$, where

$$(i_1, i_2)(j_1, j_2) = (i_1, j_2)$$

for all $(i_1, i_2), (j_1, j_2) \in I_1 \times I_2$, whereas $(\Lambda, \cdot)$ is of the form $(\Lambda_1 \times \Lambda_2, \cdot)$, where

$$(\lambda_1, \lambda_2)(\mu_1, \mu_2) = (\lambda_1, \mu_2).$$

Therefore the rectangular semiring $(S, +, \cdot)$ must be isomorphic to $(I_1 \times I_2 \times \Lambda_1 \times \times \Lambda_2, +, \cdot)$, where

$$(i_1, i_2, \lambda_1, \lambda_2)+(j_1, j_2, \mu_1, \mu_2) = (i_1, i_2, \mu_1, \mu_2),$$

$$(i_1, i_2, \lambda_1, \lambda_2)(j_1, j_2, \mu_1, \mu_2) = (i_1, j_2, \lambda_1, \mu_2)$$

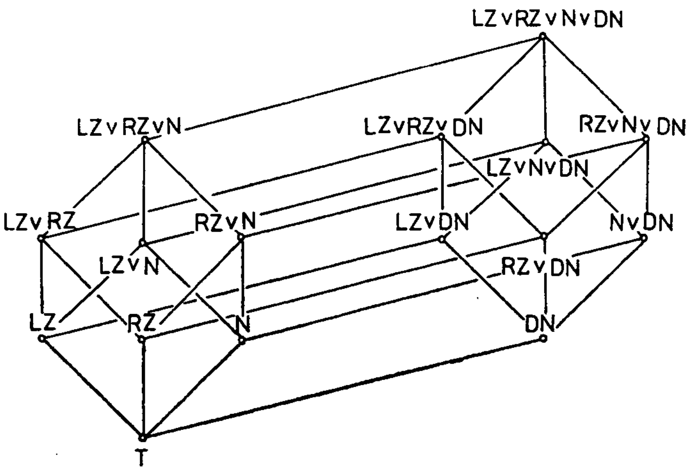
for all $i_1, j_1 \in I_1$, $i_2, j_2 \in I_2$, $\lambda_1, \mu_1 \in \Lambda_1$, and $\lambda_2, \mu_2 \in \Lambda_2$. Hence $(S, +, \cdot)$ is isomorphic to the direct product of the left-zero semiring I_1, the right-zero semiring Λ_2, the nest Λ_1 and the dual nest I_2. Conversely, any such direct product must yield a rectangulan semiring.

Let us consider the variety N of nests, DN of dual nests, LZ of left-zero semirings and RZ of right-zero semirings. Since all the equivalence relations of algebras in these classes are congruence relations, all subdirectly irreducible algebras contain exactly two elements. Hence the varieties *LZ, RZ, N* and *DN* are atoms in the lattice of varieties of rectangular semirings. It is easy to see that an arbitrary algebra in a join of some of the varieties *N, DN, LZ* and *RZ* is a direct product of algebras in the component varieties. This follows from [11] Theorem 1 (the suitable polynomials are just xy or $x+y$). Hence any subdirectly irreducible algebra in a join of some of the varieties *N, DN LZ* and *RZ* belongs to one of the component varieties. These remarks imply the following description of the lattice of varieties of rectangular semirings.

Corollary 3. *There are exactly* 16 *varieties of rectangular semirings. They form a Boolean lattice.*

The variety $LZ \vee RZ \vee N \vee DN$ *is equal to the variety of rectangular semirings. Each one of the "join" subvarieties of the variety of rectangular semirings is defined by the identities for rectangular semirings and by one additional identity:*

$LZ \vee RZ$	$x+y = xy$
$LZ \vee N$	$xy = x$
$LZ \vee DN$	$x+y = x$
$RZ \vee N$	$x+y = y$
$RZ \vee DN$	$xy = y$
$N \vee DN$	$x+y = yx$
$LZ \vee RZ \vee N$	$xy+x = x$
$LZ \vee RZ \vee DN$	$x+yx = x$
$LZ \vee N \vee DN$	$x+xy = x$
$RZ \vee N \vee DN$	$yx+x = x$



Proof. The first part of the corollary is clear from the previous remarks. The second part can be easily proved using standard methods. Let us show for example that the variety $LZ \vee RZ$ is defined by the identities for rectangular semirings and the identity $x+y=xy$. If the ID-semiring $(S, +, \cdot)$ belongs to the variety $LZ \vee RZ$, it is a product of a left-zero semiring and a right-zero semiring. So it is easy to see that the identity $x+y=xy$ is satified. Conversely, if a rectangular semiring $(S, +, \cdot)$ satisfies this identity, then by Corollary 2, we conclude that it is in fact isomorphic to the direct product of a left-zero semiring and a right-zero semiring.

Let us again consider the semiring $(S, +, \cdot)=(I\times \Lambda, +, \cdot)$ of Result 1. Then $(S, +)$ is a left-zero [right-zero] band if and only if $|\Lambda|=1$ $[|I|=1]$, and if this is the case then $(S, +, \cdot)$ may be identified with $(I, +, \cdot)$ $[(\Lambda, +, \cdot)]$. It follows that any semigroup may be represented as the multiplicative reduct of a semiring. In particular, any band may be represented as the multiplicative reduct of an ID-semiring. This fact has been observed by many authors (see e.g. [20] Example 2.3.4).

A semiring $(S, +, \cdot)$ is called a *mono-semiring* if

$$a+b = ab \tag{0.11}$$

holds for all $a, b \in S$, that is, if the two operations $+$ and $\cdot$ coincide. If $(S, +, \cdot)$ is an idempotent semiring which is also a mono-semiring, then $(S, +)$ and $(S, \cdot)$ must be normal bands, and $(S, +, \cdot)$ needs to be distributive. Conversely, any normal band is the additive (or multiplicative) reduct of an ID-mono-semiring ([20], Theorem 4.4.2). Left-zero semirings and right-zero semirings need to be mono-semirings. A rectangular mono-semiring is the direct product of a left-zero semiring and a right-zero semiring. An ID-mono-semiring is the Płonka sum (in the sense of [8]) of a semilattice ordered system of rectangular mono-semirings [19].

An idempotent semiring $(S, +, \cdot)$ where the reducts $(S, +)$ and $(S, \cdot)$ are semilattices has been called a $\cdot$*-distributive bisemilattice* in [4][12][13][14]. If moreover $(S, +, \cdot)$ is distributive, then $(S, +, \cdot)$ is called a *distributive bisemilattice* [15] (*distributive quasilattices* in [9]). Here the distributive lattices form the main example. Another particular case consists of the mono-semirings where the two reducts are semilattices: they will be called *mono-bisemilattices*. We remark that mono-bisemilattices could as well be identified with semilattices. A distributive bisemilattice is the Płonka sum of a semilattice ordered system of distributive lattices [9].

1. Normal semirings

In this section we investigate the structure of normal semirings, i.e. semirings satisfying the identities

(1.1) $$xyzw = xzyw,$$

(1.2) $$x+y+z+w = x+z+y+w.$$

We shall see later (Theorem 2.3) that any ID-semiring can "in principle" be constructed from normal semirings.

First, let $(S, +, \cdot)$ be any ID-semiring. On S we introduce the equivalence relations $\overset{+}{\mathscr{L}}, \overset{+}{\mathscr{R}}, \dot{\mathscr{L}}, \dot{\mathscr{R}}$ which are given by

(1.3) $a\overset{+}{\mathscr{L}}b$ if and only if $a+b = a$ and $b+a = b$

(1.4) $a\overset{+}{\mathscr{R}}b$ if and only if $a+b = b$ and $b+a = a$

(1.5) $a\dot{\mathscr{L}}b$ if and only if $ab = a$ and $ba = b$

(1.6) $a\dot{\mathscr{R}}b$ if and only if $ab = b$ and $ba = a.$

Obviously $\overset{+}{\mathscr{L}}, \overset{+}{\mathscr{R}}, \overset{+}{\mathscr{L}}\circ\overset{+}{\mathscr{R}}=\overset{+}{\mathscr{R}}\circ\overset{+}{\mathscr{L}}=\overset{+}{\mathscr{D}}$ and $\dot{\mathscr{L}}, \dot{\mathscr{R}}, \dot{\mathscr{L}}\circ\dot{\mathscr{R}}=\dot{\mathscr{R}}\circ\dot{\mathscr{L}}=\dot{\mathscr{D}}$ are the usual relations of Green for the bands $(S, +)$ and $(S, \cdot)$ respectively [7]. It is easy to see that $\dot{\mathscr{L}}$ and $\dot{\mathscr{R}}$ are congruence relations on the reduct $(S, +)$ [1]. Hence $\dot{\mathscr{D}}$ is a congruence relation on the reduct $(S, +)$. On the other hand $\dot{\mathscr{D}}$ is the least congruence on $(S, \cdot)$ for which $(S/\dot{\mathscr{D}}, \cdot)$ is a semilattice. Thus, $\dot{\mathscr{D}}$ is a congruence relation on the ID-semiring $(S, +, \cdot)$ and it is in fact the least congruence on the semiring $(S, +, \cdot)$ for which the quotient has a multiplicative reduct which is a semilattice. By interchanging the role of $+$ and $\cdot$ we can state a similar result for $\overset{+}{\mathscr{D}}$. Consequently $\dot{\mathscr{D}}\vee\overset{+}{\mathscr{D}}$ is the least congruence on the ID-semiring $(S, +, \cdot)$ for which the quotient is a distributive bisemilattice.

Theorem 1. *Every normal semiring* $(S, +, \cdot)$ *is a subdirect product of* (i) *a normal semiring* S_1 *which has a left normal additive reduct and a right normal multiplicative reduct*, (ii) *a left normal semiring* S_2, (iii) *a normal semiring* S_3 *which has a right normal additive reduct and a left normal multiplicative reduct and* (iv) *a right normal semiring* S_4, *where* S, S_1, S_2, S_3 *and* S_4 *have the same greatest bisemilattice homomorphic image.*

Proof. Since $(S, +)$ is a normal band, $\overset{+}{\mathscr{R}}$ and $\overset{+}{\mathscr{L}}$ are congruences on $(S, +)$ [6][19]. Therefore $\overset{+}{\mathscr{R}}$ and $\overset{+}{\mathscr{L}}$ are congruences on S, and since $\overset{+}{\mathscr{R}}\cap\overset{+}{\mathscr{L}}$ is the equality, we have that S is a subdirect product of the normal semirings $V=S/\overset{+}{\mathscr{R}}$ and $W=S/\overset{+}{\mathscr{L}}$. Since $\overset{+}{\mathscr{R}}\subseteq\overset{+}{\mathscr{D}}\vee\dot{\mathscr{D}}$ and $\overset{+}{\mathscr{L}}\subseteq\overset{+}{\mathscr{D}}\vee\dot{\mathscr{D}}$, it follows that S, V and W have the same greatest bisemilattice homomorphic image. Let $\overset{+}{\mathscr{L}}_V, \overset{+}{\mathscr{R}}_V, \overset{+}{\mathscr{D}}_V, \dot{\mathscr{L}}_V, \dot{\mathscr{R}}_V, \dot{\mathscr{D}}_V$ and $\overset{+}{\mathscr{L}}_W, \overset{+}{\mathscr{R}}_W, \overset{+}{\mathscr{D}}_W, \dot{\mathscr{L}}_W, \dot{\mathscr{R}}_W, \dot{\mathscr{D}}_W$ be the corresponding relations of Green for the normal semirings V and W respectively. Since $\overset{+}{\mathscr{L}}_V=\overset{+}{\mathscr{D}}_V$, and $\overset{+}{\mathscr{R}}_V$ is trivial, we know that the additive reduct of V is a left normal band. Dually, the additive reduct of W is a right normal band. Again, $\dot{\mathscr{L}}_V$ and $\dot{\mathscr{R}}_V$ are congruences on V. Putting $S_1=V/\dot{\mathscr{L}}_V$ and $S_2=V/\dot{\mathscr{R}}_V$, we have that V is a subdirect product of S_1 and S_2, where V, S_1 and S_2 have the same greatest bisemilattice homomorphic image, and where S_1 and S_2 are as stated in the theorem. On the other hand, W is a subdirect product of $S_3=W/\dot{\mathscr{R}}_W$ and $S_4=W/\dot{\mathscr{L}}_W$, where W, S_3 and S_4 have the same greatest bisemilattice homomorphic image, and where S_3 and S_4 are as stated in the theorem.

Theorem 2. *An ID-semiring $(S, +, \cdot)$ is left [right] normal if and only if S divides the direct product of* (i) *a left-zero [right-zero] semiring T_1,* (ii) *the greatest bisemilattice homomorphic image T_2 of S,* (iii) *a normal semiring T_3 for which the additive reduct is a left-zero [right-zero] band and the multiplicative reduct a semilattice, and* (iv) *a normal semiring T_4 for which the additive reduct is a semilattice and the multiplicative reduct a left-zero [right-zero] band.*

Proof. Let us suppose that S is a left normal semiring.

1. Let $(V, +, \cdot)$ be the semiring where $V=S$, such that the additive reduct $(V, +)$ coincides with the additive reduct $(S, +)$, and such that for all $a, b\in S=V$, $ab=a$ in $(V, \cdot)$. Clearly V is a normal semiring for which the additive reduct is a left normal band and the multiplicative reduct a left-zero band.

$\dot{\mathscr{L}}=\dot{\mathscr{D}}$ is the least congruence relation on S for which the multiplicative reduct of the quotient is a semilattice. Hence $W=S/\dot{\mathscr{L}}$ is a normal semiring for which the additive reduct is a left normal band and the multiplicative reduct a semilattice, and the greatest bisemilattice homomorphic image of W is exactly the same as the greatest bisemilattice homomorphic image of S. For every $x\in S$, let $\dot{L}_x$ denote the $\dot{\mathscr{L}}$-class containing x. From [6][19] it follows that for any $a, x\in S$, we have $\dot{L}_x\leqq\dot{L}_a$ in $(W, \cdot)$, that is $\dot{L}_x=\dot{L}_x\dot{L}_a=\dot{L}_a\dot{L}_x$, if and only if $x=xa$ in $(S, \cdot)$ or, if and only if $x\dot{\mathscr{L}}ax$. In this case ax is the only element of $\dot{L}_x$ which commutes with a.

Let us consider the semiring $V\times W$, and the subset R of $V\times W$ which is given by $R=\{(a, \dot{L}_x)|\dot{L}_x\leqq\dot{L}_a \text{ in } S/\dot{\mathscr{L}}\}$. Let $(a, \dot{L}_x)$ and $(b, \dot{L}_y)$ be any elements of R.

Since $\dot{L}_x\dot{L}_y \leqq \dot{L}_x \leqq \dot{L}_a$, we have $(a, \dot{L}_x)(b, \dot{L}_y)=(a, \dot{L}_x\dot{L}_y)=(a, \dot{L}_{xy})\in R$. From $xa=x$ and $yb=y$, we have $(x+y)(a+y)=xa+y=x+y$ and $(a+y)(a+b)=a+yb=a+y$, thus $\dot{L}_{x+y} \leqq \dot{L}_{a+y} \leqq \dot{L}_{a+b}$, and so $(a, \dot{L}_x)+(b, \dot{L}_y)=(a+b, \dot{L}_x+\dot{L}_y)=(a+b, \dot{L}_{x+y})\in R$. We conclude that R is a subsemiring of $V\times W$.

We now introduce the mapping $\varphi\colon R\to S$, $(a, \dot{L}_x)\to ax$. This mapping is well-defined, since ax is the unique element of $\dot{L}_x$ which commutes with a. Further, φ is surjective, since $(x, \dot{L}_x)\in R$ for all $x\in S$, and $(x, \dot{L}_x)\varphi=x$. Again, let $(a, \dot{L}_x)$ and $(b, \dot{L}_y)$ be any elements of R. Since $ax\in\dot{L}_x$ and $by\in\dot{L}_y$, we have $ax+by\in \in\dot{L}_x+\dot{L}_y=\dot{L}_{x+y}$. Further, $ax+by=a(ax)+by=(a+by)(ax+by)=(a+b(by))\cdot \cdot(ax+by)=(a+b)(a+by)(ax+by)$ from which it follows that $ax+by$ is the unique element of $\dot{L}_{x+y}$ which commutes with $a+b$. Thus $(a+b, \dot{L}_{x+y})\varphi=ax+by$, and so $(a, \dot{L}_x)\varphi+(b, \dot{L}_y)\varphi=ax+by=(a+b, \dot{L}_{x+y})\varphi=((a, \dot{L}_x)+(b, \dot{L}_y))\varphi$. Since $yb=y$, we have by the left normality of $(S, \cdot)$ that $(a, \dot{L}_x)\varphi(b, \dot{L}_y)\varphi=axby=axyb=axy= =(a, \dot{L}_{xy})\varphi=((a, \dot{L}_x)(b, \dot{L}_y))\varphi$. We conclude that φ is a homomorphism of R onto S. Hence, S divides the direct product of V and W, where the greatest bisemilattice homomorphic image of V is trivial, and where S and W have the same greatest bisemilattice homomorphic image.

2. By interchanging the role of $+$ and $\cdot$, and using the same method as in 1, we can now show that V divides the direct product of a left-zero semiring T_1 and a normal semiring T_4 for which the additive reduct is a semilattice and the multiplicative reduct a left-zero band. Here T_1 is the left-zero semiring with carrier $S=V$, and T_4 is the normal semiring for which the additive reduct is the semilattice $(V/\overset{+}{\mathscr{L}}_V, +)=(S/\overset{+}{\mathscr{L}}, +)$ and the multiplicative reduct the left-zero band on the set $V/\overset{+}{\mathscr{L}}_V=S/\overset{+}{\mathscr{L}}$. Similarly, W divides the direct product of a normal semiring T_3 for which the additive reduct is a left-zero band and the multiplicative reduct a semilattice, and a normal semiring T_2 which is actually a distributive bisemilattice. Here the additive reduct of T_3 is the left-zero band on the set $S/\dot{\mathscr{L}}$, whereas the multiplicative reduct of T_3 coincides with the semilattice $(S/\dot{\mathscr{L}}, \cdot)$; on the other hand $T_2=(W/\overset{+}{\mathscr{L}}_W)\cong \cong S/(\dot{\mathscr{L}}\vee\overset{+}{\mathscr{L}})$, where $S/(\dot{\mathscr{L}}\vee\overset{+}{\mathscr{L}})$ is the greatest bisemilattice homomorphic image of S. We conclude that S divides the direct product of the semirings T_1, T_2, T_3 and T_4.

3. The semirings T_1, T_2, T_3, T_4 mentioned in the theorem are all left normal semirings. Therefore every semiring which divides their direct product must also be left normal.

In a similar fashion we can prove the following theorem.

Theorem 3. *An ID-semiring* $(S, +, \cdot)$ *has a right* [*left*] *normal additive reduct and a left* [*right*] *normal multiplicative reduct if and only if* S *divides the direct product of* (i) *a nest* [*dual nest*] T_1, (ii) *the greatest bisemilattice homomorphic image* T_2 *of*

S, (iii) *a normal semiring* T_3 *for which the additive reduct is a right-zero [left-zero] band and the multiplicative reduct a semilattice, and* (iv) *a normal semiring* T_4 *for which the additive reduct is a semilattice and the multiplicative reduct a left-zero [right-zero] band.*

By Result 0.1, Corollary 0.2, Theorems 1.1, 1.2 and 1.3, we can now conclude to the following division theorem for normal semirings in general.

Theorem 4. *An ID-semiring* $(S, +, \cdot)$ *is normal if and only if S divides the direct product of* (i) *a rectangular semiring* T_1, (ii) *the greatest bisemilattice homomorphic image* T_2 *of S*, (iii) *a normal semiring* T_3 *for which the additive reduct is a rectangular band and the multiplicative reduct a semilattice, and* (iv) *a normal semiring* T_4 *for which the additive reduct is a semilattice and the multiplicative reduct a rectangular band.*

It should be remarked that the components T_1, T_2, T_3 and T_4 can be constructed in terms of sets, semilattices and distributive lattices (Result 0.1, Corollary 0.2, [9]).

Lemma 5. *The greatest bisemilattice homomorphic image of a normal semiring* $(S, +, \cdot)$ *is a distributive lattice if and only if S satisfies the identity*

$$x(x+y+x)x = x. \tag{1.7}$$

Proof. A distributive bisemilattice which satisfies (1.7) must be a distributive lattice [9]. Therefore the greatest bisemilattice homomorphic image of a normal semiring satisfying (1.7) must be a distributive lattice. Let us conversely suppose that the greatest bisemilattice homomorphic image of the normal semiring $(S, +, \cdot)$ is a distributive lattice. By the foregoing theorem we know that S divides the direct product of T_1, T_2, T_3 and T_4, where T_1, T_3 and T_4 are as stated in Theorem 4, and where T_2 is a distributive lattice. One easily checks that T_1, T_2, T_3 and T_4 satisfy (1.7). Thus S satisfies (1.7).

Theorem 6. *An ID-semiring* $(S, +, \cdot)$ *is a normal semiring if and only if S is the Płonka sum of a semilattice ordered system of normal semirings that satisfy the generalized absorption law* (1.7).

Proof. Let E be the set of all equations which hold for all normal semirings for which the greatest bisemilattice homomorphic image is a distributive lattice. Let $R(E)$ denote the set of equations which are consequences of E and which are regular. Let K_E and $K_{R(E)}$ denote the equational classes defined by E and $R(E)$, respectively. Since (1.7) belongs to E we have from Theorem 1 of [10] that $K_{R(E)}$ consists of those algebras which are the Płonka sum of a semilattice ordered system of algebras which belong to K_E. Evidently the algebras which belong to K_E are exactly the normal semirings satisfying (1.7), and since the equations (0.1), (0.2), (0.3), (0.4), (0.5), (1.1) and (1.2) are all regular, it is obvious that $K_{R(E)}$ consists of normal semirings. The normal semirings T_1, T_3 and T_4 which appear in Theorem 4 all belong to K_E since their greatest bisemilattice homomorphic image is trivial. It follows from [9] that distributive bisemilattices belong to $K_{R(E)}$. Thus the normal semirings T_1, T_2, T_3 and T_4 of Theorem 4 all belong to $K_{R(E)}$. Since $K_{R(E)}$ is an equational class it follows from Theorem 4 that every normal semiring belongs to $K_{R(E)}$.

Remark. If we restrict ourselves to mono-semirings, then Theorem 6 is equivalent with the result of [19] which states that every normal band is the Płonka sum of a semilattice ordered system of rectangular bands. Our Theorem 6 also generalizes Płonka's decomposition of distributive bisemilattices [9]; this latter generalization goes in another direction than Padmanabhan's generalization of Płonka's result: we keep distributivity but abandon commutativity, Padmanabhan keeps commutativity but abandons distributivity [5].

It will be manifest for the reader that we actually used [19] and [9] in our proofs towards Theorem 6. We could have given a direct proof by showing that for every normal semiring $(S, +, \cdot)$

$$f\colon S^2 \to S, \quad (x, y) \to x(x+y+x)x$$

is a partition function of $(S, +, \cdot)$ (in the sense of [8]). Our procedure via Result 0.1, Corollary 0.2, Theorems 1 to 4, has the advantage that it gives insight in part of the lattice of subvarieties of the variety of normal semirings.

In the next section we generalize Theorem 6 for arbitrary ID-semirings.

2. A decomposition of ID-semirings

Let $(S, +, \cdot)$ be any ID-semiring. Then $S/\overset{+}{\mathscr{D}}$ and $S/\dot{\mathscr{D}}$ are ID-semirings where one of the reducts is a semilattice. Our first result here states that $S/\overset{+}{\mathscr{D}}$ and $S/\dot{\mathscr{D}}$ must be normal semirings.

Theorem 1. *Let $(S, +, \cdot)$ be an ID-semiring where $(S, +)$ is a semilattice. Then $(S, \cdot)$ is a normal band.*

Proof. Clearly $\dot{\mathscr{D}}$ is the least congruence on $(S, +, \cdot)$ for which the quotient is a distributive bisemilattice. For any $a \in S$ we denote the $\dot{\mathscr{D}}$-class containing a by $\bar{a}$. The distributive bisemilattice $T = S/\dot{\mathscr{D}}$ is the Płonka sum of a semilattice ordered system of distributive lattices

$$\langle Y, \langle T_\alpha \rangle_{\alpha \in Y}, \langle \varphi_{\beta,\alpha} \rangle_{\alpha \leqq \beta;\, \alpha, \beta \in Y} \rangle$$

[9].

Let a and b any elements of S, such that $ab=ba=b$. Clearly $\bar{a}\bar{b}=\bar{b}\bar{a}=\bar{b}$ in the distributive bisemilattice T, and so $\bar{a}\in T_\alpha, \bar{b}\in T_\beta$ for some $\alpha, \beta\in Y$, with $\beta\leqq\alpha$. We distinguish two cases: 1. $\alpha=\beta$, and 2. $\alpha>\beta$.

1. If $\alpha=\beta$, then $\bar{a}$ and $\bar{b}$ belong to the same distributive lattice T_α. Let x be any element of $\bar{b}$ which satisfies $xa=ax=x$. Then $x+a=a+x\in\bar{a}+\bar{b}=\bar{a}$. One readily checks that also $bx+b\in\bar{b}$, and $(bx+b)a=a(bx+b)=bx+b$, and so $bx+b+a=a$. Consequently $b=ba=b(bx+b+a)=bx+b$ and dually $b=xb+b$. If we interchange the role of x and b we also have $x=xb+x=bx+x$. Further, $xb+bx=(xb+bx)^2=(xb)^2+(xb)(bx)+(bx)(xb)+(bx)^2=xb+x+b+bx=(b+x)^2= =b+x$ and so $b=b+xb+bx=xb+bx=xb+bx+x=x$. We conclude that b is the only element of its $\dot{\mathcal{D}}$-class which is a solution of $xa=ax=x$.

2. Let us now suppose that $\alpha>\beta$. From $ab=b$ it follows that $(\bar{a}\varphi_{\alpha,\beta})\bar{b}=\bar{a}\bar{b}=\bar{b}$, where $\bar{a}\varphi_{\alpha,\beta}, \bar{b}\in T_\beta$. Since T_β is a distributive lattice, we also have $\bar{a}+\bar{b}=\bar{a}\varphi_{\alpha,\beta}+\bar{b}= =\bar{a}\varphi_{\alpha,\beta}$, and consequently $a+b\in\bar{a}\varphi_{\alpha,\beta}$. One can see that $(a+b)a=a(a+b)=a+b$. Let y be any element of $\bar{a}\varphi_{\alpha,\beta}$ which satisfies $ya=ay=y$. Clearly $y+a=a+y\in\bar{a}\varphi_{\alpha,\beta}$, and since $\bar{a}\varphi_{\alpha,\beta}$ forms a multiplicative rectangular band, we have $a+y=y+a= =(a+y)y(a+y)=(a+y)(ya+y^2)=(a+y)y=ay+y^2=y$. Similarly $y(a+b)$, $(a+b)y\in \in\bar{a}\varphi_{\alpha,\beta}$ and $a+y(a+b)=y(a+b)$, $a+(a+b)y=(a+b)y$.

Therefore,

$$a+b=(a+b)y(a+b)=(a+b)y\big((a+b)+a\big)=(a+b)y(a+b)+(a+b)y=$$
$$=(a+b)+(a+b)y=(a+b)(a+y)=(a+b)y=\big(a+(a+b)\big)y=$$
$$=y+(a+b)y=y(a+b)y+(a+b)y=(y+a)(a+b)y=y(a+b)y=y,$$

and we conclude that $a+b$ is the only element of the $\dot{\mathcal{D}}$-class $\bar{a}\varphi_{\alpha,\beta}$ which is a solution of $ya=ay=y$. Let x be any element of $\bar{b}$ which satisfies $xa=ax=x$. Again $a+x\in\bar{a}\varphi_{\alpha,\beta}$ and $a(a+x)=(a+x)a=a+x$, and so by the foregoing we have $a+x= =a+b$. Furthermore $b=(a+b)b=b(a+b)$, $x=(a+x)x=(a+b)x=x(a+x)= =x(a+b)$, where $x, b\in\bar{b}$, and where $\bar{b}$ and $\overline{a+b}$ belong to the same distributive lattice T_β. By 1. we conclude that $b=x$. Therefore, b is the only element of its $\dot{\mathcal{D}}$-class which is a solution of $ax=xa=x$.

From 1. and 2. we conclude that $\bar{a}\bar{b}=\bar{b}$ in $S/\dot{\mathcal{D}}$ implies that with every $a\in\bar{a}$ there corresponds a unique $b\in\bar{b}$ which is a solution of $ax=xa=x$. Therefore $(S, \cdot)$ is a normal band [6].

Theorem 2. *For an ID-semiring $(S, +, \cdot)$ the following are equivalent:*

(i) $\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$ *is the equality,*

(ii) $(S, +, \cdot)$ *satisfies*

(2.1) $$xy+yx=yx+xy,$$

(iii) $(S, +, \cdot)$ *is a subdirect product of ID-semirings for which one of the reducts is a semilattice,*

(iv) $(S, +, \cdot)$ *divides the direct product of* (1) *a distributive bisemilattice,* (2) *a normal semiring for which the additive reduct is a rectangular band and the multiplicative reduct a semilattice, and* (3) *a normal semiring for which the additive reduct is a semilattice and the multiplicative reduct a rectangular band.*

Proof. Let $(S, +, \cdot)$ be an ID-semiring which satisfies (i). Then $(S, +, \cdot)$ is a subdirect product of $S/\dot{\mathscr{D}}$ and $S/\overset{+}{\mathscr{D}}$, where the multiplicative reduct of $S/\dot{\mathscr{D}}$ is a semilattice, and the additive reduct of $S/\overset{+}{\mathscr{D}}$ is a semilattice. Thus (i) implies (iii). It follows from Theorems 1.1 to 1.4 that (iii) implies (iv). Let us now suppose that $(S, +, \cdot)$ satisfies (iv). Clearly the normal semirings listed in (iv) all satisfy (2.1), and so $(S, +, \cdot)$ satisfies (2.1): consequently (iv) implies (ii). Let us suppose that $(S, +, \cdot)$ satisfies (2.1). We know that the $\dot{\mathscr{D}}\cap\overset{+}{\mathscr{D}}$-classes are rectangular. On the other hand, the $\dot{\mathscr{D}}\cap\overset{+}{\mathscr{D}}$-classes form subalgebras which satisfy (2.1). It is an easy matter to check that a left-zero semiring, a right-zero semiring, a nest or a dual nest can only satisfy (2.1) if they are trivial. From Corollary 0.2 we conclude that the $\dot{\mathscr{D}}\cap\overset{+}{\mathscr{D}}$-classes of $(S, +, \cdot)$ must be trivial. Thus (ii) implies (i).

Let K be a fixed class of algebras of finite type and K_1, K_2 subclasses of K. The *product* $K_1 \circ K_2$ is the class of all algebras A from K on each of which one can find a congruence θ such that $A/\theta \in K_2$ and such that all θ-classes form subalgebras which belong to K_1 [3].

Theorem 3. *The variety of ID-semirings is the product of the variety of rectangular semirings and the variety of ID-semirings which satisfy the equivalent conditions of Theorem* 2.

Proof. The canonical homomorphism $\theta\colon S \to S/\dot{\mathscr{D}} \times S/\overset{+}{\mathscr{D}}$ induces the congruence relation $\dot{\mathscr{D}}\cap\overset{+}{\mathscr{D}}$ on S. Clearly $S/\dot{\mathscr{D}}\cap\overset{+}{\mathscr{D}} \cong S\theta$ satisfies the equivalent conditions (i) to (iv) of Theorem 2, and the $(\dot{\mathscr{D}}\cap\overset{+}{\mathscr{D}})$-classes form rectangular semirings.

Remark that we can construct rectangular semirings and the semirings which appear in (iv) of Theorem 2 in terms of sets, semilattices and distributive lattices by our results of Section 0. We now give a decomposition for arbitrary ID-semirings; we use Theorem 1.6, and we shall generalize it.

Lemma 4. *The greatest bisemilattice homomorphic image of an ID-semiring* $(S, +, \cdot)$ *is a distributive lattice if and only if S satisfies the generalized absorption law* (1.7).

Proof. It should be clear that the greatest bisemilattice homomorphic image of an ID-semiring which satisfies (1.7) must be a distributive lattice. Let us now suppose that the greatest bisemilattice homomorphic image of the ID-semiring $(S, +, \cdot)$ is a distributive lattice $S/\dot{\mathcal{D}}\vee\overset{+}{\mathcal{D}}$. Since $\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}\subseteq\dot{\mathcal{D}}\vee\overset{+}{\mathcal{D}}$, we have that $\dot{\mathcal{D}}\vee\overset{+}{\mathcal{D}}/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$ is the least congruence on $S/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$ for which the quotient is a bisemilattice, and $(S/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}})/(\dot{\mathcal{D}}\vee\overset{+}{\mathcal{D}}/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}})\cong S/\dot{\mathcal{D}}\vee\overset{+}{\mathcal{D}}$ is a distributive lattice. By Theorem 1 and Theorem 3 we know that $S/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$ is a normal semiring, and so by Lemma 1.5 we may conclude that the normal semiring $S/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$ satisfies (1.7). Let a and b be any elements of S. We denote the $\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$-class containing a and b by $\bar{a}$ and $\bar{b}$ respectively. Then $a(a+b+a)a\in\bar{a}(\bar{a}+\bar{b}+\bar{a})\bar{a}=\bar{a}$ since $S/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$ satisfies (1.7). Then a, $a(a+b+a)a\in\bar{a}$, and since the multiplicative reduct of $\bar{a}$ is rectangular we have $a(a+b+a)a=a$. Thus $(S, +, \cdot)$ satisfies (1.7).

Theorem 5. *Every ID-semiring is the Płonka sum of a semilattice ordered system of ID-semirings which satisfy the generalized absorption law* (1.7).

Proof. Let $(S, +, \cdot)$ be any ID-semiring, and let us consider the normal semiring $S/\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}=\bar{S}$. For any $a\in S$, $\bar{a}$ will denote the $\dot{\mathcal{D}}\cap\overset{+}{\mathcal{D}}$-class containing a. It follows from Theorem 1.6 that $\bar{S}$ is the Płonka sum of a semilattice ordered system of normal semirings satisfying (1.7)

$$\langle Y, \langle \bar{S}_\alpha\rangle_{\alpha\in Y}, \langle \varphi_{\alpha,\beta}\rangle_{\beta\leqq\alpha;\,\alpha,\beta\in Y}\rangle. \tag{2.2}$$

For any $\alpha\in Y$, let $S_\alpha=\{a|\bar{a}\in\bar{S}_\alpha\}$.

Let $a\in S_\alpha$, and let $\beta\leqq\alpha$ in Y. Let x be any element of $\bar{a}\varphi_{\alpha,\beta}$, and consider $b=axa$. Then $axa\in\bar{a}\bar{x}\bar{a}=\bar{a}\varphi_{\alpha,\beta}$, and so b is an element of $\bar{a}\varphi_{\alpha,\beta}$ which satisfies $ab=ba=b$. Then $a+b\in\bar{a}+\bar{b}=\bar{a}\varphi_{\alpha,\beta}$, and so $a+b=(a+b)b(a+b)=(a+b)(ba+b^2)=$ $=(a+b)b=ab+b^2=b$, and similarly $b+a=b$. By symmetry we may conclude that an element b of $\bar{a}\varphi_{\alpha,\beta}$ is a solution of $ay=ya=y$ if and only if b is a solution of $a+y=y+a=y$. Let b and b' be any elements of $\bar{a}\varphi_{\alpha,\beta}$ such that $ab=ba=b$ and $ab'=b'a=b'$. Since $\bar{a}\varphi_{\alpha,\beta}$ forms a rectangular semiring, we have from the foregoing

$$b = (b+b')+b = (b+b')+aba = (b+b'+a)(b+b'+b)(b+b'+a) =$$
$$= (b+b')b(b+b') = b+b' = (b+b')b'(b+b') =$$
$$= (a+b+b')(b'+b+b')(a+b+b') = ab'a+(b+b') = b'+b+b' = b'.$$

Hence $a+y=y+a=y$ has a unique solution in $\bar{a}\varphi_{\alpha,\beta}$, and this solution coincides with the unique solution of $ay=ya=y$ in $\bar{a}\varphi_{\alpha,\beta}$. We define the mapping $\psi_{\alpha,\beta}$: $S_\alpha\to S_\beta$ by the requirement that for every $a\in S_\alpha$, $a\psi_{\alpha,\beta}$ is the unique solution in

$\bar{a}\varphi_{\alpha,\beta}$ of the above considered equations. It follows from our considerations that $\psi_{\alpha,\beta}$ is well-defined.

Let us consider the system

$$(2.3) \qquad \langle Y, \langle S_\alpha\rangle_{\alpha\in Y}, \langle\psi_{\beta,\alpha}\rangle_{\alpha\leqq\beta;\alpha,\beta\in Y}\rangle.$$

It should be obvious that for every $\alpha\in Y$, S_α and $\bar{S}_\alpha$ have the same greatest bisemilattice homomorphic image which is a distributive lattice. Therefore S_α is an ID-semiring which satisfies (1.7) for all $\alpha\in Y$. Evidently $S_\alpha\cap S_\beta=\varnothing$ if $\alpha\neq\beta$ since $\bar{S}_\alpha\cap\bar{S}_\beta=\varnothing$ in that case.

Let $\alpha, \beta\in Y$, with $\beta\leqq\alpha$, and let us suppose that $a, a'\in S_\alpha$. Then $b=a\psi_{\alpha,\beta}$, $b'=a'\psi_{\alpha,\beta}\in S_\beta$, and $bb'\in\bar{b}\bar{b}'=\bar{a}\varphi_{\alpha,\beta}\bar{a}'\varphi_{\alpha,\beta}=\bar{a}\bar{a}'\varphi_{\alpha,\beta}$. Further,

$$bb' = (b+a)(b'+a') = bb'+ba'+ab'+aa'$$

implies that $bb'+aa'=bb'$, and similarly $aa'+bb'=bb'$. Hence $bb'=(aa')\psi_{\alpha,\beta}$, and so $(aa')\psi_{\alpha,\beta}=(a\psi_{\alpha,\beta})(a'\psi_{\alpha,\beta})$. Analogously $(a+a')\psi_{\alpha,\beta}=a\psi_{\alpha,\beta}+a'\psi_{\alpha,\beta}$. We conclude that $\langle\psi_{\alpha,\beta}\rangle_{\beta\leqq\alpha;\alpha,\beta\in Y}$ is a family of homomorphisms. It should be obvious that for every $\alpha\in Y$, $\psi_{\alpha,\alpha}$ is the identity mapping on S_α. Further, let $\gamma\leqq\beta\leqq\alpha$ in Y, and let $a\in S_\alpha$. Let us put $a\psi_{\alpha,\beta}=b$ and $b\psi_{\beta,\gamma}=c$. Then $b\in\bar{a}\varphi_{\alpha,\beta}$, and $c\in\bar{b}\varphi_{\beta,\gamma}=\bar{a}\varphi_{\alpha,\beta}\varphi_{\beta,\gamma}=\bar{a}\varphi_{\alpha,\gamma}$. Further $ab=ba=b$ and $bc=cb=c$ imply $ac=ca=c$, and so $c=a\psi_{\alpha,\gamma}$. Thus $\psi_{\alpha,\beta}\psi_{\beta,\gamma}=\psi_{\alpha,\gamma}$. We conclude that (2.3) is a semilattice ordered system of ID-semirings which satisfy the generalized absorption law (1.7).

Finally, let $a\in S_\alpha$ and $b\in S_\beta$, and put $a\psi_{\alpha,\alpha\beta}=a'$ and $b\psi_{\beta,\alpha\beta}=b'$. Then $a'b'\in\bar{a}'\bar{b}'=\bar{a}\varphi_{\alpha,\alpha\beta}\bar{b}\varphi_{\beta,\alpha\beta}=\bar{a}\bar{b}=\overline{ab}$, hence $a'b'$ and ab belong to the same rectangular semiring $\overline{a'b'}=\overline{ab}$. Further, $a'b'=(a'+a)(b'+b)=a'b'+a'b+ab'+ab$ and so $a'b'+ab=a'b'$. Similary $ab+a'b'=a'b'$, and thus $ab=ab+a'b'+ab=a'b'$ We conclude that $ab=(a\psi_{\alpha,\alpha\beta})(b\psi_{\beta,\alpha\beta})$ and by symmetry, $a+b=(a\psi_{\alpha,\alpha\beta})+(b\psi_{\beta,\alpha\beta})$. Consequently, $(S, +, \cdot)$ is the Płonka sum of the system (2.3).

The reader may check that for any $\alpha, \beta\in Y$, with $\beta\leqq\alpha$, and $a\in S_\alpha$, we have $a\psi_{\alpha,\beta}=b$ if and only if $a(a+b+a)a=b$. Thus we can conclude to the following.

Corollary 6. *For any ID-semiring $(S, +, \cdot)$,*

$$f\colon S^2\to S, \quad (x,y)\to x(x+y+x)x$$

is a partition function.

Proof. Immediate from the proof of the foregoing and from [8].

Problem. Let $(S, +, \cdot)$ be an ID-semiring where $(S, +)$ is a semilattice. By Theorem 2.1 $(S, \cdot)$ must be a normal band. Characterize the normal bands which can be realized in this way. Remark that [16] and [17] give information about the normal bands $(S, \cdot)$ where $(S, +, \cdot)$ is an ID-semiring for which $(S, +)$ is a chain.

Acknowledgement. This paper was started when the first author visited the Politechnika Warszawska in April 1980, sponsored by the Polish Mathematical Society, and finished when the second author visited the Technische Hochschule Darmstadt in 1980, as a research fellow of the A. v. Humboldt Foundation.

References

[1] M. P. GRILLET, Semirings with a completely simple additive semigroup, *J. Austral. Math. Soc.*, **20** (1975), 257—267.

[2] P. JORDAN, Halbgruppen von Idempotenten und nichtkommutative Verbände, *J. Reine Angew. Math.*, **211** (1962), 136—161.

[3] A. I. MAL'CEV, *The Metamathematics of Algebraic Systems*, North-Holland (Amsterdam, 1971).

[4] R. MCKENZIE, A. ROMANOWSKA, Varieties of ·-distributive bisemilattices, *Contributions of General Algebra* (Proceedings of the Klagenfurt Conference 1978), 213—218.

[5] R. PADMANABHAN, Regular identities in lattices, *Trans. Amer. Math. Soc.*, **158** (1971), 179—188.

[6] M. PETRICH, A construction and a classification of bands, *Math. Nachr.*, **48** (1971), 263—274.

[7] M. PETRICH, *Lectures in Semigroups* (London, 1977).

[8] J. PŁONKA, On a method of construction of abstract algebras, *Fund. Math.*, **61** (1967), 183—189.

[9] J. PŁONKA, On distributive quasilattices, *Fund. Math.*, **60** (1967), 191—200.

[10] J. PŁONKA, On equational classes of abstract algebras defined by regular equations, *Fund. Math.*, **24** (1969), 241—247.

[11] J. PŁONKA, A note on the join and subdirect product of equational classes, *Algebra Universalis*, **1** (1974), 163—164.

[12] A. ROMANOWSKA, On bisemilattices with one distributive law, *Algebra Universalis*, **10** (1980), 36—47.

[13] A. ROMANOWSKA, Subdirectly irreducible ·-distributive bisemilattices I, *Demonstratio Math.*, **13** (1980), 767—785.

[14] A. ROMANOWSKA, Free ·-distributive bisemilattices, *Demonstratio Math.*, **13** (1980), 565—572.

[15] A. ROMANOWSKA, J. D. H. SMITH, Bisemilattices of subsemilattices, *J. Algebra*, **70** (1981), 78—88.

[16] T. SAITÔ, Ordered idempotent semigroups, *J. Math. Soc. Japan*, **4** (1962), 150—169.

[17] T. SAITÔ, The orderability of idempotent semigroups, *Semigroup Forum*, **7** (1974), 264—285.

[18] V. SLAVÍK, On skew lattices II, *Comment. Math. Univ. Carolinae*, **14** (1973), 493—506.

[19] M. YAMADA, N. KIMURA, Note on idempotent semigroups II, *Proc. Japan Acad.*, **34** (1958), 110—112.

[20] J. ZELEZNIKOW, On regular semigroups, semirings and rings, *Dissertation* (Monash University, 1979).

(F. P.)
DIENST HOGERE MEETKUNDE,
RIJKSUNIVERSITEIT TE GENT,
KRIJGSLAAN 271,
9000 GENT, BELGIUM

(A. R.)
INSTYTUT MATEMATYKI,
POLITECHNIKA WARSZAWSKA
PLAC JEDNOŚCI ROBOTNICZEJ 1
00661 WARSZAWA, POLAND

Acta Sci. Math. 44 (1982), 255—270

Definable principal congruence relations: Kith and kin

JOHN T. BALDWIN and JOEL BERMAN

This paper has two aims. Firstly, it seeks to illumine the way in which principal congruence relations are constructed. To this end, a hierarchy of "definability" of congruences is presented. Notions both weaker and stronger than (first order) *definable principal congruences* (dpc) are considered. Secondly, it attacks the problem posed by BURRIS and LAWRENCE [13], "If **K** is a class of algebras and if the quasivariety generated by **K**, $Q(\mathbf{K})$, has definable principal congruences must the variety generated by **K**, $V(\mathbf{K})$, also have dpc"? These two aims are linked by several results of the following form, "If $V(\mathbf{K})$ has (some weak notion of) definable principal congruences and $Q(\mathbf{K})$ has definable principal congruences, then $V(\mathbf{K})$ has definable principal congruences."

The discussion of the hierarchy mentioned above includes a survey of the literature on such notions and it attempts to connect these properties with others of a quite different nature. For example, several levels of the hierarchy are linked with n-permutability of (principal) congruence relations.

We acknowledge a number of very helpful communications with Peter Köhler in connection with this paper.

1. Definitions and notation

In general, we denote algebras by A, B, and C while **K** denotes a class of algebras of some fixed similarity type. The variety generated by **K** is $V(\mathbf{K})$; and $Q(\mathbf{K})$ denotes the quasivariety generated by the class **K**. A class of algebras is said to be *locally finite* if every finitely generated algebra in the class is finite; the class is *uniformly locally finite* if there exists a function f such that for all n every n-generated algebra in the class has cardinality at most $f(n)$.

If A is an algebra, and if x and y are elements of A, then the principal congruence relation generated by x and y is the smallest congruence relation on A for which x

Received May 29, 1980.

and y are congruent. It is denoted by $\Theta(x, y)$. For a given algebra A of some similarity type τ, Malcev has provided a description of the principal congruence generated by the elements a_0 and a_1 solely in terms of the polynomials of the algebra A, e.g. [23] or [16, p. 54]. Namely,

$$b_0 \equiv b_1 \quad \Theta(a_0, a_1) \leftrightarrow \exists n \exists p_1, \ldots, p_n \exists s \exists \mathbf{z}_1, \ldots, \mathbf{z}_n \quad \text{such that}$$

$$\text{(M)} \quad b_0 = p_1(a_{s(1)}, \mathbf{z}_1), \quad p_i(a_{1-s(i)}, \mathbf{z}_i) = p_{i+1}(a_{s(i+1)}, \mathbf{z}_{i+1}) \quad \text{for} \quad 1 \leqq i < n, \quad \text{and}$$

$$b_1 = p_n(a_{1-s(n)}, \mathbf{z}_n)$$

where n is a positive integer, the p_i are k_i-ary polynomials of type τ, s is a switching function, i.e. $s\colon \{1, 2, \ldots, n\} \to \{0, 1\}$, and $\mathbf{z}_i$ are $k_i - 1$ tuples from A, for $1 \leqq i \leqq n$. Each fixed instance of the polynomials p_i and the switching function s is called a *Malcev formula*.

We write $\psi(a_0, a_1, b_0, b_1, p_1, \ldots, p_n, s, \mathbf{z}_1, \ldots, \mathbf{z}_n)$ for the last 2 lines of (M). Note that the integer n is implicit in this formula. Also observe that $\exists \mathbf{z}_1, \ldots, \mathbf{z}_n\, \psi$ is a positive existential formula. Such a formula was called a congruence formula in [3].

There is an easy abstract form of this result which is implicit in [25]. (Also see [26] for a nice application.) Let $\text{Diag}^+(A) = \{R(\mathbf{z}) \mid R \text{ is atomic, } A \models R(\mathbf{z})\}$. Note that $\text{Diag}^+(A)$ is in a language $L(A)$ which has names for all members of A.

Lemma 1.1. (Folklore) *If a, b, c, d are in A and $c \equiv d\ \Theta(a, b)$, then there is a positive existential L-formula $S(x, y, u, v)$ such that*

i) $\models S(x, y, u, u)$ *implies* $x = y$,

ii) $A \models S(c, d, a, b)$.

Proof. Note $\text{Diag}^+(A) \cup (a = b) \cup (c \neq d)$ is consistent if and only if there is a homomorphic image of A which identifies a with b but does not identify c with d. Thus if $c \equiv d\ \Theta(a, b)$ then $\models \bigwedge \{R_i(c, d, a, b, \mathbf{e}) \mid 1 \leqq i \leqq m\} \,\&\, (a = b)$ implies $(c = d)$ for some finite set $R_1, \ldots, R_m$ of atomic formulas. But then

$$\models \forall x, y, u, v\big(\exists \mathbf{w}\big(\bigwedge \{R_i(x, y, u, v, \mathbf{w}) \mid 1 \leqq i \leqq m\}\big) \,\&\, (u = v)\big)$$

implies $x = y$ and thus $A \models \exists \mathbf{w} \bigwedge \{R_i(c, d, a, b) \mid 1 \leqq i \leqq m\}$.

Either of these characterizations allow us to "define" principal congruences in the language $L(\omega_1, \omega)$ i.e. the language which extends first order logic by allowing infinite disjunctions. Namely $c \equiv d\ \Theta(a, b) \leftrightarrow \bigvee \{S(c, d, a, b) \mid S \in P\}$ where P is the collection of positive existential formulas satisfying $S(x, y, u, u) \to x = y$ (or the collection of Malcev formulas). Although some information can be obtained from this weak definability of principal congruences (cf. [2]), in this paper we want to discuss various stronger notions of definability. The first formalization of this kind occurred as follows in [3]. A class **K** is said to have *definable principal conguences* if there is a 4-ary first order formula φ in the language of **K** such that

$$\forall A\in\mathbf{K}; \forall a_0, a_1, b_0, b_1\in A, \ (b_0\equiv b_1\, \Theta(a_0, a_1) \leftrightarrow \varphi(a_0, a_1, b_0, b_1)).$$

For further details on definable principal congruences see [3] and sections 2 and 5 below.

Our earlier work on definable principal congruences focused on varieties. However, [12] shows the advantages of dealing with more general classes. Thus in the following we define these notions for arbitrary classes of algebras. Occasionally it will be necessary to assume that such a class satisfies the compactness theorem. Of course any elementary class satisfies this condition. For this reason many of our results have a dual character, we first describe the effect of a property on a specified algebra and then note the effect on algebras in a variety satisfying this condition. We call the former a local result, and the latter a global one.

Since a disjunction of all possible ψ describes the principal congruences in any class **K** of algebras, it follows from the compactness theorem that if a class **K** which satisfies the compactness theorem has definable principal congruences, then this defining formula is equivalent to some finite disjunction of the ψ, i.e.

$$\varphi(a_0, a_1, b_0, b_1) \leftrightarrow \bigvee_i \psi(a_0, a_1, b_0, b_1, p_1^i, \ldots, p_n^i, s_i, \mathbf{z}_1^i, \ldots, \mathbf{z}_n^i),$$

where i ranges over some finite index set. Note there is a uniform subscript n in this formula. This is possible since the diagonal elements are in any principal congruence relation, i.e. $w\equiv w \ \Theta(x, y)$ allowing the "padding out" of formulas ψ having different lengths to one uniform length.

2. DPC and its relatives

In this section we discuss first some weakenings and then some strengthenings of the notion of a definable principal congruence relation. These arise in a natural way via a reshuffling of the quantifiers. That is, we require that certain of the existentially quantified variables in (M) do not depend on the particular a_0, a_1, b_0, and b_1.

We start with Malcev's characterization of principal congruences. For any algebra A in the class **K**, and for all a_0, a_1, b_0, b_1 in A,

$$b_0 \equiv b_1\ \Theta(a_0, a_1) \leftrightarrow \exists n \exists p_1, \ldots, p_n \exists s \exists \mathbf{z}_1, \ldots, \mathbf{z}_n$$
$$\psi(a_0, a_1, b_0, b_1, p_1, \ldots, p_n, s, \mathbf{z}_1, \ldots, \mathbf{z}_n).$$

We proceed to pull out existential quantifiers from the right side of this expression.

A) There is a bound n on the number of steps in determining principal congruences for all algebras in $\mathbf{K}$. In this case we say $\mathbf{K}$ has *n-step principal congruences*. By previous remarks we can, without loss of generality, assume all principal congruences can be described using a Malcev formula with the same fixed n. Formally, this gives

$$\exists n \quad \text{such that} \quad \forall A \in \mathbf{K} \quad \forall a_0, a_1, b_0, b_1 \in A$$
$$(b_0 \equiv b_1 \, \Theta(a_0, a_1) \leftrightarrow \exists p_1, \ldots, p_n \exists s \exists \mathbf{z}_1, \ldots, \mathbf{z}_n$$
$$\psi(a_0, a_1, b_0, b_1, p_1, \ldots, p_n, s, \mathbf{z}_1, \ldots, \mathbf{z}_n)).$$

Note that this is not as strong as definable principal congruences since the polynomials which are to be used are not specified (and the language does not allow variables having polynomials for values). But, by restricting n in this way there are only finitely many choices for the switching functions, i.e. the 2^n functions from $\{1, \ldots, n\}$ to $\{0, 1\}$. Algebras with n-step principal congruences are discussed in section 3.

B) The class $\mathbf{K}$ has n-step principal congruences and there is a *specified list of switching functions* for determining all principal congruences of algebras in $\mathbf{K}$. This becomes in the notational pattern we have adopted:

$$\exists n, \exists s_1, \ldots, s_k \quad \text{such that} \quad \forall A \in \mathbf{K} \quad \forall a_0, a_1, b_0, b_1 \in A$$
$$(b_0 \equiv b_1 \, \Theta(a_0, a_1) \leftrightarrow \exists p_1, \ldots, p_n \exists i \exists \mathbf{z}_1, \ldots, \mathbf{z}_n$$
$$\psi(a_0, a_1, b_0, b_1, p_1, \ldots, p_n, s_i, \mathbf{z}_1, \ldots, \mathbf{z}_n)).$$

We will be mainly interested in the case that $k=1$. Results concerning this situation will be given in section 4.

C) The class $\mathbf{K}$ has n-step principal congruences with a specified list of switching functions, and there is a *specified list of polynomials* to be used for determining principal congruences of all the algebras in $\mathbf{K}$. This is of course, *definable principal congruences*. Thus it formally becomes:

$$\exists n, \exists s_1, \ldots, s_k \exists p_1^i, \ldots, p_n^i \quad (1 \leqq i \leqq k) \quad \text{such that} \quad \forall A \in \mathbf{K} \quad \forall a_0, a_1, b_0, b_1 \in A$$
$$(b_0 \equiv b_1 \quad \Theta(a_0, a_1) \leftrightarrow \exists i \quad \exists \mathbf{z}_1, \ldots, \mathbf{z}_n \quad \psi(a_0, a_1, b_0, b_1, p_1^i, \ldots, p_n^i, s_i, \mathbf{z}_1, \ldots, \mathbf{z}_n)).$$

A discussion of definable principal congruences in this context is found in section 5.

D) The class $\mathbf{K}$ has n-step principal congruences with a *single switching function* s and a specified list of polynomials $p_1, \ldots, p_n$ to be used for determining principal congruences in **K**. Formally this becomes:

$$\exists n, \exists s, \exists p_1, \ldots, p_n \quad \text{such that} \quad \forall A \in \mathbf{K} \quad \forall a_0, a_1, b_0, b_1 \in A$$
$$(b_0 \equiv b_1 \, \Theta(a_0, a_1) \leftrightarrow \exists z_1, \ldots, z_n \quad \psi(a_0, a_1, b_0, b_1, p_1, \ldots, p_n, s, z_1, \ldots, z_n)).$$

If φ defines principal congruences for a class **K** then φ is a finite disjunction of Mal-cev formulas. In D) there is only one disjunct (this is exactly the distinction implicit in [3] between the congruences being defined by a weak congruence formula or a congruence formula). This notion is now called "**K** has a uniform congruence scheme" and has been investigated in [15], [7], [8], [21]. In particular, [15] shows uniform congruence scheme is equivalent to their notion of equationally definable principal congruences.

We thus have a hierarchy of properties A), B), C) and D). It is a natural question to investigate how this hierarchy behaves in the presence of other properties the class **K** may possess. These questions will not figure in the remainder of this paper; so we briefly mention them at this point. FRIED and KISS [33] have also considered related questions.

One possibility is to restrict the variables z_i to take on values from $\{a_0, a_1, b_0, b_1\}$. By a result of DAY [14] this is equivalent to the congruence extension property. (For details on the congruence extension property see [17] or [16].) If this restriction on the z_i is added to condition A) and if **K** also is uniformly locally finite (or at least 4-generated algebras in **K** have bounded cardinality), then **K** has definable principal congruences, and the defining formula can be made quantifier free as discussed in [3]. Moreover, if in D) all $z_i \in \{a_0, a_1, b_0, b_1\}$, then this gives the restricted uniform congruence scheme discussed in [15], [9], [10], [21] and [22].

Another possible restriction is to place a bound m on the possible arities of the polynomials which appear. In this case, if **K** is also uniformly locally finite, then conditions A) and B) both collapse to C). This is similar to the property CEP_n discussed in [21]. This also has some bearing on the notion of P_0-principal congruence relations which are defined in the next section. BAKER [31] considers restrictions on the polynomials for varieties of lattices.

If **K** is a variety with distributive congruence lattices, then it is shown in [15] that conditions C) and D) are the same. We do not know how this affects conditions A) and B).

3. Bounded number of steps

In this section we investigate condition A) for a class **K** of algebras. We provide some curious examples, pose some questions, and deal with some work of Burris and Lawrence on definable principal congruences.

The simplest example of a class of algebras having a finite bound on the number of steps required in principal congruence relations is provided by sets, i.e. algebras with no fundamental operations. For sets, the only polynomials are the projection operations. If A is a set, then for $x, y \in A$, the principal congruence relation $\Theta(x, y)$ consists of the ordered pairs (x, y), (y, x), and (w, w) for all $w \in A$. It is easily seen that these three cases can be handled with polynomial projection functions; but both possible switching functions are required, and two different projection functions are needed. This simple example is instructive since it shows that condition A) is not preserved under extension of varieties, and hence is not a Malcev condition. (For details on Malcev conditions see [16], [27], and [3].) This is to be contrasted with some of the results in section 4.

An instance of the significance of n-step principal congruences may be found in the work of BURRIS and LAWRENCE [12], [13] on definable principal congruences in groups and rings. In the second of these two papers, they define the notion of P_0-projective principal congruences for a class **K** of algebras. Essentially this is condition A) with $n=1$ and with the added stipulation that the polynomials which are to be used are drawn from some class P_0 of polynomials. They posed the following problem:

Problem 1. (Burris and Lawrence) Let **K** be a class of algebras such that $Q(\mathbf{K})$ has definable principal congruences. Does $V(\mathbf{K})$ also have definable principal congruences?

They prove the following in [13]:

Theorem 3.1. (Burris & Lawrence) *Let* **K** *be a class of algebras such that* $Q(\mathbf{K})$ *has definable principal congruences and such that* $V(\mathbf{K})$ *has* P_0*-projective principal congruence relations for some set* P_0. *Then* $V(\mathbf{K})$ *also has definable principal congruences.*

For an arbitrary class **K** of algebras, the class $Q(\mathbf{K})$ and the class $ISP(\mathbf{K})$ of all algebras isomorphic to subalgebras of products of members of **K** need not be the same. For example [4] and [5] contain a discussion of this. However, for any class **K**, $HQ(\mathbf{K})=HSP(\mathbf{K})=V(\mathbf{K})$ and the following variant on Problem 1 is possible.

Problem 1a. Let **K** be a class of algebras such that $SP(\mathbf{K})$ has definable principal congruences. Does $V(\mathbf{K})$ also have definable principal congruences?

We now present a generalization of Theorem 3.1 which answers Problem 1 and 1a for certain classes **K** of algebras.

Theorem 3.2. *Let* **K** *be a class of algebras such that* **K** *has definable principal congruences and such that* $H(\mathbf{K})$ *has* n*-step principal congruences for some integer* n. *Then* $H(\mathbf{K})$ *also has definable principal congruences.*

Proof. Let φ be any positive 4-ary formula defining principal congruences for the class **K**, and suppose $H(\mathbf{K})$ has n-step principal congruences. We claim $H(\mathbf{K})$ has definable principal congruences given by

$$(*)\qquad b_0 \equiv b_1 \quad \Theta(a_0, a_1) \ \text{ iff } \ \exists w_0, \ldots, w_n$$

$$(w_0 = b_0,\ w_n = b_1, \ \text{ and } \ \varphi(a_0, a_1, w_{i-1}, w_i) \ \text{ for } \ 1 \leqq i \leqq n).$$

To verify this claim consider $B \in H(\mathbf{K})$, $A \in \mathbf{K}$, and a homomorphism h from A onto B. Let $b_0 \equiv b_1 \quad \Theta(a_0, a_1)$ in B. So B models the formula

$$\psi(a_0, a_1, b_0, b_1, p_1, \ldots, p_n, s, \mathbf{z}_1, \ldots, \mathbf{z}_n)$$

for some choice of p_j, s, and $\mathbf{z}_j$, where the p_j are k_j-ary polynomials. Arbitrarily choose $a_i' \in A$ and $k_j - 1$ tuples $\mathbf{z}_j'$ with $h(a_i') = a_i$ and $h(\mathbf{z}_j') = \mathbf{z}_j$. Note that $h(p_j(a_i', \mathbf{z}_j')) = p_j(a_i, \mathbf{z}_j)$. Also, in the algebra A it is the case that $p_j(a_0', \mathbf{z}_j') \equiv \equiv p_j(a_1', \mathbf{z}_j') \ \Theta(a_0', a_1')$ for $1 \leqq j \leqq n$. Hence $A \models \varphi(a_0', a_1', p_j(a_0', \mathbf{z}_j'), p_j(a_1', \mathbf{z}_j'))$ and since φ is positive and h is a homomorphism $B \models \varphi(a_0, a_1, p_j(a_0, \mathbf{z}_j), p_j(a_1, \mathbf{z}_j))$. Moreover, in B, $p_j(a_{1-s(j)}, \mathbf{z}_j) = p_{j+1}(a_{s(j+1)}, \mathbf{z}_{j+1})$. So choosing $w_0 = p_1(a_{s(j)}, \mathbf{z}_1)$, $w_n = p_n(a_{1-s(n)}, \mathbf{z}_n)$, and $w_i = p_{i+1}(a_{s(i+1)}, \mathbf{z}_{i+1})$ with $1 \leqq i < n$ shows $(*)$ holds in B.

Corollary 3.3. *Let* **K** *be a class of algebras such that the class* $Q(\mathbf{K})$ *has definable principal congruences and the class* $V(\mathbf{K})$ *has n-step principal congruences for some integer n. Then* $V(\mathbf{K})$ *also has definable principal congruences. Moreover, this result holds if Q is replaced by SP.*

Because of Theorem 3.2 and because the formula ψ is positive, it would be tempting to conjecture that n-step principal congruences is preserved under homomorphism. The following shows this is not the case.

Example 3.4. Let A have universe consisting of the positive integers and suppose for each positive integer i there is a unary operation g_i such that

$$g_i(1) = 2i+1, \quad g_i(2) = 2i+2, \quad \text{and} \quad g_i(k) = k \quad \text{for all} \quad k > 2.$$

Then A satisfies condition A) with $n=2$, but A has a homomorphic image B which does not satisfy condition A) for any n. In order to verify this, first observe that there are only four types of principal congruences on A (we list the nontrivial blocks):

$\Theta(1,2) = 1,2/3,4/5,6/\ldots,$
$\Theta(1,k) = /k, \text{ odds}/, \ (k > 2)$
$\Theta(2,k) = /k, \text{ evens}/, \ (k > 2)$
$\Theta(k,m) = /k, m/ \ (k, m > 2).$

It is easily seen that the first and last of these congruences can be done in one step, while for the others, two steps will suffice. Define a homomorphism h of A so that h has kernel consisting of $\bigvee_i \Theta(2i, 2i+1)$ as i ranges over all $i > 1$. Then in the algebra

$h(A)$, the principal congruence relation $\Theta(h(1), h(2))$ will require an arbitrarily large number of steps.

Problem 2. Let A be an algebra that satisfies condition A) for its principal congruences with $n=1$. Does every homomorphic image of A also have this property?

We now present another approach to Problem 1. If A is an algebra, and if h is a homomorphism of A and if Θ is some congruence relation on A, then the image of Θ under the homomorphism h, denoted $h(\Theta)$, will consist of $\{(h(x), y(h)) | (x, y) \in \Theta\}$. Note that in general, because of transitivity, $h(\Theta)$ need not be a congruence in $h(A)$. We say that the homomorphism h *preserves* the congruence Θ if $h(\Theta)$ is a congruence relation of $h(A)$. Consider the following property for an algebra A:

(∗∗) For any homomorphism h of A any principal congruence relation $\Theta(h(x), h(y))$ of the algebra $h(A)$ is the image of some principal congruence relation of A.

Note that if h preserves $\Theta(x, y)$, then $\Theta(h(x), h(y)) = h(\Theta(x, y))$. Thus if every homomorphism of A preserves every principal congruence relation of A, then A has (∗∗). In section 4 we investigate the condition that a given homomorphism preserve a given congruence relation. Our interest in (∗∗) stems from the following.

Theorem 3.5. *Let* **K** *be a class of algebras with definable principal congruences, and suppose each algebra in* **K** *has property* (∗∗). *Then* $H(\mathbf{K})$ *also has definable principal congruences.*

Proof. Let **K** have definable principal congruences given by some positive formula φ. Consider $A \in \mathbf{K}$ and some homomorphism h of A onto an algebra B. If $u \equiv v \ \Theta(x, y)$ in B, then by property (∗∗), there are u', v', x', y' in A such that $h(\Theta(x', y')) = \Theta(x, y)$ and $h(u') = u$, $h(v') = v$, and $u' \equiv v' \ \Theta(x', y')$. Hence, $A \models \varphi(x', y', u', v')$, and since φ is positive, it follows that $B \models \varphi(x, y, u, v)$. Thus φ serves to define principal congruences in B as well.

We can relate property (∗∗) to our hierarchy by the following observation. If **K** is a class of algebras for which the class $H(\mathbf{K})$ satisfies condition A) with $n=1$, then every algebra A in **K** has property (∗∗). The proof of this is immediate since transitivity will not be violated. This observation gives Theorem 3.1 as a corollary of Theorem 3.5.

We note that condition (∗∗) is not trivial.

Example 3.6. There exist finite algebras A and B and a homomorphism h from A onto B such that the algebra B has a principal congruence relation that is not the image under h of any principal congruence relation on A. For example, let $A=\{1,2,3,4,5,6\}$ and let A have two unary operations f and g given by:

	1	2	3	4	5	6
f	3	4	3	4	5	6
g	5	6	3	4	5	6.

Let h be the homomorphism of A which identifies only 4 and 5. The principal congruence $\Theta(h(1), h(2))$ in the algebra $h(A)$ is not the image under h of any principal congruence of A; for if it were, a simple argument shows it would be the image of the congruence $\Theta(1,2)$, but the image of this relation under the homomorphism h is not transitive.

We conclude this section by discussing the relation between n-step principal congruences and n-permutability of principal congruences. This is motivated, in part, by work of MAGARI [22].

If R and S are binary relations on the same set, then the composition of them, denoted $R\circ S$, consists of all pairs (x,y) for which there is some z such that $(x,z)\in R$ and $(z,y)\in S$. Let $R\circ R\circ\ldots\circ R$ with n factors be denoted by R^n. Also, $R^{-1}=\{(y,x)|(x,y)\in R\}$. The equivalence relations R and S are said to be n-permutable if

$$R\circ S\circ R\ldots=S\circ R\circ S\ldots, \text{ each side having } n \text{ factors.}$$

Thus 2-permutable is the usual permutable: $R\circ S=S\circ R$. If R and S are n-permutable then

$$R\vee S = R\circ S\circ R\ldots \quad (n \text{ factors}).$$

Theorem 3.7. *Let A be an algebra such that A and every homomorphic image of A has n-step principal congruences. Then the principal congruences of A are $2n+1$ permutable.*

Proof. We will prove a slightly stronger result: If Θ is any principal congruence relation, $\Theta=\Theta(a_0,a_1)$, and if Σ is any congruence relation, then

$$\Theta\vee\Sigma = \Sigma\circ\Theta\circ\Sigma\circ\ldots\circ\Sigma \quad (2n+1 \text{ factors}).$$

To this end, let $b_0\equiv b_1$ $(\Theta\vee\Sigma)$. So there exists a sequence of elements $t_0, t_1, \ldots, t_k$ such that $t_0=b_0$, $t_k=b_1$, $t_{2i}\equiv t_{2i+1}$ Θ, and $t_{2i+1}\equiv t_{2i+2}$ Σ. Let h be a homomorphism with kernel Σ. So $h(t_{2i+1})=h(t_{2i+2})$ and $h(t_{2i})\equiv h(t_{2i+1})$ $\Theta(h(a_0), h(a_1))$. Therefore, $h(t_0)\equiv h(t_k)$ $\Theta(h(a_0), h(a_1))$. By hypothesis $h(A)$ has n-step principal congruences, so there exist polynomials p_i and elements z_i, $1\leqq i\leqq n$, and a switching function s to establish that $h(t_0)\equiv h(t_k)$ $\Theta(h(a_0), h(a_1))$ in $h(A)$. But then in

A, $p_i(a_{1-s(i)}, z_i) \equiv p_{i+1}(a_{s(i+1)}, z_{i+1})$ Σ and $p_i(a_0, z_i) \equiv p_i(a_1, z_i)$ Θ. Therefore, the chain $b_0 = t_0$, $p_i(a_{s(1)}, z_1)$, $p_1(a_{1-s(1)}, z_1)$, $p_2(a_{s(2)}, z_2)$, ..., $t_n = b_1$, establishes the claim.

Note that the algebra $A = \langle\{0, 1, 2, ...\}, f\rangle$ where $f(i) = i-1$, and $f(0) = 0$ has permutable congruences as do all its subalgebras and homomorphic images; but principal congruences in A require an arbitrary number of steps. Hence Theorem 3.7 has no natural converse in this local form. However, a reasonable converse for the global version might be:

Problem 3. If a variety **K** has the property that there is some m for which all principal congruences are m-permutable, does **K** have n-step principal congruences for some n.

4. Bounded steps and specified switching

We next investigate condition B) for a class **K** of algebras. This general problem has not received much attention in the literature. However, if **K** is a variety and if only one switching function is allowed, and if this function is a constant function, then such **K** have been studied in some detail, although from a different point of view. For the remainder of this section we confine ourselves to classes **K** satisfying condition B) and having only one switching function s.

As in section 2, the easiest example is furnished by sets. For if **K** is the variety of sets, then all principal congruences in **K** can be obtained using only 2 steps and with a switching function $s(1) = 0$ and $s(2) = 1$. Note of course that several different polynomials will be required in the different cases. One consequence of this example is that condition B), even with only one switching function, is not a Malcev condition. This has also been observed by P. Köhler. However, compare this with Theorem 4.2 below.

In [22] Magari considers the notion of a good n-family for a class **K** of algebras. In the case of $n=1$ this reduces to **K** having m-step principal congruences for some integer m and some fixed switching function s, with all of the $z_i \in \{a_0, a_1, b_0, b_1\}$. He shows in the proof on pp. 695—696 that if an algebra has this property then if Θ is any principal congruence relation and Σ is any congruence relation, then

$$\Theta \circ \Sigma \circ \Theta \circ \ldots \circ \Theta \subseteq \Sigma \circ \Theta \circ \Sigma \circ \ldots \circ \Sigma \quad (2m+1 \text{ factors in both}).$$

It follows then that in **K** principal congruence relations are $2m+1$ permutable.

We have shown in Theorem 3.7 that this result of Magari does not depend on fixing a particular switching function. In fact, by specifying the switching function (to be a constant) and working in a global setting, an even stronger result can be obtained.

We now consider varieties with n-permutable congruences. The following is due to HAGEMANN and MITSCHKE [19] and HAGEMANN [18].

Theorem 4.1. *For a variety* $\mathbf{K}$ *of algebras, the following are equivalent:*

(i) *The congruence relations of every algebra in* $\mathbf{K}$ *are n-permutable.*

(ii) *There exist ternary algebraic operations* $q_1, \ldots, q_n$ *on* $\mathbf{K}$ *such that*

$$q_1(x, y, y) = x, \; q_{i-1}(x, x, y) = q_i(x, y, y) \;\; and \;\; q_{n-1}(x, x, y) = y.$$

(iii) *For any* A *in* $\mathbf{K}$ *and any reflexive subalgebra* R *of* A^2, $R^{-1} \subseteqq R^{n-1}$.

(iv) *For any* A *in* $\mathbf{K}$ *and for any reflexive subalgebra* R *of* A^2, $R^n = R^{n-1}$.

The following is implicit in HAGEMANN and MITSCHKE [19] and is an unpublished result of H. Lakser. Also see CHAJDA and RACHUNEK [32].

Theorem 4.2. *For a variety* $\mathbf{K}$ *of algebras, the following are equivalent:*

(i) $\mathbf{K}$ *has* $n+1$ *permutable congruence relations.*

(ii) *There is a constant function* $s\colon \{1, \ldots, n\} \to \{0, 1\}$ *such that all principal congruences of algebras in* $\mathbf{K}$ *can be done in n steps using s as the switching function.*

Proof. To show (i)→(ii) let $A \in \mathbf{K}$ and suppose $c \equiv d \; \Theta(a, b)$ in A. Define a relation $R = \{(p(a, \mathbf{z}), p(b, \mathbf{z})) \mid p$ is any k-ary polynomial and $\mathbf{z}$ is any $k-1$ tuple of elements of A, $k = 1, 2, \ldots\}$. We wish to show $(c, d) \in R^n$. Note that the relation R is reflexive and is a subalgebra of A^2. Also, by Malcev's lemma, there exists an m such that $(c, d) \in R_1 \circ R_2 \circ \ldots \circ R_m$, where each R_i is either R or R^{-1}. By Theorem 4.1 $R^{-1} \subseteqq R^n$, so $(c, d) \in R^t$ for some $t \geqq m$. But again by Theorem 4.1, $R^t \subseteqq R^n$, and hence $(c, d) \in R^n$ as desired.

For the opposite direction, assume, without loss, that $s(i) = 0$ for all i. Let F be the free $\mathbf{K}$ algebra on the three free generators a, b, and c. Note $b \equiv a \; \Theta(a, b)$. So there exist polynomials $p_1, \ldots, p_n$ such that

$$b = p_1(a, \mathbf{z}_1); \quad p_i(b, \mathbf{z}_i) = p_{i+1}(a, \mathbf{z}_{i+1}), \quad a = p_n(b, \mathbf{z}_n).$$

Each $\mathbf{z}_i$ is a sequence of elements of F, and each element of F is itself a polynomial in the variables a, b, and c. Denote this sequence of polynomials by $\mathbf{z}_i(a, b, c)$. Finally, defining $q_i(x, y, w) = p_i(y, \mathbf{z}_i(w, x, w))$ gives the desired polynomial identities of Theorem 4.1.

Combining Theorem 4.2 with Theorem 3.2 we have:

Corollary 4.3. *If the variety* $\mathbf{V}$ *has n-permutable congruences for some n,* $\mathbf{K} \subseteqq \mathbf{V}$, *and* $Q(\mathbf{K})$ *has definable principal congruences, then* $V(\mathbf{K})$ *has definable principal congruences.*

Corollary 4.4. (Burris and Lawrence) *If* $\mathbf{K}$ *is a class of groups or rings and*

$Q(\mathbf{K})$ has definable principal congruences, then $V(\mathbf{K})$ has definable principal congruences.

Problem 4. Can the condition that $\mathbf{K}$ is a variety in Theorems 4.1 or 4.2 be relaxed in some way?

Problem 5. Are there any results, analogous to Theorem 4.2, for switching functions s that are not constant? Observe that the argument using the free algebra can still be used to produce polynomial identities.

With regard to Problem 5, Peter Köhler has observed that the variety of distributive lattices has a restricted uniform congruence scheme with $n=4$ and $s(1)=s(3)=0$ and $s(2)=s(4)=1$, but by a result of WILLE [29, p. 79], distributive lattices are not m-permutable for any integer m.

We conclude this section by exhibiting a few curiosities concerning 3-permutability. It is easily seen that if an algebra A has principal congruences that are 2-permutable (i.e. permutable), then all congruence relations of A are permutable. This is not the case for 3-permutability; witness e.g., sets have all principal congruences 3-permutable, but congruences in general for sets are not. Indeed, congruences on sets are not n-permutable for any n, since a constant switching function will not suffice for generating them.

If a class $\mathbf{K}$ has the property that $H(\mathbf{K})$ has 2-permutable congruences, then by Theorem 4.2 and by the remarks following the proof of Theorem 3.5, it follows that every algebra in $\mathbf{K}$ has property (* *). A similar result is possible for 3-permutability as well. In [29, Satz 6.19] WILLE proved that a variety $\mathbf{K}$ has the property that arbitrary homomorphisms preserve congruence relations iff $\mathbf{K}$ has 3-permutable congruences. We now present a local version of his result, via a similar proof, and thereby give another sufficient condition for (* *).

Theorem 4.5. *Let A be an arbitrary algebra. A congruence relation Θ of A is preserved by a homomorphism h iff $\Theta\circ\ker(h)\circ\Theta\subseteq\ker(h)\circ\Theta\circ\ker(h)$.*

Proof. To show h preserves Θ, it suffices to show $h(\Theta)$ is transitive. So let $(h(w), h(x))$ and $(h(y), h(z))$ be in $h(\Theta)$, with $h(x)=h(y)$. Then there exist w', x', y', and z' in A such that $h(w')=h(w)$, $h(x')=h(x)$, etc. with (w', x') and (y', z') in Θ. Note $h(x')=h(y')$. By hypothesis, $\exists u, v\in A$ such that $h(w')=h(u), h(z')=h(v)$, and $(u, v)\in\Theta$. Hence $h(w)=h(u), h(z)=h(v)$, and $(h(u), h(v))\in h(\Theta)$. So $(h(w), h(z))\in h(\Theta)$ as desired.

Conversely, let $(w, x), (y, z)\in\Theta$ with $h(x)=h(y)$. Apply h to give $(h(w), h(x))\in h(\Theta)$, $(h(y), h(z))\in h(\Theta)$. So there exist $w', z'\in A$ such that $h(w')=h(w)$, $h(z')=h(z)$ and $(w', z')\in\Theta$. This gives $(w, z)\in\ker(h)\circ\Theta\circ\ker(h)$ as desired.

Note that the variety **K** of 1-unary algebras in which $f(x)=f(y)$ for all x and y has the property that any homomorphism preserves principal congruences, but not arbitrary congruence relations.

A specialized version of Theorem 4.5 gives the next result. The proof is an easy modification of the proof of 4.5.

Theorem 4.6. *Let* **K** *be a class of algebras. The following are equivalent;*
(i) *every homomorphism of an algebra in* **K** *preserves principal congruences;*
(ii) *principal congruences are 3-permutable for algebras in* **K**.

5. Definable principal congruences

The notion of definable principal congruences was introduced in [3] in an attempt to describe the behavior of subdirectly irreducible algebras in a variety. Interest in the concept has continued, not only for its own sake, but also as a crucial hypothesis for other theorems in universal algebra and equational logic. One direction of research has been to classify varieties by whether or not they have definable principal congruences. This was done in [3] and has continued with the previously cited work of Burris and Lawrence [12], [13] for groups and rings, and by Baker [1] for groups. Negative results have also been obtained by Burris [11] who exhibited a 4-element algebra that generates a variety without definable principal congruences (but which does have distributive congruence lattices); by McKenzie [24] who showed that every nondistributive variety of lattices fails to have definable principal congruences; and by Taylor [28] who showed that the variety of commutative semigroups satisfying the law $xy=uv$ (which is generated by a 3-element semigroup), does not have definable principal congruences. In [6] it is shown that every 2-element algebra generates a variety with definable principal congruences. A useful theorem of McKenzie [24] states that if a variety **V** of finite similarity type has definable principal congruences and if there is a finite bound on the cardinality of subdirectly irreducible members of **V**, then **V** has a finite basis for its polynomial identities. (Also Jónsson [20] has a similar result.) This was used, for example, in [30] to show that a certain variety of upper bound algebras is finitely axiomatizable. Several strengthened versions of definable principal congruence relations have been given in the literature; some of these were discussed in section 2. Recently Tulipani [34] has shown that if a variety has definable principal congruences, then for any n there is a first order formula for describing the join of n principal congruence relations.

One possible way to obtain a positive solution to Problem 1 would be to show that whenever an algebra A has definable principal congruences with defining formula φ, then every homomorphic image of this algebra also has its principal congruences defined by φ. We now present another example to show that this is not the case. Note that we cannot use Example 3.4 since the algebra in that example did not have definable principal congruences.

Example 5.1. There exists a groupoid A and a homomorphic image B of A such that A has definable principal congruences and B does not.

Let the groupoid A have universe $\{0, 1, 2, \ldots\}$ and define $\oplus$ on A by

$$\begin{aligned} &1 \oplus x = x \oplus 1 = x \quad \forall x, \\ &0 \oplus 0 = 0, \\ &x \oplus y = x \quad \forall x, y > 1, \\ &0 \oplus x = x \oplus 0 = x \quad \forall x = 3i \pm 1, \quad i > 0, \\ &0 \oplus 3i = 3i-1 \quad \forall i > 0, \\ &3i \oplus 0 = 3i+1 \quad \forall i > 0. \end{aligned}$$

The principal congruence relations of A are listed below, where only nontrivial blocks are given and $x, y > 1$.

$$\begin{aligned} \Theta(0,1) &= /0, 1/2, 3, 4/5, 6, 7/\ldots \\ \Theta(0, x) &= /0, 2, 3, 4, \ldots/ \\ \Theta(1, x) &= /0; 1, 2, 3, \ldots/ \\ \Theta(x, y) &= /x, y/ \quad x, y \not\equiv 0 \pmod 3) \\ \Theta(x, y) &= /x-1, x, x+1, y/ \quad x \equiv 0 \pmod 3, \quad y \not\equiv 0 \pmod 3 \\ \Theta(x, y) &= /x-1, y-1/x, y/x+1, y+1/ \quad x, y \equiv 0 \pmod 3. \end{aligned}$$

All of these principal congruences can be achieved in at most six steps using unary algebraic polynomials of the form $q(x) = z_4 \oplus (z_2 \oplus (x \oplus z_1) \oplus z_3)$ where the z_i are in A. Thus, A has definable principal congruences. Let Θ denote the congruence relation

$$\Theta = \vee_i \Theta(3i+1, 3i+2), \quad i = 1, 2, \ldots.$$

Let h be any homomorphism with kernel Θ and let $B = h(A)$. Consider the principal congruence relation $\Theta(h(0), h(1))$ in B. $x \equiv y \;\; \Theta(0, 1)$ in the algebra A implies $h(x) \equiv h(y) \;\; \Theta(h(0), h(1))$. So the block of $\Theta(h(0), h(1))$ includes $\{h(3i-1), h(3i), h(3i+1)\}$, for all $i > 0$. Thus $\Theta(h(0), h(1)) = /h(0), h(1)/$ the rest of $B/$. But for any polynomial p,

$$\{p(h(0), h(\mathbf{z})), p(h(1), h(\mathbf{z}))\} \subseteq h(\{p(0, \mathbf{z}'), p(1, \mathbf{z}')|\; \mathbf{z}' \equiv \mathbf{z} \;\; \Theta\}) \subseteq$$

$$\subseteq h(\{3(i-1)-1,\; 3(i-1), 3(i-1)+1, 3i-1, 3i, 3i+1)\})$$

for some i, and so an arbitrarily large number of steps are required.

With regard to Example 5.1, note that $SP(A)$ does not have definable principal congruences. For let C be the subalgebra of A^n generated by $\{\mathbf{3}, \mathbf{4}, \mathbf{6}, \text{and } t^i\ (1 \leq i \leq n)\}$, wehre $\mathbf{k}$ is the n-tuple of all k's, and t^i consists of all 8's, except for $t_i^i=0$. We leave it to the reader to verify that $\mathbf{4} \equiv \mathbf{6}\ \Theta(\mathbf{3}, \mathbf{6})$ in the algebra C, and that polynomials of arity $n+1$ are required.

References

[1] K. BAKER, Definable normal closures in locally finite varieties of groups, preprint.

[2] J. T. BALDWIN, The number of subdirectly irreducibles in a variety. II, *Algebra Universalis,* **11** (1980), 1—6.

[3] J. T. BALDWIN and J. BERMAN, The number of subdirectly irreducible algebras in a variety, *Algebra Universalis,* **5** (1975), 379—389.

[4] B. BANASCHEWSKI and E. NELSON, Boolean powers as algebras of continuous functions, *Dissertationes Math.,* **179** (1980).

[5] B. BANASCHEWSKI and H. HERRLICH, Subcategories defined by implications, *Houston J. Math.,* **2** (1976), 149—171.

[6] J. BERMAN, A proof of Lyndon's finite basis theorem, *Discrete Math.,* **29** (1980), 229—233.

[7] J. BERMAN and G. GRÄTZER, Uniform representations of congruence schemes, *Pacific J. Math.,* **76** (1978), 301—311.

[8] J. BERMAN, G. GRÄTZER and C. PLATT, Extending algebras to model congruence schemes, preprint.

[9] W. BLOK, P. KÖHLER and D. PIGOZZI, The deduction theorem in algebraic logic. II, preprint.

[10] W. BLOK and D. PIGOZZI, The deduction theorem in algebraic logic. I, preprint.

[11] S. BURRIS, An example concerning definable principal congruences, *Algebra Universalis,* **8** (1977), 403—404.

[12] S. BURRIS and J. LAWRENCE, Definable principal congruences in varieties of groups and rings, *Algebra Universalis,* **9** (1979), 152—164.

[13] S. BURRIS and J. LAWRENCE, A correction to "Definable principal congruences in groups and rings", *Algebra Universalis,* to appear.

[14] A. DAY, A note on the congruence extension property, *Algebra Universalis,* **1** (1970), 234—235.

[15] E. FRIED, G. GRÄTZER and R. QUACKENBUSH, Uniform Congruence Schemes, *Algebra Universalis,* **10** (1980), 176—180.

[16] G. GRÄTZER, *Universal Algebra,* 2nd edition, Springer-Verlag (Berlin—New York, 1979).

[17] G. GRÄTZER and H. LAKSER, The structure of pseudocomplemented distributive lattices II. Congruence extension and amalgamation, *Trans. Amer. Math. Soc.,* **156** (1971), 343—358.

[18] J. HAGEMANN, Grundlagen der allgemeinen topologischen Algebra, preprint.

[19] J. HAGEMANN and A. MITSCHKE, On n-permutable congruences, *Algebra Universalis,* **3** (1973), 8—12.

[20] B. JÓNSSON, On finitely based varieties of algebras, *Colloq. Math.,* **42** (1979), 255—261.

[21] P. KÖHLER and D. PIGOZZI, Varieties with equationally defined principal congruences, *Algebra Universalis,* **11** (1980), 213—219.

[22] R. MAGARI, Classi e schemi ideali, *Ann. Scuola Norm. Sup. Pisa,* **27** (1973), 687—706.

[23] A. I. MALCEV, On the general theory of algebraic systems, *Mat. Sb.,* **35** (1954), 3—20. (Russian)

[24] R. MCKENZIE, Para primal varieties: A study of finite axiomatizability and definable principal congruences in locally finite varieties, *Algebra Universalis,* **8** (1978), 336—348.

[25] R. MCKENZIE and S. SHELAH, The cardinals of simple models for universal theories, in: *Tarski*

Symposium, Proc. of Symp. in Pure Math., vol. 25, Amer. Math. Soc. (Princeton, R. I., 1974), 53—74.

[26] R. MEKLER, On residual properties, *Proc. Amer. Math. Soc.,* **78** (1980), 187—188.

[27] W. TAYLOR, Residually Small Varieties, *Algebra Universalis,* **2** (1972), 33—51.

[28] W. TAYLOR, Letter dated 14 December, 1977.

[29] R. WILLE, *Kongruenzklassengeometrien,* Lecture Notes in Mathematics, vol. 113, Springer-Verlag (Berlin, 1970).

[30] S. WINKER and J. BERMAN, Finite basis and related results for an upper bound algebra, *Notices of Amer. Math. Soc.,* **26** (1979), A—427.

[31] K. BAKER, Nondefinability of projectivity in lattice varieties, preprint.

[32] I. CHAJDA and J. RACHUNEK, Relational characterizations of permutable and *n*-permutable varieties, preprint.

[33] E. FRIED and E. KISS, Connection between the congruence lattices and polynomial properties, *Algebra Universalis,* to appear.

[34] S. TULIPANI, On classes of algebras with the definability of congruences, preprint.

MATHEMATICS DEPARTMENT
UNIVERSITY OF ILLINOIS CHICAGO CIRCLE
CHICAGO, ILLINOIS 60680, U.S.A.

Acta Sci. Math. 44 (1982), 271—279

Rectangular bands in universal algebra: Two applications

MATTHEW GOULD

The concept of rectangular band, well known in semigroup theory (see, e.g., CLIFFORD and PRESTON [5]), has been utilized, implicitly or explicitly, in a few contexts in universal algebra, for example by CHANG, JÓNSSON, TARSKI [4], FAJTLOWICZ [7], GOULD [12], PŁONKA [23], NEUMANN [22], and TAYLOR [26]. In this note we employ rectangular bands to obtain two results concerning automorphism groups of universal algebras. The first of these results is a new proof of E.T. SCHMIDT's [24] theorem establishing the abstract independence of the concepts of automorphism group and subalgebra lattice. Although Schmidt's result has been re-proved and generalized in several ways (as in GOULD and PLATT [14], FRIED and GRÄTZER [9], LAMPE [21], and STONE [25]), the present proof is, in the author's view, simpler than the others and yields a stronger result in the finite case, namely the following:

Given a finite group G and a finite, non-trivial lattice L, there is a finite algebra $\mathfrak{A}$ of three binary operations, such that $G\cong \operatorname{Aut}\mathfrak{A}$ and $L\cong \operatorname{Sub}\mathfrak{A}$. Moreover, the three binary operations may be replaced by a single quaternary operation without altering the automorphisms or subalgebras.

Our second result is concerned with automorphism groups of direct products. It establishes, in a somewhat stronger form, the following statement:

Given a group G, there exist multi-unary algebras $\mathfrak{A}$ and $\mathfrak{B}$, each having a trivial automorphism group, such that $G\cong \operatorname{Aut}(\mathfrak{A}\times\mathfrak{B})$. Moreover, $\mathfrak{A}$ and $\mathfrak{B}$ are finite if G is finite.

Concepts and notations of universal algebra used here are taken from GRÄTZER [16], except for the notations $\operatorname{End}\mathfrak{A}$, $\operatorname{Aut}\mathfrak{A}$, $\operatorname{Sub}\mathfrak{A}$, $\operatorname{Con}\mathfrak{A}$, respectively denoting the endomorphism monoid, automorphism group, subalgebra lattice, and congruence lattice of a universal algebra $\mathfrak{A}$. Moreover, for sets A and B and an element $x\in A\times B$, the respective components of x will be denoted x_0 and x_1, that is, $x=\langle x_0, x_1\rangle$. The projections $\pi_0\colon A\times B\to A$ and $\pi_1\colon A\times B\to B$ are then given by $x\pi_i=x_i$, for $i=0, 1$.

Received July 25, 1980.

1. Automorphism groups and subalgebra lattices. Given sets A and B, define a binary operation on $A\times B$ by $x*y=\langle x_0, y_1\rangle$. The resulting groupoid is a semigroup known as the *rectangular band on* $A\times B$. In KIMURA [20] the following conditions on a semigroup S are proved to be equivalent:

(i) S satisfies the identities $xyz=xz$ and $x^2=x$.
(ii) S satisfies the identity $xyx=x$.
(iii) There exist sets A and B such that S is isomorphic to the rectangular band on $A\times B$.

A semigroup satisfying any of these equivalent conditions is called a *rectangular band.* (A generalization to Cartesian products of finitely many sets was introduced by PŁONKA [23] and termed *diagonal algebra;* see also FAJTLOWICZ [7].)

When A and B are endowed with operations of the same similarity type, the imposition of the rectangular band operation on the direct product of the algebras results in an algebra whose endomorphisms, subalgebras, and congruences readily decompose. Specifically, we have the following lemma, in which the notation $L_0\otimes L_1$ denotes, for algebraic lattices L_0 and L_1, the lattice obtained by adjoining a zero to the partial sublattice of $L_0\times L_1$ given by $\{x\in L_0\times L_1 \mid x_0\neq 0\neq x_1\}$. (The lemma contains a statement about congruence lattices that is included only for the sake of completeness, as it will not be used in the sequel. It was essentially noted by TAYLOR [26] and is a ready consequence of results of FRASER and HORN [8].)

Lemma 1.1. *Let* $\mathfrak{A}=\langle A; F\rangle$ *and* $\mathfrak{B}=\langle B; F\rangle$ *be universal algebras of the same similarity type, let* $\langle A\times B; F\rangle$ *denote their direct product, and let* $*$ *be the rectangular-band operation on* $A\times B$. *Then the algebra* $\mathfrak{C}=\langle A\times B; F\cup\{*\}\rangle$ *has the following properties.*

(1.1.1) $\mathrm{End}\,\mathfrak{C}\cong\mathrm{End}\,\mathfrak{A}\times\mathrm{End}\,\mathfrak{B}$, *and likewise for automorphisms.*
(1.1.2) $\mathrm{Sub}\,\mathfrak{C}\cong\mathrm{Sub}\,\mathfrak{A}\times\mathrm{Sub}\,\mathfrak{B}$ *if* F *contains nullary operations.*
(1.1.3) $\mathrm{Sub}\,\mathfrak{C}\cong\mathrm{Sub}\,\mathfrak{A}\otimes\mathrm{Sub}\,\mathfrak{B}$ *if* F *contains no nullary operations.*
(1.1.4) $\mathrm{Con}\,\mathfrak{C}\cong\mathrm{Con}\,\mathfrak{A}\times\mathrm{Con}\,\mathfrak{B}$.

Proof. As all parts of the lemma are proved in a very straightforward manner, we prove only (1.1.1) in detail and note that a common proof of (1.1.2) and (1.1.3) is achieved by verifying that $\mathrm{Sub}\,\mathfrak{C}=\{U\times V \mid U\in\mathrm{Sub}\,\mathfrak{A}$ and $V\in\mathrm{Sub}\,\mathfrak{B}\}$. The distinction between (1.1.2) and (1.1.3) arises from the convention that $\emptyset\in\mathrm{Sub}\,\mathfrak{A}$ if and only if $\mathfrak{A}$ has no nullary operations.

Given $\alpha\in\mathrm{End}\,\mathfrak{A}$ and $\beta\in\mathrm{End}\,\mathfrak{B}$, define $\gamma\colon A\times B\to A\times B$ pointwisely: $x\gamma=\langle x_0\alpha, x_1\beta\rangle$ for all $x\in A\times B$. Obviously $\gamma\in\mathrm{End}\,(\mathfrak{A}\times\mathfrak{B})$. Moreover, $(x*y)\gamma=\langle x_0, y_1\rangle\gamma=\langle x_0\alpha, y_1\beta\rangle=\langle x_0\alpha, x_1\beta\rangle*\langle y_0\alpha, y_1\beta\rangle=x\gamma*y\gamma$ for all $x, y\in A\times B$, and therefore $\gamma\in\mathrm{End}\,\mathfrak{C}$. Writing $\gamma=\alpha\times\beta$, we thus see that $\{\alpha\times\beta \mid \alpha\in\mathrm{End}\,\mathfrak{A}, \beta\in\mathrm{End}\,\mathfrak{B}\}$

is a submonoid of End $\mathfrak{C}$, and is obviously isomorphic to End $\mathfrak{A}\times$End $\mathfrak{B}$. Hence, to prove (1.1.1) it suffices to show that every endomorphism of $\mathfrak{C}$ belongs to this submonoid.

Let $\varphi\in$End $\mathfrak{C}$. Fixing $u\in A\times B$, define maps φ_0: $A\to A$ and φ_1: $B\to B$ by $a\varphi_0=\langle a, u_1\rangle\varphi\pi_0$ and $b\varphi_1=\langle u_0, b\rangle\varphi\pi_1$ for all $\langle a, b\rangle\in A\times B$. If we can show that φ_0 and φ_1 are independent of the choice of u, it will follow that φ_0 and φ_1 are endomorphisms of $\mathfrak{A}$ and $\mathfrak{B}$ respectively and that $\varphi=\varphi_0\times\varphi_1$. Now, for $v\in A\times B$ and a, b, u as above, we have $a\varphi_0=\langle a, u_1\rangle\varphi\pi_0=(\langle a, v_1\rangle*u)\varphi\pi_0=(\langle a, v_1\rangle\varphi*u\varphi)\pi_0= =\langle\langle a, v_1\rangle\varphi\pi_0, u\varphi\pi_1\rangle\pi_0=\langle a, v_1\rangle\varphi\pi_0$, and likewise $b\varphi_1=\langle v_0, b\rangle\varphi\pi_1$.

We can now prove the main result of this section.

Theorem 1.2. *Let L be an algebraic lattice of at least two elements and let G be a group.*

(1.2.1) (E. T. Schmidt [24]) *There is an algebra whose subalgebra lattice is isomorphic to L and whose automorphism group is isomorphic to G.*

(1.2.2) *If G is at most countable and each compact element of L contains at most countably many compact elements, then there is an infinite groupoid meeting the requirements of* (1.2.1).

(1.2.3) *If G and L are finite, there exists a finite algebra of three binary operations meeting the requirements of* (1.2.1); *there is also such a finite algebra having only a single quaternary operation.*

Proof. By classical results of Birkhoff and Frink [3] and Birkhoff [2] there exist algebras $\mathfrak{A}$ and $\mathfrak{B}$ having no nullary operations, such that Sub $\mathfrak{A}\cong L$, Aut $\mathfrak{B}\cong G$, and Sub $\mathfrak{B}\cong\mathbf{2}$, the two-element lattice. (For the sake of completeness we may specify that $\mathfrak{A}=\langle A;\ \{f_a|\,a\in A\}\cup\{\vee\}\rangle$, where A is the set of non-zero compact elements of L, the operation $\vee$ is the join inherited from L, and the unary operations f_a are given by $f_a(x)=a$ if $a\leqq x$ and $f_a(x)=x$ otherwise. Moreover, we may take $\mathfrak{B}=\langle G;\ \{\lambda_g|\,g\in G\}\rangle$, where $\lambda_g(x)=gx$ for all $x\in G$.)

Unfortunately, $\mathfrak{A}$ may have non-trivial automorphisms, and hence must be modified to eliminate such automorphisms without altering the subalgebras. To this end, let ϱ be a well-ordering of A and define a binary operation f_ϱ on A by setting $f_\varrho(x, y)=x$ if $x\varrho y$ and $f_\varrho(x, y)=y$ otherwise. Adjoining f_ϱ to the operations of $\mathfrak{A}$ we obtain an algebra $\mathfrak{A}_\varrho$ whose subalgebras are precisely those of $\mathfrak{A}$. Moreover, any automorphism of $\mathfrak{A}_\varrho$ is an automorphism of the well-ordered set $\langle A, \varrho\rangle$ and hence must be identity.

Although $\mathfrak{A}_\varrho$ and $\mathfrak{B}$ need not be of the same similarity type, the difference is, for our purposes, purely superficial. We can increase operational rank (e.g., by substituting for an n-ary operation f the $(n+1)$-ary operation f' given by $f'(x_1, \dots, x_{n+1})=f(x_1, \dots, x_n)$) and introduce projection operations (such as

$p(x_1, \dots, x_n) = x_1$) as needed in order to convert $\mathfrak{A}_\varrho$ and $\mathfrak{B}$ into algebras $\mathfrak{A}'$ and $\mathfrak{B}'$ sharing a common similarity type and maintaining the desired properties: $\mathrm{Sub}\,\mathfrak{A}' = \mathrm{Sub}\,\mathfrak{A}_\varrho = \mathrm{Sub}\,\mathfrak{A} \cong L$ and $\mathrm{Aut}\,\mathfrak{A}' = \mathrm{Aut}\,\mathfrak{A}_\varrho = \{1_A\}$; while $\mathrm{Sub}\,\mathfrak{B}' = \mathrm{Sub}\,\mathfrak{B} \cong \mathbf{2}$ and $\mathrm{Aut}\,\mathfrak{B}' = \mathrm{Aut}\,\mathfrak{B} \cong G$. As neither $\mathfrak{A}'$ nor $\mathfrak{B}'$ has nullary operations, the above lemma gives an algebra $\mathfrak{C}$ satisfying $\mathrm{Aut}\,\mathfrak{C} \cong G$ and $\mathrm{Sub}\,\mathfrak{C} \cong L \otimes \mathbf{2} \cong L$, whereupon (1.2.1) is proved.

To prove (1.2.2) and (1.2.3) we assume the hypotheses of (1.2.2) and begin by choosing the above $\mathfrak{A}$ and $\mathfrak{B}$ to be groupoids, of cardinality at most $|L|$ and $|G|$ respectively. By a result of HANF [17], such a groupoid $\mathfrak{A} = \langle A; f\rangle$ with $\mathrm{Sub}\,\mathfrak{A} \cong L$ indeed exists: essentially as noted by WHALEY [27] (see also JÓNSSON [19]) it suffices to take A to be the set of non-zero compact elements of L and to fix for each $x \in A$ an enumeration $\{x_0, \dots, x_n, \dots\}$ of the set $\{y \in A \mid y \leqq x\}$ in such a way that $x_i = x_j$ implies $x_{i+1} = x_{j+1}$. The required binary operation f can then be defined by $f(x, y) = x_{n+1}$ if $y = x_n$ for some n, and $f(x, y) = x \vee y$ otherwise.

In GOULD [10], a groupoid $\mathfrak{B}$ satisfying $\mathrm{Aut}\,\mathfrak{B} \cong G$ was defined on the set G as follows. First, choose an enumeration $\{g_0, \dots, g_n, \dots\}$ of G satisfying $g_0 = 1$ and $g_{i+1} = g_{j+1}$ whenever $g_i = g_j$. Then define the binary operation f by $f(x, y) = g_{n+1}x$, where $g_n = yx^{-1}$. Setting $\mathfrak{B} = \langle G; f\rangle$ we note that $\mathrm{Sub}\,\mathfrak{B} \cong \mathbf{2}$. Indeed, it suffices to show that each $x \in G$ generates (in $\mathfrak{B}$) the entire set G. Since $g_0 x = x$ and $f(x, g_n x) = g_{n+1} x$ for all n, it follows by induction that x generates $\{g_n x \mid n < \omega\} = Gx = G$.

Starting with these groupoids $\mathfrak{A}$ and $\mathfrak{B}$, we obtain as in the proof of (1.2.1) an algebra $\mathfrak{C}$ with $\mathrm{Sub}\,\mathfrak{C} \cong L$ and $\mathrm{Aut}\,\mathfrak{C} \cong G$. Note that $\mathfrak{C}$ has precisely three operations, all of them binary, and that the cardinality of $\mathfrak{C}$ is at most $|L| \cdot |G|$. By a result of JEŽEK [18], given any algebra $\mathfrak{C} = \langle C; F\rangle$ having at most countably many operations and no nullary operations, there is a groupoid $\mathfrak{C}'$ of cardinality equal to $\aleph_0 \cdot |C|$, such that $\mathrm{End}\,\mathfrak{C}' \cong \mathrm{End}\,\mathfrak{C}$ and $\mathrm{Sub}\,\mathfrak{C}' \cong \mathrm{Sub}\,\mathfrak{C}$. Applying this result to the $\mathfrak{C}$ given above, we have proved (1.2.2).

If L and G are finite, the above $\mathfrak{C}$ suffices to establish the first statement of (1.2.3). To prove the second statement, we distinguish two cases.

Case 1. Supposing $|G| = 1$, we in fact obtain a ternary operation with the required properties. Recall the algebra $\mathfrak{A}_\varrho = \langle A; f, f_\varrho\rangle$, and define a ternary operation t by setting $t(x, y, z) = f(x, y)$ if $y = z$ and $t(x, y, z) = f_\varrho(x, y)$ if $y \neq z$. It is easily verified that $\mathrm{Sub}\,\langle A; t\rangle = \mathrm{Sub}\,\langle A; f\rangle \cong L$ and $\mathrm{Aut}\,\langle A; t\rangle = \mathrm{Aut}\,\langle A; f_\varrho\rangle = \{1_A\} \cong G$.

Case 2. Supposing $|G| \neq 1$, we note that the groupoid $\mathfrak{B} = \langle B; f\rangle$ defined above has no one-element subgroupoids. It follows that in the direct product $\mathfrak{A} \times \mathfrak{B} = \langle A \times B; f\rangle$ there are no one-element subgroupoids. Passing to the algebra

$\mathfrak{C}$ above, we may write $\mathfrak{C}=\langle C; f_1, f_2, f_3\rangle$ where $C=A\times B$ and $f_1=f$. Now define a quaternary operation q on C by

$$q(x, y, z, w) = \begin{cases} f_1(x, y) & \text{if } y=z=w, \\ f_2(x, y) & \text{if } y=z\neq w, \\ f_3(x, y) & \text{if } y\neq z. \end{cases}$$

It is readily verified that $\operatorname{Aut}\langle C; q\rangle=\operatorname{Aut}\mathfrak{C}\cong G$ and that $\operatorname{Sub}\mathfrak{C}\subseteq\operatorname{Sub}\langle C; q\rangle$. To verify the reverse inclusion, let $X\in\operatorname{Sub}\langle C; q\rangle$. As the empty set belongs to $\operatorname{Sub}\mathfrak{C}$, we assume $X\neq\emptyset$. Since f_1 is a polynomial of $\langle C; q\rangle$ it follows that X is closed under f_1 and therefore contains more than one element. Thus, given $x, y\in X$ we may choose $w\in X$ with $w\neq y$, whereupon $f_2(x, y)=q(x, y, y, w)\in X$ and $f_3(x, y)= =q(x, y, w, w)\in X$. Hence $\operatorname{Sub}\langle C; q\rangle=\operatorname{Sub}\mathfrak{C}\cong L$ and the theorem is proved.

We close this section with some remarks concerning the above theorem. First, the hypothesis $|L|>1$ is obviously justified by the fact that an algebra with only one subalgebra can have only one automorphism. Second, the hypothesis in (1.2.2) concerning the compact elements is necessary because in the subalgebra lattice of any algebra of at most countably many operations, the compact elements are the finitely generated subalgebras, each of which is at most countable. The other hypothesis in (1.2.2), namely $|G|\leqq\aleph_0$, cannot in general be dispensed with. For example, if the unit element of L is compact, then every algebra having L as its subalgebra lattice must be finitely generated. Given an algebra $\mathfrak{A}=\langle A; F\rangle$ generated by a finite set S, the mapping that associates with each automorphism of $\mathfrak{A}$ its restriction to S is a one-to-one function of $\operatorname{Aut}\mathfrak{A}$ into A^S. As above, if F is at most countable, the fact that $\mathfrak{A}$ is finitely generated implies that A is at most countable, whereupon the same holds for A^S and hence for $\operatorname{Aut}\mathfrak{A}$ as well.

Finally, we note that the second statement in (1.2.3) improves a result of the author [11] establishing the existence of a finite algebra with the desired properties that has only one operation. We ask whether the rank of this operation can be reduced to two, thereby combining, in the finite case, the nicest features of (1.2.2) and (1.2.3). Precisely stated: *given a finite group G and a finite lattice L, does there exist a finite groupoid $\mathfrak{A}$ satisfying $G\cong\operatorname{Aut}\mathfrak{A}$ and $L\cong\operatorname{Sub}\mathfrak{A}$?* We conjecture that the answer is affirmative. (In this conjecture and in (1.2.2), "groupoid" is best possible in the sense that an algebra whose operations all have rank less than two must have a distributive subalgebra lattice.)

2. Automorphism groups of direct products. Given sets A and B, a binary operation $*$ is readily defined on the set of all functions of $A\times B$ into itself: for two such functions α and β, simply define $\alpha*\beta\colon A\times B\to A\times B$ to be the map that sends each $x\in A\times B$ to $\langle x\alpha\pi_0, x\beta\pi_1\rangle$. Straightforward calculation shows that $*$ is an associative operation satisfying the rectangular-band identities (i) cited in the previous section.

Moreover, it is readily observed that composition of mappings is left-distributive over $*$ in the sense that $\alpha(\beta*\gamma)=(\alpha\beta)*(\alpha\gamma)$ for all transformations α, β, γ of $A\times B$.

If now A and B are the carrier-sets of algebras $\mathfrak{A}$ and $\mathfrak{B}$ respectively, of the same similarity type, it is readily observed that $\mathrm{End}\,(\mathfrak{A}\times\mathfrak{B})$ is closed under $*$. Thus, we enrich the endomorphism monoid of $\mathfrak{A}\times\mathfrak{B}$ to form the *endomorphism system* $\mathcal{M}(\mathfrak{A}\times\mathfrak{B})=\langle \mathrm{End}\,(\mathfrak{A}\times\mathfrak{B});\ \cdot, *, 1\rangle$, an algebraic system of type $\langle 2,2,0\rangle$ in which $\cdot$ denotes composition of mappings and 1 is the identity endomorphism, here regarded as a nullary operation on $\mathrm{End}\,(\mathfrak{A}\times\mathfrak{B})$.

The following lemma shows that the aforementioned equational properties of $\mathcal{M}(\mathfrak{A}\times\mathfrak{B})$ actually characterize such endomorphism systems. Here and in the sequel, expressions of the form $x\cdot y$ will be written as xy, and $(xy)*(xz)$ will be written as $xy*xz$.

Lemma 2.1. *Let* $\mathcal{M}=\langle M;\ \cdot, *, 1\rangle$ *be an algebraic system of type* $\langle 2,2,0\rangle$. *The following conditions are jointly equivalent to the existence of algebras* $\mathfrak{A}$ *and* $\mathfrak{B}$ *such that* $\mathcal{M}\cong\mathcal{M}(\mathfrak{A}\times\mathfrak{B})$.

(2.1.1) $\langle M;\ \cdot, 1\rangle$ *is a monoid;*

(2.1.2) $\langle M;\ *\rangle$ *is a rectangular band;*

(2.1.3) $x(y*z)=xy*xz$ *for all* $x, y, z\in M$.

Moreover, given (2.1.1)—(2.1.3) *the algebras* $\mathfrak{A}$ *and* $\mathfrak{B}$ *can be chosen to have unary operations only and to be finite if* M *is finite.*

Proof. Having noted the converse, let us suppose that (2.1.1)—(2.1.3) hold. By the result of KIMURA [20] quoted in the previous section, there exist sets A and B such that $\langle M;\ *\rangle$ is isomorphic to $\langle A\times B;\ *\rangle$, the rectangular band on $A\times B$. As the other operations in $\mathcal{M}$ can be transferred to $A\times B$ by means of this isomorphism, we have a system $\mathcal{M}'=\langle A\times B;\ \cdot, *, 1\rangle$ satisfying (2.1.1)—(2.1.3) and isomorphic to $\mathcal{M}$.

Now define a multi-unary algebra $\mathfrak{C}=\langle A\times B;\ \{f_u|u\in A\times B\}\rangle$ in which f_u denotes left multiplication by u, that is, f_u maps each $x\in A\times B$ to ux. It readily follows (as in BIRKHOFF [2] or ARMBRUST and SCHMIDT [1]) that $\mathrm{End}\,\mathfrak{C}=\{\varrho_u|u\in A\times B\}$, where $x\varrho_u=xu$, for all $x\in A\times B$. The map $u\to\varrho_u$ is an isomorphism of $\langle A\times B;\ \cdot, 1\rangle$ onto $\mathrm{End}\,\mathfrak{C}$.

By (2.1.3) each f_u is in fact an endomorphism of $\langle A\times B;\ *\rangle$, whence it follows as in Lemma 1.1 that for each u there exist maps $f_u^A\colon A\to A$ and $f_u^B\colon B\to B$ such that $f_u(\langle a,b\rangle)=\langle f_u^A(a), f_u^B(b)\rangle$ for all $\langle a,b\rangle\in A\times B$. Hence $\mathfrak{C}=\mathfrak{A}\times\mathfrak{B}$, where $\mathfrak{A}=\langle A;\ \{f_u^A|u\in A\times B\}\rangle$ and $\mathfrak{B}=\langle B;\ \{f_u^B|u\in A\times B\}\rangle$. To conclude that the map $u\to\varrho_u$ is an isomorphism of $\mathcal{M}'$ onto $\mathcal{M}(\mathfrak{A}\times\mathfrak{B})$, it remains only to note that $\varrho_{u*v}=\varrho_u*\varrho_v$ for all $u, v\in A\times B$. Indeed, $x\varrho_{u*v}=x(u*v)=xu*xv=\langle (xu)\pi_0, (xv)\pi_1\rangle=\langle x\varrho_u\pi_0, x\varrho_v\pi_1\rangle=x(\varrho_u*\varrho_v)$ for all $x\in A\times B$, whereupon the lemma is proved.

Before stating the main result of this section, we borrow from EHRENFEUCHT and GRZEGOREK [6] the following definition. Given sets A and B, a function $\alpha\colon A\times B\to A\times B$ is said to be *axial* if either $\alpha\pi_0=\pi_0$ *(left axial)* or $\alpha\pi_1=\pi_1$ *(right axial)*. Clearly α is left axial if and only if $1*\alpha=\alpha$, and right axial if and only if $\alpha*1=\alpha$.

We now show that an arbitrary group can be realized as the automorphism group of the direct product of algebras whose automorphism groups are trivial. In fact, we have the following stronger result.

Theorem 2.2. *Given a group G, there exist multi-unary algebras $\mathfrak{A}$ and $\mathfrak{B}$, both finite if G is finite, such that $G\cong\mathrm{Aut}(\mathfrak{A}\times\mathfrak{B})$ and $\mathfrak{A}\times\mathfrak{B}$ has no axial automorphisms other than the identity; thus* Aut $\mathfrak{A}$ *and* Aut $\mathfrak{B}$ *are both trivial.*

Proof. By the above lemma, it suffices to construct a system $\langle M;\ \cdot, *, \bar{1}\rangle$ satisfying (2.1.1)—(2.1.3) and an isomorphism $g\to\bar{g}$ of G onto the group of units of $\langle M;\ \cdot, \bar{1}\rangle$, such that no $g\in G\setminus\{1\}$ satisfies $\bar{1}*\bar{g}=\bar{g}$ or $\bar{g}*\bar{1}=\bar{g}$. To this end, set $M=G\times G$, and for each $g\in G$ set $\bar{g}=\langle g, g\rangle$. Define multiplication in M by:

$$xy=\begin{cases}\langle x_0y_0, x_0y_1\rangle & \text{if} \quad x_0=x_1,\\ x \text{ otherwise}\end{cases}$$

and let $\langle M;\ *\rangle$ be the rectangular band on $G\times G$. It is evident that $\bar{1}$ is an identity element with respect to multiplication, and that the map $g\to\bar{g}$ is an isomorphism of G onto the group of invertible elements of $\langle M;\ \cdot, \bar{1}\rangle$. Moreover, if an element $g\in G$ satisfies $\bar{g}=\bar{1}*\bar{g}$, it follows that $\langle g, g\rangle=\langle 1, g\rangle$, whence $g=1$; likewise $\bar{g}=\bar{g}*\bar{1}$ implies $g=1$. Thus, (2.1.3) and the associativity of multiplication are all that remains to be proved.

Let $x, y, z\in M$. If $x_0=x_1$ we have $x(y*z)=x\cdot\langle y_0, z_1\rangle=\langle x_0y_0, x_0z_1\rangle=$ $=\langle x_0y_0, x_0y_1\rangle*\langle x_0z_0, x_0z_1\rangle=xy*xz$, while if $x_0\neq x_1$ we have simply $x(y*z)=$ $=x=x*x=xy*xz$, whence (2.1.3) is proved. As for associativity, first note that $x_0\neq x_1$ implies $x(yz)=x=xz=(xy)z$. Thus, we now assume $x_0=x_1$. If $y_0=y_1$ we then have $x(yz)=x\cdot\langle y_0z_0, y_0z_1\rangle=\langle x_0y_0z_0, x_0y_0z_1\rangle=(xy)z$. If $y_0\neq y_1$ it follows that $x_0y_0\neq x_0y_1$, whence $x(yz)=xy=\langle x_0y_0, x_0y_1\rangle=\langle x_0y_0, x_0y_1\rangle\cdot z=(xy)z$. Thus associativity is proved and Lemma 2.1 now gives the desired direct product having no non-trivial axial automorphisms. As a non-trivial automorphism of either factor would obviously give rise to a non-trivial axial automorphism of the direct product, the theorem is proved.

We close this section by remarking that the finiteness of the algebra constructed above, in the case where G is finite, stands in striking contrast to the author's construction in [12] establishing the fact that any group having an element of order two is isomorphic to Aut $(\mathfrak{A}\times\mathfrak{A})$ for some multi-unary algebra $\mathfrak{A}$ having only the trivial endomorphism: if the group has more than two elements the algebra $\mathfrak{A}$ in that construction will be infinite. However, a different construction given in [12] (and subsequently generalized to Aut ($\mathfrak{A}^n$) by the author and H. H. JAMES in [15]) produces a finite $\mathfrak{A}$ in the case where the given finite group retracts onto a two-element subgroup. One can easily verify that both constructions are free of non-trivial axial automorphisms.

The fact that finiteness is preserved in Theorem 2.3 makes it plausible that a finiteness-preserving construction can be found for Aut ($\mathfrak{A}\times\mathfrak{A}$) as well. By a result of the author in [13], representing a group as Aut ($\mathfrak{A}\times\mathfrak{A}$) for finite $\mathfrak{A}$ is equivalent to representing it as Aut $\mathfrak{B}$ for a finite algebra $\mathfrak{B}$ that is free on a two-element basis. Moreover, it is sufficient to allow only unary operations in the former case and binary in the latter. In view of the retraction theorem quoted above, the cyclic group of order four is the first group for which the question of such a representation is open.

Acknowledgments. The author gratefully acknowledges a grant from the Research Council of Vanderbilt University which, along with the generous hospitality of the Universität Mannheim and the Technische Hochschule Darmstadt, enabled him to pursue this work in the stimulating mathematical ambience of these host institutions.

Note added August 31, 1981. At the Symposium in Honor of Professor Bjarni Jónsson (Vanderbilt University, August 4—7, 1981), Ralph McKenzie announced a result that implies an affirmative solution to the conjecture stated at the end of §1.

References

[1] M. ARMBRUST and J. SCHMIDT, Zum Cayleyschen Darstellungssatz, *Math. Ann.*, **154** (1964), 70—72.

[2] G. BIRKHOFF, Sobre los grupos de automorfismos, *Revista Unión Mat. Argentina*, **11** (1946), 155—157.

[3] G. BIRKHOFF and O. FRINK, Representations of lattices by sets, *Trans. Amer. Math. Soc.*, **64** (1948), 299—316.

[4] C. C. CHANG, B. JÓNSSON and A. TARSKI, Refinement properties for relational structures, *Fund. Math.*, **55** (1964), 249—281.

[5] A. H. CLIFFORD and G. B. PRESTON, *The Algebraic Theory of Semigroups*, Mathematical Surveys no 7, vol. I-II, Amer. Math. Soc. (Providence, R. I., 1961, 1967).

[6] A. EHRENFEUCHT and E. GRZEGOREK, On axial maps of direct products (I), *Colloq. Math.*, **32** (1974), 1—11.

[7] S. FAJTLOWICZ, *n*-dimensional dice, *Rend. Mat.* (6), 4 (1971), 855—865.

[8] G. FRASER and A. HORN, Congruence relations in direct products, *Proc. Amer. Math. Soc.*, **26** (1970), 390—394.

[9] E. FRIED and G. GRÄTZER, On automorphisms of the subalgebra lattice induced by automorphisms of the algebra, *Acta. Sci. Math.*, **40** (1978), 49—52.

[10] M. GOULD, A note on automorphisms of groupoids, *Algebra Universalis*, 2 (1972), 54—56.

[11] M. GOULD, Automorphism and subalgebra structure in algebras of finite type, *Algebra Universalis,* **2** (1972), 369—374.

[12] M. GOULD, Endomorphism and automorphism structure of direct squares of universal algebras, *Pacific J. Math.,* **59** (1975), 69—84.

[13] M. GOULD, Automorphism groups of free algebras and direct powers, in: *Universal Algebra* (Proc. Conf. Esztergom, Hungary, 1977), Colloq. Math. Soc. J. Bolyai, vol. 29, North-Holland (Amsterdam, 1982); 331—334.

[14] M. GOULD and C. PLATT, Versatile monoids and versatile categories, *Algebra Universalis,* **1** (1971), 54—62.

[15] M. GOULD and H. H. JAMES, Automorphism groups retracting onto symmetric groups, *Pacific J. Math.,* **81** (1979), 93—100.

[16] G. GRÄTZER, *Universal Algebra,* second edition, Springer-Verlag (Berlin—New York, 1979).

[17] W. HANF, Representations of lattices by subalgebras, *Bull. Amer. Math. Soc.,* **62** (1956), 402.

[18] J. JEŽEK, Endomorphism semigroups and subgroupoid lattices, in: *Contributions to Universal Algebra,* (Proc. Conf. Szeged, Hungary, 1975), Colloq. Math. Soc. J. Bolyai, vol. 17, North-Holland (Amsterdam, 1977); 209—212.

[19] B. JÓNSSON, *Topics in universal algebra,* Lecture Notes in Mathematics, v. 250, Springer-Verlag (Berlin—New York, 1972).

[20] N. KIMURA, The structure of idempotent semigroups (I), *Pacific J. Math.,* **8** (1958), 257—275.

[21] W. A. LAMPE, The independence of certain related structures of a universal algebra, I—IV, *Algebra Universalis,* **2** (1972), 99—112, 270—283, 286—295, 296—302.

[22] W. D. NEUMANN, Representing varieties of algebras by algebras, *J. Austral. Math. Soc.,* **11** (1970), 1—8.

[23] J. PŁONKA, Diagonal algebras, *Fund. Math.,* **58** (1966), 309—321.

[24] E. T. SCHMIDT, Universale Algebren mit gegebenen Automorphismengruppen und Unteralgebrenverbänden, *Acta. Sci. Math.,* **24** (1963), 251—254.

[25] M. G. STONE, Subalgebra and automorphism structure in universal algebras: a concrete characterization, *Acta. Sci. Math.,* **33** (1972), 45—48.

[26] W. TAYLOR, The fine spectrum of a variety, *Algebra Universalis,* **5** (1975), 263—303.

[27] T. P. WHALEY, Algebras satisfying the descending chain condition for subalgebras, Ph. D. thesis, Vanderbilt University, 1968.

VANDERBILT UNIVERSITY
NASHVILLE, TENNESSEE 37235, U.S.A.

Acta Sci. Math. 44 (1982), 281—286

On derivations in generalized matroid lattices

M. STERN

1. Introduction. The notion of a generalized matroid lattice (GML for short) was introduced in [8]. A GML is a finite lattice L which satisfies the property that for each join-irreducible element $u \in L$ and for each element $b \in L$ the transposed intervals $[u \wedge b, u]$ and $[b, u \vee b]$ are isomorphic. The name "generalized matroid lattice" stems from the fact that many properties of matroid lattices (=geometric lattices= finite atomistic lattices with covering condition) can be proved in the original form or in a somewhat modified form also for GMLs. For results in this direction we refer to [8] and [2]. Similar results in a somewhat more general context appear in [7].

In the present paper we shall have a closer look on the uniquely determined lower cover u' of a join-irreducible element $u(\neq 0)$ in a GML. This lower cover u' will be called the (first) derivation of the join-irreducible element u. Concerning derivations of join-irreducible elements, we generalize in Section 3 some results of [6]. In the following Section 2 we deal with some basic facts which will be used in the sequel.

Acknowledgment. The author thanks the referee for his helpful remarks.

2. Some basic facts. By $x \prec y$ we mean that x is a lower cover of y.

Dealing with finite lattices, we define an element u to be join-irreducible if it has exactly one lower cover $u' \prec u$. In this context the least element 0 is not considered as join-irreducible. The uniquely determined lower cover u' of a join-irreducible element u will also be called the (first) derivation of u. By the n-th derivation $u^{(n)}$ of a join-irreducible element u we mean the element $(u^{(n-1)})'$. The element $u^{(n)}$ exists exactly if the elemente $u, u', u'', \ldots, u^{(n-1)}$ are all join-irreducible. By $J(L)$ we mean the set of all join-irreducible elements of a finite lattice L.

We restrict ourselves now to the abovementioned class of generalized matroid lattices.

Definition. (cf. [8]) Let L be a finite lattice and denote by $J(L)$ the set of all joinirreducible elements of L. We call L a *generalized matroid lattice* (briefly: GML) if the following isomorphism property (I) is satisfied:

Received January 19, 1981, and in revised form March 16, 1982.

(I) For all $u \in J(L)$ and for all $b \in L$ the transposed intervals $[u \wedge b, u]$ and $[b, u \vee b]$ are isomorphic.

The isomorphism indicated by this definition will be denoted by $[u \wedge b, u] \cong [b, u \vee b]$. The isomorphism property (I) implies (upper) semimodularity in the sense of [1].

Lemma 1. [8, Theorem 3] *A generalized matroid lattice is upper semimodular, that is, the implication*

$$a \wedge b \prec a \Rightarrow b \prec a \vee b$$

holds true.

By Lemma 1 the Jordan-Dedekind chain condition holds in a GML, i.e. all maximal chains of an interval $[x, y]$ have the same length (or dimension) which will be denoted by $d[x, y]$. This uniquely determined nonnegative integer $d[x, y]$ is called the *length* (or dimension) of the interval $[x, y]$.

Using arguments concerning the lengths of intervals, we show now that in a GML arbitrary join-irreducible elements and arbitrary elements form modular pairs in the sense of the following

Definition. Let a, b be elements of a lattice L. We say that (a, b) is a *modular pair*, and we write $(a, b)M$, if the implication

$$c \leqq b \Rightarrow (c \vee a) \wedge b = c \vee (a \wedge b) \quad (c \in L)$$

holds in L.

Lemma 2. *Let L be a generalized matroid lattice. Then $(u, b)M$ holds for each join-irreducible element $u \in J(L)$ and for each $b \in L$.*

Proof. We show that the implication

$$c \leqq b \Rightarrow (c \vee u) \wedge b = c \vee (u \wedge b)$$

holds in L. By [3, Lemma 1.4] it is sufficient to show that $(u, b)M$ holds in the interval $[u \wedge b, u \vee b]$. We may therefore assume that $u \wedge b \leqq c \leqq b$. By the isomorphism property (I) we have

$$[u \wedge b, u] = [u \wedge c, u] \cong [c, c \vee u],$$

which implies

$$d[c, c \vee u] = d[u \wedge b, u]. \tag{1}$$

By the isomorphism property (I) we have moreover

$$[u \wedge b, u] = [u \wedge \{b \wedge (c \vee u)\}, u] \cong [b \wedge (c \vee u), u \vee \{b \wedge (c \vee u)\}]$$

which yields

$$d[b \wedge (c \vee u), u \vee \{b \wedge (c \vee u)\}] = d[u \wedge b, u]. \tag{2}$$

Furthermore we have

$$c \leqq b\wedge(c\vee u) \leqq u\vee\{b\wedge(c\vee u)\} \leqq u\vee c. \tag{3}$$

The relations (1), (2), and (3) together imply now $u\vee\{b\wedge(c\vee u)\}=u\vee c$ and $c=b\wedge(c\vee u)$. This latter equality together with the trivial $c=c\vee(u\wedge b)$ implies the assertion.

We remark that Lemma 2 has also been proved by other methods in [8] and [2].

Finally we shall need

Lemma 3. *Let L be a generalized matroid lattice, $u\in J(L)$, $a\in L$ and $u'\vee a<u\vee a$. Then $u'\vee a\prec u\vee a$.*

Proof. From $u'\vee a<u\vee a$ it follows that $u\nleqq u'\vee a$. Hence it follows that $u\wedge(u'\vee a)=u'\prec u$. By the isomorphism property (I) we have then

$$[u', u] = [u\wedge(u'\vee a), u] \cong [u'\vee a, u\vee(u'\vee a)] = [u'\vee a, u\vee a],$$

that is, $u'\vee a\prec u\vee a$.

3. Properties of derivations. In this section we generalize some results of [6]. For other properties of derivations in a generalized matroid lattice we refer to [8].

Theorem 4. *Let L be a generalized matroid lattice, and let $u, u_1, \dots, u_k$ be join-irreducible elements of L. Moreover, let $u \leqq u_1\vee\dots\vee u_k$ where the join $u_1\vee\dots\vee u_k$ is (without loss of generality) assumed to be irredundant, and suppose*

$$u \nleqq u_1\vee\dots\vee u_{i-1}\vee u_i'\vee u_{i+1}\vee\dots\vee u_k$$

holds for all i $(1\leqq i\leqq k)$. Then

$$u' \leqq u_i'\vee\dots\vee u_k'.$$

Proof. If we had

$$u_1'\vee u_2\vee\dots\vee u_k = u_1\vee u_2\vee\dots\vee u_k$$

then it would follow that

$$u \leqq u_1\vee u_2\vee\dots\vee u_k = u_1'\vee u_2\vee\dots\vee u_k,$$

contradicting the assumptions of the theorem. Hence we have

$$u_1'\vee u_2\vee\dots\vee u_k < u_1\vee u_2\vee\dots\vee u_k.$$

By Lemma 3 we obtain now that

$$u_1'\vee u_2\vee\dots\vee u_k \prec u_1\vee u_2\vee\dots\vee u_k. \tag{4}$$

Moreover, $u \nleqq u_1'\vee u_2\vee\dots\vee u_k$ implies by the isomorphism property (I) that

$$[u\wedge(u_1'\vee u_2\vee\dots\vee u_k), u] \cong [u_1'\vee u_2\vee\dots\vee u_k, u_1\vee u_2\vee\dots\vee u_k].$$

Because of (4) we obtain from this that

$$(5) \qquad u \wedge (u_1' \vee u_2 \vee \ldots \vee u_k) \prec u.$$

But u is join-irreducible and has therefore exactly one lower cover u'. Hence (5) yields

$$u' = u \wedge (u_1' \vee u_2 \vee \ldots \vee u_k) \leqq u_1' \vee u_2 \vee \ldots \vee u_k.$$

Similarly, one shows that also

$$u' \leqq u_1 \vee \ldots \vee u_{j-1} \vee u_j' \vee u_{j+1} \vee \ldots \vee u_k$$

holds for $j=2, \ldots, k$. Thus we get

$$(6) \qquad u' \leqq \bigwedge_{j=1}^{k} (u_1 \vee \ldots \vee u_{j-1} \vee u_j' \vee u_{j+1} \vee \ldots \vee u_k).$$

We show now that

$$(7) \qquad \bigwedge_{j=1}^{k} (u_1 \vee \ldots \vee u_{j-1} \vee u_j' \vee u_{j+1} \vee \ldots \vee u_k) = u_1' \vee \ldots \vee u_k'$$

holds. Because of (6) this already implies the assertion of the theorem. In order to prove (7) we define for the sake of brevity

$$\bar{u}_m \stackrel{\text{def}}{=} u_1 \vee \ldots \vee u_{m-1} \vee u_{m+1} \vee \ldots \vee u_k.$$

According to the assumptions of the theorem we have $\bar{u}_m \vee u_m' < \bar{u}_m \vee u_m$. Applying again Lemma 3 we get

$$\bar{u}_m \vee u_m' \prec u_1 \vee \ldots \vee u_k.$$

From this we obtain by the isomorphism property (I) that

$$u_m \wedge (\bar{u}_m \vee u_m') \prec u_m$$

holds. Since u_m is join-irreducible, we have therefore

$$(8) \qquad u_m \wedge (u_1 \vee \ldots \vee u_{m-1} \vee u_m' \vee u_{m+1} \vee \ldots \vee u_k) = u_m'.$$

We now define

$$x_m \stackrel{\text{def}}{=} u_1 \vee \ldots \vee u_m \vee u_{m+1}' \vee u_{m+2} \vee \ldots \vee u_k, \quad z_m \stackrel{\text{def}}{=} u_{m+1},$$

$$y_m \stackrel{\text{def}}{=} u_1' \vee \ldots \vee u_m' \vee u_{m+2} \vee \ldots \vee u_k$$

with $m=1, \ldots, k-1$. Since we have $y_m \leqq x_m$, Lemma 2 implies $(z_m, x_m)M$. This means that we have

$$(9) \qquad (y_m \vee z_m) \wedge x_m = y_m \vee (z_m \wedge x_m).$$

According to the definitions of z_m and x_m we get from (8) the equation

$$z_m \wedge x_m = u'_{m+1}. \tag{10}$$

Substituting (10) in (9), by the definition of y_m we obtain the relation

$$(y_m \vee z_m) \wedge x_m = u'_1 \vee \dots \vee u'_m \vee u'_{m+1} \vee u_{m+2} \vee \dots \vee u_k.$$

This implies in particular

$$(y_1 \vee z_1) \wedge x_1 = (u'_1 \vee u_2 \vee u_3 \vee \dots \vee u_k) \wedge (u_1 \vee u'_2 \vee u_3 \vee \dots \vee u_k) =$$
$$= u'_1 \vee u'_2 \vee_3 \vee \dots \vee u_k = y_2 \vee z_2.$$

Hence we obtain

$$(y_2 \vee z_2) \wedge x_2 = (u'_1 \vee u'_2 \vee u_3 \vee \dots \vee u_k) \wedge (u_1 \vee u_2 \vee u'_3 \vee u_4 \vee \dots \vee u_k) =$$
$$= u'_1 \vee u'_2 \vee u'_3 \vee u_4 \vee \dots \vee u_k.$$

Continuing this procedure, we obtain finally (7) and the theorem is proved.

Now we consider as a special case those GMLs in which every join-irreducible element is a cycle. (By a *cycle* we mean a join-irreducible element z for which the interval $[0, z]$ is a chain). To these lattices we apply the preceding theorem in order to obtain an estimation for the dimension (=length) of a cycle. We remark that other properties of these special GMLs are considered in [4] and [5] in a slightly more general setting.

Corollary 5. *Let L be a generalized matroid lattice in which every join-irreducible element is a cycle. If $z, z_1, \dots, z_k$ are cycles of L and if $z \leqq z_1 \vee \dots \vee z_k$, then*

$$d[0, z] \leqq \max d[0, z_j] \quad (1 \leqq j \leqq k).$$

Proof. By Lemma 1 the lattice L is upper semimodular. Hence the interval $[0, z_1 \vee \dots \vee z_k]$ is also upper semimodular. The function d introduced in the preceding section is therefore a dimension function on the interval $[0, z_1 \vee \dots \vee z_k]$. Thus we have

$$d[0, z] \leqq d[0, z_1 \vee \dots \vee z_k] \leqq \sum_{j=1}^{n} d[0, z_j].$$

Without loss of generality we may assume that no z_j $(1 \leqq j \leqq k)$ can be replaced by a cycle $y_j < z_j$ in such a way that

$$z \leqq z_1 \vee \dots \vee z_{j-1} \vee y_j \vee z_{j+1} \vee \dots \vee z_k$$

also holds. This yields in particular that every derivation $z_j^{(n)}$ $(n = 1, 2, \dots)$ of z_j satisfies the relation

$$z \not\leqq z_1 \vee \dots \vee z_{j-1} \vee z_j^{(n)} \vee z_{j+1} \vee \dots \vee z_k. \tag{11}$$

Moreover, the join representation $z_1 \vee \ldots \vee z_k$ must be irredundant. Otherwise we would have a relation

$$z \leqq z_1 \vee \ldots \vee z_{j-1} \vee z_{j+1} \vee \ldots \vee z_k$$

contradicting (11) since there exists a natural number m with $z_j^{(m)}=0$.

Hence all assumptions of Theorem 4 are satisfied. Since all join-irreducible elements are cycles, Theorem 4 can be applied repeatedly. This yields for the n-th derivation

$$z^{(n)} \leqq z_1^{(n)} \vee \ldots \vee z_k^{(n)}. \tag{12}$$

Putting $t=\max d[0, z_j]$ $(1 \leqq j \leqq k)$ we have $z_j^{(t)}=0$ $(1 \leqq j \leqq k)$. Because of (12) we obtain therefore

$$z^{(t)} \leqq z_1^{(t)} \vee \ldots \vee z_k^{(t)} = 0,$$

that is, $z^{(t)}=0$. This implies $d[0, z] \leqq t$ which proves the corollary.

We remark that Theorem 4 and Corollary 5 were proved in [6] for the special case of modular lattices (of finite length) in which every join-irreducible element is a cycle. It should also be remarked that our proofs carry over without alteration to the case of lattices of finite length with isomorphism property (I).

References

[1] G. Grätzer, *General Lattice Theory*, Akademie-Verlag (Berlin, 1978).

[2] L. Kohl—M. Stern, A characterization of certain finite lattices in which every element is a join of cycles, in: *Contributions to Lattice Theory* (Proc. Conf. Szeged, 1980), Coll. Math. Soc. J. Bolyai, vol. 33, North-Holland, Amsterdam (to appear).

[3] F. Maeda—Sh. Maeda, *Theory of Symmetric Lattices*, Springer-Verlag (Berlin—Heidelberg—New York, 1970).

[4] G. Richter—M. Stern, A unified theory of cyclically generated modular lattices and *AC*-lattices, *Studia Sci. Math. Hungar.*, **11** (1976), 413—429.

[5] G. Richter—M. Stern, *ZI*-lattices, in: *Universal Algebra* (Proc. Conf. Esztergom, 1977), Colloq. Math. Soc. J. Bolyai, vol. 29, North-Holland (Amsterdam, 1982); 635—652.

[6] V. P. Soltan, Jordan elements of a lattice and subordinate sums (Russian), *Mat. Issled.*, **10** (1975), 230—237.

[7] M. Stern, Baer lattices I, II, III, *Wiss. Z. Martin-Luther-Univ. Halle*, **22** (1978), 135—140; **27** (1978), 65—69; **28** (1979), 21—24.

[8] M. Stern, Generalized matroid lattices, *Algebraic Methods in Graph Theory* (Proc. Conf. Szeged, 1978), Colloq. Math. Soc. J. Bolyai, vol. 25, North-Holland (Amsterdam, 1981); 727—748.

SEKTION MATHEMATIK
MARTIN-LUTHER-UNIVERSITÄT
UNIVERSITÄTSPLATZ 6
402 HALLE, GDR

Acta Sci. Math. 44 (1982), 287—297

The finite interpolation property for small sets of classical polynomials

ARTHUR KNOEBEL

0. Introduction

LAGRANGE's [7] interpolation formula tells us that an arbitrary operation on the real numbers may be matched at a finite number of points by some polynomial, that is, by an operation composed solely from addition, multiplication, and constants. Let us abstract the essence of this definition. Following FOSTER [3] and PIXLEY [9], we shall say that such a collection F of operations over a set S has the finite interpolation property if any other arbitrary operation can be matched at an arbitrary finite set of arguments by some composition of the operations of F together with constants of S. In other words, any partial operation defined on a finite subset of S can be extended to a composition of operations of $F \cup S$, defined on all of S.

The formulation of this concept immediately provokes the question of whether there are apparently weaker sets of operations which nevertheless have the finite interpolation property. For example, in KNOEBEL [4], less was required when the finite interpolation property was established over the reals for the set $\{+, {}^2\}$, where "2" is the operation of squaring. More generally, $\{+, {}^2\}$ has the finite interpolation property in any field not of characteristic 2. Similarly for multiplication, it was proven in KNOEBEL [5] that the set $\{\times, s\}$ has the finite interpolation property in any field where s is unit translation: $s(x)=x+1$.

The object of this article is to generalize these two results by replacing squaring and translation by rather arbitrary classical polynomials. We investigate in this paper four settings determined by two dichotomies: multiplication or addition over the complex or real numbers. In each of the four cases, we characterize those polynomials of one argument which together with the given binary operation yield the finite interpolation property over the given set.

The specific results are these. If p is a polynomial of degree at least two over the complex numbers $\mathbf{C}$, then the set $\{+, p\}$ has the finite interpolation property over $\mathbf{C}$, and conversely. Restricting the operations to the real numbers $\mathbf{R}$, we find that p

Received January 5, 1981, and in revised form April 8, 1982.

must also be of even degree or of odd degree with the leading coefficient negative for the set $\{+, p\}$ to have the finite interpolation property over $\mathbf{R}$.

To describe the situation with multiplication, we shall say that a polynomial is *cyclic of order* k if it is of the form

$$p(x) = c_0(x^k - c_1)(x^k - c_2)\dots(x^k - c_j)$$

for some constants $c_0, c_1, \dots, c_j$; the polynomial p is *cyclic* if it is cyclic of some order $k \geqq 2$. Then the set $\{\times, p\}$ has the finite interpolation property over $\mathbf{C}$ if, and only if, the polynomial p is not cyclic and not constant, and $p(0) \neq 0$. Over $\mathbf{R}$, we need only to avoid p being cyclic of order 2, but p should cross the x-axis.

The order of presentation is in order of increasing difficulty of the proofs. The method of proof applies standard results about classical polynomials to the theorems found in KNOEBEL [5], [6]. Numerous examples will illustrate the tightness of the hypotheses. Three open problems close the paper.

Needed in the sequel are certain definitions. By *polynomial* we mean a polynomial function in one variable in the classical sense, that is, a one-place function

$$p(x) = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \dots + \alpha_0$$

on $\mathbf{R}$ or $\mathbf{C}$, composed from addition, multiplication and constants. The degree of p is abbreviated 'deg p'. An *operation* ω on a set S is any n-place function $\omega: S^n \to S$ for some finite n. If F is a family of operations on a set S, then by an *F-polynomial* we understand a composition of operations from F together with constants from S. For example, over $\mathbf{R}$, a $\{+, \times\}$-polynomial is just a classical polynomial of $\mathbf{R}[x]$; a $\{+\}$-polynomial is a multilinear operation. With a little effort one can show that a $\{+, {}^2\}$-polynomial over $\mathbf{R}$ or $\mathbf{C}$ is any monic polynomial whose degree is a power of 2. We say that a family F of operations on a set S has the *finite interpolation property*, if, for every positive integer n, for every finite subset $T \subseteqq S^n$ and for every function $f: T \to S$, there is an F-polynomial ω such that f agrees with ω on T, that is, $f = \omega | T$. Briefly, we say that F has the f.i.p. over S.

A highly restricted version of this concept is that of (m, n)-transitivity, where m and n are positive integers. We say our family F of operations is *(m, n)-transitive* over S if, for every subset $T_m \subseteqq S$ of m elements, for every subset $T_n \subseteqq S$ of n elements and for every function $f: T_m \to T_n$, there is a composition $\omega: A \to A$ of operations in F such that $f = \omega | T_m$. Oftentimes, we wish to obtain (m, n)-transitivity by means of constants as well; this is most easily accomplished by the phrase, '$F \cup S$ is (m, n)-transitive.'

To prove that in our theorems the conditions on polynomials are tight enough and really necessary, we introduce the idea of preservation of properties. Let P be a property, that is, a finitary relation on S, say $P \subseteqq S^m$. We say P is *preserved* by an operation $\omega: S^n \to S$ if, whenever

$$(s_j^1, s_j^2, \dots, s_j^m) \in P \quad (\text{for all } j = 1, \dots, n),$$

and

$$s_0^i = \omega(s_1^i, s_2^i, \dots, s_n^i) \quad (\text{for all } i = 1, \dots, m),$$

then

$$(s_0^1, s_0^2, \dots, s_0^m) \in P.$$

Clearly, preservation passes through composition; that is, if P is preserved by $n+1$ operations $\omega_0, \omega_1, \dots, \omega_n$ where ω_0 is n-ary and $\omega_1, \dots, \omega_n$ are p-ary, then P is preserved by the composition $\omega_0(\omega_1, \dots, \omega_n)$. Typically, we apply this to show that a particular constraint is necessary for the f.i.p. by finding a property P which is preserved by all operations of F (and constants) and yet is *not* preserved by every conceivable operation needed to establish the f.i.p..

Some historical comments are in order. To gain more information on the origins and subsequent development of Lagrange's interpolation formula, one should start by looking in the index of GOLDSTINE [4]. Robin McLeod has pointed out to me that the finite interpolation property has acquired a meaning in contemporary universal algebra different from that in numerical analysis; see DAVIS [2] for the classical definition. Infinite universal algebras with the finite interpolation property are called 'functionally complete in the small' by FOSTER [3]. Each such algebra generates, in a natural way by the subextension of identities, a class of algebras each of which is isomorphic to a bounded subdirect power of the generating algebra. This theorem of Foster (Theorem 19.1, loc. cit.) is an infinite analog of STONE's [12] representation theorem for Boolean algebras. For related work on the finite interpolation property, the interested reader should consult the recent surveys by KNOEBEL [6], PIXLEY [9] QUACKENBUSH [10] and ROSENBERG [11].

1. Multiplication

In this section we give necessary and sufficient conditions, over both the real and complex numbers, for the set $\{\times, p\}$ to have the f.i.p.. Needed to prove these is the following result from an earlier paper.

Theorem 1.0 (KNOEBEL [5]). *If $F \cup S$ is (2,2)-transitive over S, and there is a binomial $\times$ in F with a null element 0 and a unit element 1 in S, that is, $0 \times s = 0 = s \times 0$, $1 \times s = s = s \times 1$ $(s \in S)$, then F has the finite interpolation property over S.*

With this we can prove the next theorem.

Theorem 1.1. *Let $p \in \mathbf{C}[x]$. Then $\{\times, p\}$ has the finite interpolation property over $\mathbf{C}$ if, and only if,*

(i) *p is not constant,*
(ii) $p(0)\neq 0$, *and*
(iii) *p is not cyclic.*

Proof. $\Rightarrow$ Let $F=\{\times, p\}$. By way of contradiction first assume $p(0)=0$. Then any non-constant F-monomial composed from $\times, p$ and constants will have the same property, and thus we do not even have (1,1)-transitivity.

Similarly, if p is cyclic of order k, let P_k be the binary relation holding for pairs (a, b) in $\mathbf{C}^2$ whenever there is a c in $\mathbf{C}$ such that a and b are both roots of z^k-c, i.e., $b=ae^{2\pi ij/k}$ for some integer j. Clearly p, as well as multiplication and constants, preserve P_k, so F is not (2,2)-transitive.

$\Leftarrow$ In view of Theorem 1.0, we need only to establish the (2,2)-transitivity of $F\cup\mathbf{C}$. To this end, let a and b in $\mathbf{C}$ be distinct with A and B also in $\mathbf{C}$; we are looking for a $\{\times, p\}$-polynomial q such that $q(a)=A$ and $q(b)=B$. Without loss of generality, assume $a\neq 0$. We claim such a q can be found in the form

$$q(z) = \gamma\times p\,(\beta\times p(\alpha\times z)),$$

where α, β, γ are constants to be determined. Let r be a root of p such that $\frac{br}{a}$ is not also a root. Such a root always exists, since otherwise $r, \frac{b}{a}r, \left(\frac{b}{a}\right)^2 r, \ldots$ must all be roots. Since a polynomial has a finite number of roots, this must imply that $\left(\frac{b}{a}\right)^k=1$ for some $k\geqq 2$, and therefore p is cyclic, a contradiction.

Now a goes into A and b goes into B by the following sequence of polynomial transformations:

$$\begin{array}{ccccccccccc}
a & \longrightarrow & r & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & p(0) & \longrightarrow & A \\
 & \frac{r}{a}\times(\ ) & & p(\ ) & & \beta\times(\ ) & & p(\ ) & & \frac{A}{p(0)}\times(\ ) & \\
b & \longrightarrow & \frac{br}{a} & \longrightarrow & p\left(\frac{br}{a}\right) & \longrightarrow & p^{-1}\left(\frac{Bp(0)}{A}\right) & \longrightarrow & \frac{Bp(0)}{A} & \longrightarrow & B.
\end{array}$$

The last multiplication is possible since $p(0)\neq 0$. We may choose

$$\beta = \frac{p^{-1}\left(\frac{Bp(0)}{A}\right)}{p\left(\frac{br}{a}\right)}$$

since $\frac{br}{a}$ is not a root; by $p^{-1}\left(\frac{Bp(0)}{A}\right)$ we mean a fixed root z_0 of $p(z)-\frac{Bp(0)}{A}=0$, which always exists over the complex numbers. The foregoing does not work when

$A=0$. In this case choose

$$\beta = \frac{B}{p\left(\frac{b}{a}r\right)},$$

and finish two steps earlier.

For example, the set $\{\times, x^2+x+1\}$ has the f.i.p. over **C**, but $\{\times, x^2+1\}$ does not. Similarly, $\{\times, x+1\}$ has the f.i.p., whereas $\{\times, x^3+1\}$ does not.

When **C** is replaced by **R**, the conditions must be formulated differently, but the proof is similar. By the phrase '*f crosses the x-axis*' we mean that there are a and b in **R** such that

$$f(a) < 0 < f(b).$$

Note that this is a stronger condition than merely saying that f has a real root.

Theorem 1.2. *Let $p \in \mathbf{R}[x]$. Then $\{\times, p\}$ has the finite interpolation property over* **R** *if, and only if,*

(i) *p is not constant,*
(ii) $p(0) \neq 0$,
(iii) *p is not cyclic of order 2, and*
(iv) *p crosses the x-axis.*

Proof. $\Rightarrow$ We show the contrapositive. If $p(0)=0$, then 0 cannot be taken into a nonzero element by any $\{\times, p\}$-polynomial.

If p does not cross the x-axis, then it is all of one sign, and consequently no polynomial in $\times$ and p can take two numbers of the same sign into numbers of opposite sign. More precisely, letting $P=\{\langle a, b\rangle | ab \geqq 0\}$, we see that all operations of $\{\times, p\} \cup \mathbf{R}$ preserve P; hence, e.g., $x+1$ is not a $\{\times, p\}$-polynomial in this case.

If p is cyclic of order 2, that is,

$$p(x) = r_0(x^2-r_1)(x^2-r_2)\dots(x^2-r_j),$$

and also $a=-b\neq 0$, then any $\{\times, p\}$-polynomial q will give $|q(a)|=|q(b)|$. Thus, for example, 1 and -1 cannot be taken into 1 and 2 respectively. The preservation relation in this case is $P=\{\langle a, \pm a\rangle | a \in \mathbf{R}\}$.

$\Leftarrow$ We need only modify the proof developed in the complex case. We assume that the reader now has before himself the sequence of polynomial transformations of the previous proof. Note that the only real roots of unity are $r=\pm 1$ and therefore, for the proof to work over **R**, we need only rule out polynomials which are cyclic of order 2 in the first transformation of multiplying by $\frac{r}{a}$.

The only other steps that might be different for the real case are the third and fourth, which depend on finding a root of

$$p(x) = \frac{Bp(0)}{A}.$$

Such a root may not exist when p is of even degree. In such an eventuality, certainly

$$p(x) = -\frac{Bp(0)}{A}$$

has a root, since p crosses the x-axis. Using this root instead, we will end up with $\langle a, b\rangle$ going to $\langle A, -B\rangle$. However, redefining the third transformation as multiplication by $-\beta$ will rectify this unwanted sign.

By way of example, notice that if $p(x)=x^2+1$ or x^2+x+1, then $\{\times, p\}$ does *not* have the f.i.p. over $\mathbf{R}$, but if $p(x)=x+1$ or x^3+1, then it does.

2. Addition

We now turn to addition to see which polynomials achieve the f.i.p.. The proofs now are more complicated; we do the complex case first since it is simpler than the real. Both depend on the following theorem.

Theorem 2.0 (Knoebel [6]). *If* $F\cup S$ *is (3,2)-transitive over* S, *and there is a binomial* $+$ *in* F *with a unit element* 0 *in* S, *that is,* $0+s = s = s+0$ $(s\in S)$, *and such that* $s+s\neq 0$ *for some* $s\in S$, *then* F *has the finite interpolation property over* S.

Theorem 2.1. *Let* $p\in \mathbf{C}[z]$. *Then* $\{+, p\}$ *has the finite interpolation property over* $\mathbf{C}$ *if, and only if,* p *is of degree at least two.*

Proof. $\Rightarrow$ Let $F=\{+, p\}$. On the contrary if $\deg p \leqq 1$, then p is constant or linear, in which case only linear operations are obtainable by composition from $+$, p and elements of $\mathbf{C}$.

$\Leftarrow$ Let us show (2,2)-transitivity first. Because we have sums and constants, translations are available for use anywhere. If we wish to take a to A and b to B, it suffices to find a z_0 such that

$$p(z_0+\delta)-p(z_0) = \Delta$$

where $\delta=a-b$ and $\Delta=A-B$. Since $\deg p\geqq 2$, the left side has degree at least 1. Among the complex numbers there is a solution z_0 to this difference equation. Hence $\{+, p\}\cup\mathbf{C}$ is (2,2)-transitive.

For (3, 2)-transitivity, let us prove that for any distinct a, b, and c in $\mathbf{C}$ there must be a $\{+, p\}$-polynomial q for which $q(a)=q(b)\neq q(c)$. By repeated addition, Nz is a $\{+\}$-polynomial for any positive integer N. Consider the family of polynomials q_λ^N indexed by $N\in\mathbf{N}$:

$$q_\lambda^N(z) = p(z+\lambda)+Nz.$$

We claim one of these will do the trick. For each positive integer N there is a root λ_N in $\mathbf{C}$ of the equation

$$p(a+\lambda_N)+Na = p(b+\lambda_N)+Nb$$

since $\deg p\geqq 2$. Thus $q_{\lambda_N}^N(a)=q_{\lambda_N}^N(b)$ for all positive integers N. If for some N, $q_{\lambda_N}^N(b)\neq q_{\lambda_N}^N(c)$, we are finished.

If not, we will reach a contradiction. For in this worst case, we would have for all positive integers N,

$$q_{\lambda_N}^N(a) = q_{\lambda_N}^N(b) = q_{\lambda_N}^N(c).$$

In terms of p, this is the infinite family of equations

$$p(a+\lambda_N)+Na = p(b+\lambda_N)+Nb = p(c+\lambda_N)+Nc.$$

From these, upon eliminating N, we derive

$$\frac{p(a+\lambda_N)-p(b+\lambda_N)}{a-b} = \frac{p(b+\lambda_N)-p(c+\lambda_N)}{b-c}$$

for all positive integers N. But each side is a polynomial agreeing with the other side at an infinite number of points. Since a nonzero polynomial has a finite number of zeros, they must agree at all points:

$$\frac{p(a+z)-p(b+z)}{a-b} = \frac{p(b+z)-p(c+z)}{b-c}.$$

Hence the coefficients of z^{n-2} on each side are equal:

$$\alpha_n n(n-1)(a+b)+(n-1)\alpha_{n-1} = \alpha_n n(n-1)(b+c)+(n-1)\alpha_{n-1},$$

where the α_n come from $p(z)=\alpha_n z^n+\alpha_{n-1}z^{n-1}+\dots$. Therefore, $a=c$, a contradiction to the distinctness of a and c.

The preceding paragraph, together with the (2,2)-transitivity proven earlier, establishes the (3,2)-transitivity of $\{+, p\}\cup\mathbf{C}$. From Theorem 2.0 follows the f.i.p. of $\{+, p\}$.

By the way of illustration, note that $\{+, p\}$ has the f.i.p. over $\mathbf{C}$ when $p(x) = x^2+1$, x^2+x+1 or x^3+1, but not when $p(x)=x+1$.

Theorem 2.2. *Let* $p \in \mathbf{R}[x]$. *Then* $\{+, p\}$ *has the finite interpolation property over* $\mathbf{R}$ *if, and only if,*

(i) $\deg p \geqq 2$, *and*

(ii) $\deg p$ *is even; or when* $\deg p$ *is odd, the leading coefficient* α_n *of* p *is negative.*

Proof. $\Rightarrow$ As in Theorem 2.1, condition (i) cannot be dropped.

Now if p satisfies the first condition but fails the second, it must be that p is a nonlinear polynomial of odd degree with positive leading coefficient. For sufficiently large differences between a and b, p will magnify this difference, and thus there is no way in which such a large difference may be decreased. In more detail, one can find a positive real number m such that $p(a)-p(b)>m$ whenever $a-b>m$. Let $P=\{\langle a, b\rangle | b-a>m\}$. Then P is preserved by both $+$ and p. Hence $\{+, p\} \cup \mathbf{R}$ is not (2,2)-transitive, and consequently $\{+, p\}$ does not have the f.i.p..

$\Leftarrow$ Let us consider only the case of a polynomial of odd degree at least 3 with negative leading coefficient:

$$p(x) = \alpha_n x^n + \dots + \alpha_0$$

with $\alpha_n < 0$ and n odd and $n \geqq 3$. The case of a nonconstant polynomial of even degree is similar but less complicated, and so the necessary modifications will be left to the reader. Notice that, with addition and constants appropriately composed, all translations and their inverses are available. By suitable translations on both the x- and y-coordinates, we may thus, without loss of generality, safely assume that

(i) $p(0) = 0$,

(ii) $p(x) > 0$ if $x < 0$,

(iii) $p''(x) > 0$ if $x \leqq 0$.

We first show the (2,2)-transitivity of $\{+, p\} \cup \mathbf{R}$. Then later the argument will be modified to accommodate (3,2)-transitivity by showing that in any triple the first two numbers may be identified by some polynomial which keeps the third one distinct.

Assume $a, b, A, B \in S$ and $a>b$; we will try to find a $\{+, p\}$-polynomial f such that $f(a)=A, f(b)=B$. Set $\delta=a-b$ and $\Delta=A-B$. Again by the use of translations, we would be finished if we could find a λ in $\mathbf{R}$ such that

$$p(\lambda+\delta)-p(\lambda) = \Delta.$$

The required polynomial f would be $f(x)=p(x+\lambda-b)-p(\lambda+\delta)+A$. Such a root λ would exist if p were of even degree, since then the difference is of odd degree and always has a root in $\mathbf{R}$.

However, when p is of odd degree, this won't work in general, but it can be made to work with the following modifications. Remember that Nx is a $\{+\}$-polynomial for any positive integer N. Set

$$q_\lambda^N(x) = p(x+\lambda)+Nx.$$

To obtain (2,2)-transitivity now with q_λ^N instead of p, the constant λ should be chosen to be a root of

$$q_\lambda^N(a)-q_\lambda^N(b) = \Delta.$$

In terms of p, this is

$$(*) \qquad p(a+\lambda)-p(b+\lambda)+N\delta = \Delta.$$

Whether such a λ exists depends on the value of N. Now $p(a+\lambda)-p(b+\lambda)$ has leading term $n\alpha_n(a-b)\lambda^{n-1}$ with negative coefficient and even exponent. Since $a>b$, the equation $(*)$ will have a root λ if N is chosen sufficiently large. But N can be any positive integer, so this is always possible. Hence $\{+,p\}\cup\mathbf{R}$ is (2,2)-transitive.

To establish (3,2)-transitivity, we argue as follows. Still assuming p to be a polynomial of odd degree satisfying conditions (i) to (iii), we will take a and b both to 0 (possible by the (2,2)-transitivity just established), but we will do it carefully enough so c goes to a nonzero real number. First of all, the ordering on a, b, and c may be reversed by translating to the left beyong 0 and applying p. Therefore, without loss of generality, we may assume $a>b>c$ or $a>c>b$. Secondly, in the next paragraph we need $\lambda+a$ to be negative. By choosing N sufficiently large — perhaps larger than before — we may guarantee $\lambda+a$ to be negative.

Proceeding as before with (2,2)-transitivity, we transform both a and b into 0 by the $\{+,p\}$-polynomial

$$q(x) = q_\lambda^N(x)-q_\lambda^N(a)$$

where $\lambda+a<0$, and $N>0$. Set $C=q(c)$ and $\gamma=q_\lambda^N(a)$. Thus

$$\begin{aligned} 0 &= q(a) = p(a+\lambda)+Na-\gamma, \\ 0 &= q(b) = p(b+\lambda)+Nb-\gamma, \\ C &= q(c) = p(c+\lambda)+Nc-\gamma. \end{aligned}$$

Notice that the arguments of p are all negative. Recall that $p''(x)>0$ when $x<0$, and hence this is also true for q. If by some fluke $C=0$, we would have a concave segment agreeing with a straight line at three points, which is nonsense. Thus $\{+,p\}\cup\mathbf{R}$ is (3,2)-transitive.

We make two comments. This convexity argument for (3,2)-transitivity using three points in the real case could be replaced by the argument using an infinite number of points in the complex case. Secondly, as an illustration, notice that both

$\{+, x+1\}$ and $\{+, x^3+1\}$ can be shown not to have the f.i.p. over **R** by directly observing that both functions are monotonically increasing, and so all compositions must have this property. On the other hand, the sets $\{+, x^2+1\}$, $\{+, x^2+x+1\}$ and $\{+, -x^3+1\}$ all have the f.i.p..

3. Open problems

We close with three open problems suggested by the theorems of this paper.

1. The set $\{+, \cosh\}$ can be shown to have the f.i.p. over **R** by arguments similar to those used in this paper since the hyperbolic cosine is shaped like a parabola. The problem is to find an appropriate definition of 'polynomial-like' so that the results of this paper are still true for functions which are not polynomials but similar to them in behavior.

2. Replace addition or multiplication by an arbitrary polynomial in k variables, and give necessary and sufficient conditions for the set $\{p, q\}$ to have the f.i.p. when q is a polynomial in one variable. More generally, which sets F of polynomials in any number of variables have the f.i.p. over **C** or **R**? Probably for most polynomials p of two or more arguments, $\{p\}$ by itself has the f.i.p. The evidence for this is Murskiĭ's [8] theorem that on a finite set, the proportion of two-place operations with the f.i.p. to all two-place operations approaches 1 as the cardinality of the set increases without bound. Compare this with the result of Davies [1] that the proportion of two-place Sheffer operations to all two-place operations approaches $1/e$ as the size of the finite set increases without bound. (An operation is *Sheffer* if all other operations are obtainable from it by composition without the help of constants.) Most likely, algebraic geometry will be needed to settle the exceptional cases.

3. Extend these results beyond **R** and **C** to more general structures, say, all fields.

References

[1] R. O. Davies, On n-valued Sheffer Functions, *Z. Math. Logik Grundlag. Math.*, **25** (1979), 293—298.

[2] P. J. Davis, *Interpolation and Approximation.*, Dover Publ. (New York, 1975).

[3] A. L. Foster, Functional completeness in the small — algebraic structure theorems and identities, *Math. Ann.*, **143** (1961), 29—58.

[4] H. H. Goldstine, *A History of Numerical Analysis from the 16th through the 19th Century*, Springer-Verlag (Berlin—Heidelberg—New York, 1977).

[5] R. A. Knoebel, Simplicity vis-à-vis functional completeness, *Math. Ann.*, **189** (1970), 299—307.

[6] R. A. Knoebel, Further conclusions on functional completeness, *Fund. Math.*, **99** (1978), 93—112.

[7] J.-L. Lagrange, Sur l'usage des courbes dans la solution des problèmes (1795), *Oeuvres de Lagrange* (ed. by J.-A. Serret and G. Darboux), Vol. VII, 1877; pp. 271—287.

[8] V. L. Murskiĭ, The existence of a finite basis of identities, and other properties of "almost all" finite algebras (Russian), *Problemy Kibernet.*, **30** (1975), 43—56.

[9] A. F. Pixley, A survey of interpolation in universal algebra, in: *Universal Algebra* (Proc. Colloq. Esztergom, 1977), Colloq. Math. Soc. János Bolyai, vol. 29, North-Holland (Amsterdam, 1982); 583—607.

[10] R. W. Quackenbush, Primality: the influence of Boolean algebras in universal algebra, Appendix 5 in George Grätzer's, *Universal Algebra*, 2nd ed, Springer-Verlag (Berlin—Heidelberg—New York, 1979).

[11] I. G. Rosenberg, Completeness properties of multiple-valued logic algebras, Chapter 6 of *Computer Science and Multiple-valued Logic. Theory and Applications* (D. C. Rine, ed.), North-Holland (Amsterdam, 1977).

[12] M. H. Stone, A theory of representations for Boolean algebras, *Trans. Amer. Math. Soc.*, **40** (1936), 37—111.

DEPARTMENT OF MATHEMATICAL SCIENCES
NEW MEXICO STATE UNIVERSITY
LAS CRUCES, NEW MEXICO 88003, U.S.A.

Acta Sci. Math. 44 (1982), 299—305

Additive functions with regularity properties

I. KÁTAI

1. Recently J. L. MAUCLAIRE and LEO MURATA [1] proved that a multiplicative function $g(n)$, satisfying the conditions

$$|g(n)| = 1 \quad (n = 1, 2, \ldots),$$

and

$$\frac{1}{x}\sum_{n\le x}|g(n+1)-g(n)| \to 0 \quad (x\to\infty)$$

has to be completely multiplicative. For a real z let $\|z\|$ denote its distance from the nearest integer. Their theorem is equivalent with the following assertion: If f is additive and

$$\frac{1}{x}\sum_{n\le x}\|f(n+1)-f(n)\| \to 0 \quad (x\to\infty) \tag{1.1}$$

then f is completely additive.

I conjecture that the following assertion is true: If f is an additive function satisfying (1.1), then $f(n)=c\log n+g(n)$, where $g(n)$ is an integer valued additive function.

In [2] the following simple assertion was proved: If $f(n)$ is additive and $n\|f(n+1)-f(n)\|=O(1)$, then $f(n)=c\log n+g(n)$, where $g(n)$ is an integer valued additive function. Now we prove the following stronger

Theorem 1. *If $f(n)$ is additive and*

$$n\|f(n+1)-f(n)\| = O(n^{\gamma}) \tag{1.2}$$

with a constant $\gamma<1$, then $f(n)=c\log n+g(n)$, where $g(n)$ is integer valued.

Proof. By the cited result of Mauclaire and Murata, we may assume that f is completely additive. Let I_n be the nearest integer to $(f(n+1)-f(n))$, and $\sigma(n)= =(f(n+1)-f(n))-I_n$. Then we have $\sigma(n)\in\left[-\frac{1}{2},\frac{1}{2}\right]$, and from (1.2), $n|\sigma(n)|= =O(n^{\gamma})$. Let

$$T(x) = \max_{m\le x} m|\sigma(m)|.$$

Received October 1, 1981.

We shall prove step by step the following assertions:

(1) The assertion is true if $T(x)=O(1)$ $(x\to\infty)$.

(2) If $T(x)=O(x^\gamma)$, $\gamma<1$, then $T(x)=O(\log x)$.

(3) If $T(x)\to\infty$, then the fractional parts of $m\sigma(m)$ are everywhere dense in $[0, 1)$.

(4) Completion of proof.

We start from the identity

$$f((n+1)^2-1)-f((n+1)^2)=f(n)-f(n+1)+f(n+2)-f(n+1),$$

which by $\sigma(n)\to 0$ implies that

$$(*) \qquad \sigma(n+1)=\sigma(n)-\sigma((n+1)^2-1) \quad \text{if} \quad n>n_0.$$

Applying this identity for $n+1, \ldots, n+H-1$ instead of n, we get

$$\sigma(n+H)-\sigma(n)=\sum_{j=0}^{H-1}(\sigma(n+j+1)-\sigma(n+j))=-\sum_{j=1}^{H}\sigma((n+j)^2-1),$$

so that

$$\sum_{H=0}^{R-1}\sigma(n+H)-R\sigma(n)=-\sum_{l=1}^{R-1}\sigma((n+l)^2-1)(R-l) \quad (n>n_0).$$

Let $n=mR$ and observe that

$$(1.3) \qquad \sum_{H=0}^{R-1}\sigma(mR+H)=f(mR+R)-f(mR)-\sum_{H=0}^{R-1}I_{mR+H}= \\ =\sigma(m)+(I_m-\sum I_{mR+H}).$$

The absolute value of the left hand side of (1.3) is not greater than

$$\frac{RT((m+1)R)}{mR}=\frac{T((m+1)R)}{m},$$

that is, less than 1/2 provided $m>m_0$, and R is not too large. Consequently it is $\sigma(m)$; therefore

$$(1.4) \qquad \sigma(m)-R\sigma(mR)=-\sum_{l=1}^{R-1}\sigma((mR+1)^2-1)(R-1),$$

if $T((m+1)R)<\frac{m}{2}$. The right hand side of (1.4) is majorated by

$$\frac{R^2T((m+1)^2R^2)}{m^2R^2},$$

hence

$$(1.5) \qquad |m\sigma(m)-Rm\sigma(mR)|\leqq\frac{T((m+1)^2R^2)}{m}.$$

Assume now that $T(x)$ is bounded, $T(x) \leqq K$. Putting $m=N_1$, $r=N_2$, and $m=N_2$, $R=N_1$ into (1.5) and using the triangle inequality we get

$$|N_1\sigma(N_1)-N_2\sigma(N_2)| \leqq \frac{K}{N_1}+\frac{K}{N_2}$$

for every large N_1, N_2. This shows that $N\sigma(N)$ is a Cauchy sequence, consequently $N\sigma(N)\to A$.

Let $\sigma(m)=\frac{A}{m}+\frac{\varepsilon_m}{A}$, $\varepsilon_m\to 0$. Furthermore, let p and q be arbitrary integers satisfying the relations: $1<q/p$, $A\log q/p<1/2$. Consider the relation

$$f(q)-f(p) = \sum_{n=pU}^{qU-1}(f(n+1)-f(n)) = \sum_{n=pU}^{qU-1}\sigma(n)+J(U),$$

where $J(U)$ is an integer depending on U. The sum on the right is

$$A\sum_{m=pU}^{qU-1}\frac{1}{m}+\sum_{m=pU}^{qU-1}\frac{\varepsilon m}{m} = A\log\frac{p}{q}+o_U(1)$$

as $U\to\infty$. Hence

$$f(q)-f(p)-A\log\frac{q}{p} = J(U)+o_U(1),$$

which shows that $J(U)$ is constant for $U>U_0(p,q)$. Consequently for $U\to\infty$ we get that $f(q)-f(p)-A\log q/p$ is an integer, which immediately implies our assertion.

Assume now that $T(x)=O(x^\gamma)$, $T(x)>Kx^\gamma$. Using (1.4) with $R=2$ we get

$$2m\sigma(2m) = m\sigma(m)+m\sigma((2m+1)^2-1). \tag{1.6}$$

Furthermore, from (*) we get

$$(2m+1)\sigma(2m+1) = \left(m+\frac{1}{2}\right)\sigma(m)-\frac{2m+1}{2}\sigma((2m+1)^2-1). \tag{1.7}$$

Let $x>x_0$ and assume that $T(2x)>T(x)$. The maximum of $|n\sigma(n)|$ in $[1,2x]$ is reached in $\left(\frac{x}{2}, x\right]$. If the maximum is taken for even n, then from (1.6),

$$T(2x) \leqq T(x)+\max_{m\in\left(\frac{x}{2},x\right]} m|\sigma((2m+1)^2-1)|.$$

Since $(2m+1)^2-1=(2m)(2m+2)\leqq 2x(2x+2)=4x(x+1)$, the last term is majorated by $x^{-1}T(4x(x+1))$, therefore

$$T(2x) \leqq T(x)+\frac{T(4x(x+1))}{x}.$$

Assume that the maximum is reached for $2m+1\in(x, 2x]$. Applying (1.7), as earlier, we deduce

$$T(2x) \leqq \left(1+\frac{1}{x-1}\right)T(x)+\frac{1}{2(x-1)}T(4x^2).$$

Since $4x(x+1)<8x^2$, $2(x-1)>x$ for large x and $T(x)/x\to 0$, we have

$$T(8x) \leqq T(x)+\frac{T(8x^2)}{x}+\varepsilon_x,$$

where $\varepsilon_x = \frac{T(x)}{x} \to 0.$

Assume that $\gamma>1/2$. Then

$$\frac{T(8x^2)}{x} < x^{2\gamma-1}, \quad \varepsilon_x \ll x^{2\gamma-1},$$

so that

(1.8) $$T(2x) \leqq T(x)+cx^{2\gamma-1}$$

for $x\geqq x_0$. Putting $x_k=2^k x_0$ $(k=0, 1, 2, \dots, N-1)$, we deduce that

(1.9) $$T(2^N x_0) \leqq T(x_0)+c\sum x_k^{2\gamma-1} \ll (2^N x_0)^{2\gamma-1}.$$

By the monotonity of T, we have $T(x)=O(x^{2\gamma-1})$. So we have proved the following assertion: If $\frac{1}{2}<\gamma<1$, and $T(x)\ll x^\gamma$, then $T(x)\ll x^{2\gamma-1}$. Repeating this argument for $\gamma=\gamma_1$, $2\gamma_1-1=\gamma_2, \dots$, in finitely many steps we get an exponent $\gamma_i\in(0, 1/2)$ such that $T(x)\ll x^{\gamma_i}\ll x^{1/2}$. Assume now that $\gamma=1/2$. Then (1.8) holds, i.e. $T(2x)\leqq T(x)+c$, and instead of (1.9) we get

$$T(2^N x_0) \leqq T(x_0)+O(N).$$

Consequently $T(x)=O(\log x)$. Since for $\gamma<1/2$ we have $T(x)\ll x^\gamma\ll x^{1/2}$, therefore we have $T(x)=O(\log x)$ whenever $T(x)\ll x^\gamma$, $\gamma<1$.

Now let m_1, m_2 be chosen so that

$$T\big((m_1+1)m_2\big) < \frac{1}{2}m_2, \quad T\big((m_2+1)m_1\big) < \frac{1}{2}m_1.$$

This implies (1.4). From (1.5) we deduce that

(1.10) $$|m_1\sigma(m_1)-m_2\sigma(m_2)| \leqq K(\log m_1 m_2)\left(\frac{1}{m_1}+\frac{1}{m_2}\right)$$

holds with a suitable constant K for every pair m_1, m_2 satisfying

(1.11) $$(c_1<)\, m_1 < m_2 < e^{\sigma m_1}$$

with a small positive constant σ, and a positive constant c_1. From (1.10) we get

$$\sigma(m_2)-\frac{m_1}{m_2}\sigma(m_1)=B\frac{\log m_1 m_2}{m_1 m_2}$$

for m_1, m_2 satisfying (1.11), where B is a bounded variable.

Now let $m_1=U$ and m_2 run over the interval $[U, 2U-1]$. Then we have

$$f(2)=J(U)+\sigma(U)U\sum_{m_2=U}^{2U-1}\frac{1}{m_2}+O\left(\frac{\log U}{U}\right). \tag{1.12}$$

From (1.12) we get immediately that $m\sigma(m)$ varies slowly. Consequently, if $m\sigma(m)$ is not bounded then the set of the fractional parts of $m\sigma(m)$ is a dense subset in [0, 1). Let $\alpha\in[0, 1)$ be chosen so that $\{f(2)\}\neq\{\alpha\log 2\}$. Let U_j be an infinite sequence such that $\{\sigma(U_j)U_j\}\to\alpha$. Putting $U=U_j$ into (1.12), and taking into account that

$$\sum_{U_j}^{2U_j-1} 1/m_2\to\log 2,$$

we get that $\{f(2)\}=\{\alpha\log 2\}$, which contradicts our assumption.

The proof of our theorem is complete.

2. Let $f(n)$ be a completely additive function, $N_1<N_2<\dots$ an infinite sequence of integers, $J_N=[N, N+(2+\varepsilon)\sqrt{N}]$, and $\varepsilon>0$ a constant.

Theorem 2. *If*

$$f(n)\equiv\alpha_j \pmod 1 \quad \textit{for} \quad n\in J_{N_j} \quad (j=1,2,\dots) \tag{2.1}$$

where $\alpha_1, \alpha_2, \dots$ are arbitrary real numbers, then $\alpha_1=\alpha_2=\dots=0$ and $f(n)$ takes on integer values only.

Proof. The method of proof is almost the same as that used in [3]. First we prove the following

Lemma. *Let $1<v_1<v$, v_1, v be constants. Assume that $f(x)\equiv\alpha \pmod 1$ in the interval $J_N=[N, vN]$. Then for every $N\geq N_0(v_1, v)$ we have*

$$f(n)\equiv 0 \pmod 1, \quad n<(v-v_1)N.$$

Let p, q be arbitrary integers satisfying the conditions

$$p<q<v_1 p, \quad q<(v-v_1)N. \tag{2.2}$$

For $m=\left[\frac{N}{p}\right]+1$ we get

$$N<pm<qm<q\left(\frac{N}{p}+1\right)<\frac{q}{p}N+q<v_1 N+(v-v_1)N=vN,$$

consequently for every pair p, q satisfying (2.2),

$$f(p) \equiv f(n+1) \pmod 1, \quad n \in \left(\frac{1}{v_1-1}, (v-v_1)N-1\right).$$

Let $f(n) \equiv \gamma \pmod 1$. Let n be chosen so that $(v_1-1)^{-2} < n^2 < (v-v_1)N-1$. Then $f(n^2) \equiv \gamma \pmod 1$, and so $\gamma = 0$. Hence

$$f(n) \equiv 0 \pmod 1, \quad \frac{1}{v_1-1} \leqq n \leqq (v-v_1)N-1.$$

It remains to prove that $f(k) \equiv 0 \pmod 1$ for $k \leqq (v_1-1)^{-1}$. Putting $m = \left[\frac{1}{v_1-1}\right]+1$, and letting N to be large, we have $f(km) \equiv 0 \pmod 1$, and $f(m) \equiv 0 \pmod 1$, implying that $f(k) \equiv 0 \pmod 1$. This proves the Lemma.

Now we prove the theorem. Let $N_j = N$ be temporarily fixed. For an integer k let

$$I_k = \left[\frac{N}{k}, \frac{N}{k} + (2+\varepsilon)\frac{\sqrt{N}}{k}\right].$$

If the intervals I_k, I_{k+1} contain a common integer element m, then $f(k) \equiv f(k+1) \pmod 1$. Indeed, $mk, m(k+1) \in J_N$, and (2.1) holds.

There is a common element m, if

$$\frac{N+(2+\varepsilon)\sqrt{N}}{k+1} - \frac{N}{k} \geqq 0,$$

i.e., if $k^2 - ((2+\varepsilon)\sqrt{N}-1)k + N \leqq 0$. This inequality holds in the interval $k \in [k_1, k_2]$, where

$$k_1 = \frac{1}{2}\{(2+\varepsilon)\sqrt{N}-1\} - \frac{1}{2}\sqrt{\{(2+\varepsilon)\sqrt{N}-1\}^2 - 4N}, \tag{2.3}$$

$$k_2 = \frac{1}{2}\{(2+\varepsilon)\sqrt{N}-1\} + \frac{1}{2}\sqrt{\{(2+\varepsilon)\sqrt{N}-1\}^2 - 4N}, \tag{2.4}$$

and so

$$f(k) \equiv \gamma_j \pmod 1, \quad k \in [k_1, k_2].$$

It is obvious that $k_1 = k_1(N)$ as $N = N_j \to \infty$. Furthermore, in view of (2.3) and (2.4), $\frac{k_2}{k_1} \geqq 1 + \varepsilon_1$, $\varepsilon_1 > 0$ for every large N. Now we may apply the Lemma with

$k_1=N$, $v=(1+\varepsilon_1)$, $v_1=1+\frac{\varepsilon_1}{2}$. Putting $N=N_j$ we deduce that $f(n)\equiv 0 \pmod 1$ for $n<\frac{\varepsilon_1}{2}N_j$, i.e., for every n.

This completes the proof.

References

[1] J. L. MAUCLAIRE and LEO MURATA, On the regularity of arithmetic multiplicative functions. I, *Proc. Japan Acad. Ser. A*, **56** (1980), 438—440.

[2] I. KÁTAI, Some problems in number theory, *Studia Sci. Math. Hungar.*, to appear.

[3] I. KÁTAI, On the determination of an additive arithmetical function by its local behaviour, *Colloq. Math.*, **20** (2) (1969), 265—267.

EÖTVÖS LORÁND UNIVERSITY
DEPARTMENT OF NUMERICAL ANALYSIS AND COMPUTER SCIENCE
MÚZEUM KRT. 6—8
1088 BUDAPEST, HUNGARY

Acta Sci. Math. 44 (1982), 307—311

Beitrag zu den Arbeiten "Bemerkung zu einem Satz von S. Kaczmarz" und "Über einen Satz von Alexits und Sharma"

KÁROLY TANDORI

1. Für ein System $\varphi=\{\varphi_k(x)\}_1^\infty$ der Funktionen $\varphi_k(x)\in L(0, 1)$ sei

$$L_n^*(\varphi; x)=\int_0^1 \max_{1\le i\le n}\left|\sum_{k=1}^{i}\varphi_k(x)\varphi_k(t)\right|dt \quad (x\in(0, 1);\ n=1, 2, \ldots).$$

In der Note [1] haben wir den folgenden Satz bewiesen.

Satz A. *Ist* $a=\{a_k\}_1^\infty\notin l^2$, *dann gibt es ein orthonormiertes System* φ *in* (0,1) *mit* $L_n^*(\varphi; x)=O(1)$ $(x\in(0, 1);\ n=1, 2, \ldots)$ *derart, daß die Reihe*

(1) $$\sum_{k=1}^{\infty} a_k\varphi_k(x)$$

in (0, 1) *überall divergiert.*

Mit der in [1] angewandten Methode kann man die folgende, schärfere Behauptung beweisen.

Satz I. *Ist* $a\notin l^2$, *dann gibt es ein orthonormiertes System* φ *in* (0, 1) *mit*

(2) $$\int_0^1 \sum_{k=1}^{\infty}|\varphi_k(x)\varphi_k(t)|\,dt\le K<\infty \quad (x\in(0, 1))$$

derart, daß die Reihe (1) *in* (0, 1) *überall divergiert.*

Beweis des Satzes I. Es sei $0=n(1)<\ldots<n(l)<\ldots$ eine Indexfolge mit der Eigenschaft

(3) $$A_l^2=\sum_{k=n(l-1)+1}^{n(l)} a_k^2\ge 4^l \quad (l=2, 3, \ldots).$$

Mit $k_1<\ldots<k_i<\ldots$ bezeichnen wir die Indizes k, für die $a_k=0$ ist. Es sei $Z(l)$ die Menge der Indizes k mit $n(l-1)<k\le n(l)$ und $a_k\ne 0$ $(l=2, 3, \ldots)$.

Eingegangen am 16. Oktober 1981.

Es seien weiterhin $I_k(l), J_k(l)$ $(k\in Z(l);\ l=2, 3, \ldots)$, J_i $(i=1, 2, \ldots)$ Teilintervalle von $(0, 1)$ mit den Eigenschaften (für $l, l_1, l_2=2, 3, \ldots$)

$$I_{k_1}(l)\cap I_{k_2}(l)=\emptyset \quad (k_1, k_2\in Z(l),\ k_1\neq k_2), \quad \bigcup_{k\in Z(l)} I_k(l)=(0, 1),$$

$$\operatorname{mes} I_k(l)=a_k^2/A_l^2 \quad (k\in Z(l)), \quad I_k(l)\cap J_k(l)=\emptyset \quad (k\in Z(l)),$$

$$J_{k_1}(l_1)\cap J_{k_2}(l_2)=\emptyset \quad (k_1\in Z(l_1),\ k_2\in Z(l_2),\ (k_1-k_2)^2+(l_1-l_2)^2\neq 0),$$

$$\operatorname{mes} J_k(l)=\operatorname{mes} I_k(l)/l^2 \quad (k\in Z(l)), \quad J_{i_1}\cap J_{i_2}=\emptyset \quad (i_1, i_2=1, 2, \ldots, i_1\neq i_2).$$

Unter den obigen Bedingungen kann man solche Intervalle leicht angeben.

Es sei $\varphi=\{\varphi_k(x)\}_1^\infty$ ein orthonormiertes System von Treppenfunktionen in $(0, 1)$ mit den Eigenschaften

$$|\varphi_k(x)|=\begin{cases} A_l/|a_k|l, & x\in I_k(l), \\ (1-1/l^2)^{1/2}/\sqrt{\operatorname{mes} J_k(l)}, & x\in J_k(l), \\ 0 & \text{sonst} \end{cases} \quad (k\in Z(l)),$$

$$|\varphi_{k_i}(x)|=\begin{cases} 1/\sqrt{\operatorname{mes} J_i}, & x\in J_i, \\ 0 & \text{sonst} \end{cases} \quad (i=1, 2, \ldots).$$

Ein solches System kann leicht angegeben werden; man hat die Gruppe der Funktionen $\varphi_k(x)$ $(k\in Z(l))$, $\varphi_{k_{l-1}}(x)$ durch Rekursion zu definieren.

Es sei $x\in(0, 1)$. Auf Grund der Definition der Intervalle $J_k(l), J_i$ und der Funktionen $\varphi_k(x)$ gibt es einen Index l_0 derart, daß

$$x\notin\Big(\bigcup_{l=l_0}^{\infty}\bigcup_{k\in Z(l)} J_k(l)\Big)\cup\Big(\bigcup_{\substack{i\\ k_i>n(l_0-1)}} J_i\Big).$$

Ist $l\geqq l_0$, dann gibt es auf Grund von (3) und der Definition von $\varphi_k(x)$ einen Index $k(x, l)\in Z(l)$ mit

$$\Big|\sum_{k=n(l-1)+1}^{n(l)} a_k\varphi_k(x)\Big|=|a_{k(x,l)}\varphi_{k(x,l)}(x)|=A_l/l\geqq 2^l/l.$$

Daraus folgt, daß die Reihe (1) im Punkt x divergiert.

Es sei $x\in(0, 1)$. Auf Grund der Definition der Intervalle $I_k(l)$, $J_k(l)$, J_i und der Funktionen $\varphi_k(x)$ gibt es für jedes l einen Index $k(x, l)\in Z(l)$ mit $x\in I_{k(x,l)}(l)$; weiterhin eventuell existiert einen Index $k_0(x, l_0)\in Z(l_0)$ mit $x\in J_{k_0(x,l_0)}(l_0)$, bzw. einen Index i_0 mit $x\in J_{i_0}$. Dann gilt

$$\sum_{k=1}^{\infty}|\varphi_k(x)\varphi_k(t)|=\sum_{l=2}^{\infty}|\varphi_{k(x,l)}(x)\varphi_{k(x,l)}(t)|+$$

$$+|\varphi_{k_0(x,l_0)}(x)\varphi_{k_0(x,l_0)}(t)|+|\varphi_{k_{i_0}}(x)\varphi_{k_{i_0}}(t)|.$$

(Wenn solcher Index $k_0(x, l_0)$, bzw. i_0 nicht existiert, dann ist der zweite, bzw. dritte Glied an der rechten Seite gleich mit Null.) Auf Grund der Definition der Funktion $\varphi_k(x)$ ergibt sich dann

$$\int_0^1 \sum_{k=1}^{\infty} |\varphi_k(x)\varphi_k(t)|\,dt =$$

$$= \sum_{l=2}^{\infty}\Big(|\varphi_{k(x,l)}(x)| \int_{I_{k(x,l)}(l)} |\varphi_{k(x,l)}(t)|\,dt + |\varphi_{k(x,l)}(x)| \int_{J_{k(x,l)}(l)} |\varphi_{k(x,l)}(t)|\,dt +$$

$$+ |\varphi_{k_0(x,l_0)}(x)| \int_{I_{k_0(x,l_0)}(l_0)} |\varphi_{k_0(x,l_0)}(t)|\,dt + |\varphi_{k_0(x,l_0)}(x)| \int_{J_{k_0(x,l_0)}(l_0)} |\varphi_{k_0(x,l_0)}(t)|\,dt +$$

$$+ |\varphi_{k_{i_0}}(x)| \int_{J_{i_0}} |\varphi_{k_{i_0}}(t)|\,dt =$$

$$= \sum_{l=2}^{\infty}\left(\frac{A_l^2}{a_{k(x,l)}^2 l^2} \operatorname{mes} I_{k(x,l)}(l) + \frac{A_l}{|a_{k(x,l)}|\,l}\left(1-\frac{1}{l^2}\right)^{\frac12} \frac{1}{\sqrt{\operatorname{mes} J_{k(x,l)}(l)}} \operatorname{mes} J_{k(x,l)}(l)\right) +$$

$$+\left(1-\frac{1}{l_0^2}\right)^{\frac12} \frac{1}{\sqrt{\operatorname{mes} J_{k_0(x,l_0)}(l_0)}}\, \frac{A_{l_0}}{|a_{k_0(x,l_0)}|\,l_0} \operatorname{mes} I_{k_0(x,l_0)}(l_0) +$$

$$+\left(1-\frac{1}{l_0^2}\right) \frac{1}{\operatorname{mes} J_{k_0(x,l_0)}(l_0)} \operatorname{mes} J_{k_0(x,l_0)}(l_0) + \frac{1}{\operatorname{mes} J_{i_0}} \operatorname{mes} J_{i_0} =$$

$$= \sum_{l=2}^{\infty}\left(\frac{1}{l^2} + \left(1-\frac{1}{l^2}\right)^{\frac12} \frac{1}{l^2}\right) + \left(1-\frac{1}{l_0^2}\right)^{\frac12} + \left(1-\frac{1}{l_0^2}\right) + 1 \leqq 3\left(1+\sum_{l=2}^{\infty}\frac{1}{l^2}\right) = K < \infty.$$

Damit haben wir bewiesen, daß (2) für das System φ erfüllt ist.

2. Es sei $\lambda = \{\lambda_k\}_1^\infty$ eine monoton nichtabnehmende Folge von positiven Zahlen mit $\lambda_k \to \infty$ $(k \to \infty)$. Ohne Beschränkung der Allgemeinheit können wir $\lambda_1 \geqq 1$ voraussetzen. Für jede positive ganze Zahl l bezeichnet $Z(l)$ die Menge der positiven ganzen Zahlen k, mit $2^l < \lambda_k \leqq 2^{l+1}$. Es seien $l_1 < \ldots < l_i < \ldots$ diejenigen Indizes, für die $Z(l_i) \neq \emptyset$ ist; die Elemente von $Z(l)$ seien in der natürlichen Anordnung $\nu(i)+1, \ldots, \nu(i+1)$. Für eine reelle Zahlenfolge $a = \{a_k\}_1^\infty$ setzen wir

$$A_i^2 = \sum_{k=\nu(i)+1}^{\nu(i+1)} a_k^2 \lambda_k \quad (i = 1, 2, \ldots).$$

In der Arbeit [2] haben wir den folgenden Satz bewiesen.

Satz B. *Gilt*

(4) $$\sum_{i=1}^{\infty} A_i = \infty,$$

dann gibt es ein System $\varphi=\{\varphi_k(x)\}_1^\infty$ *von reellen Funktionen in* $L(0,1)$ *derart, daß*

$$L_n^*(\varphi;x)\leqq 16\lambda_n \quad (x\in(0,1);\ n=1,2,\ldots) \tag{5}$$

besteht, und die Reihe (1) *in* $(0,1)$ *überall divergiert.*

Mit der in [2] angewandten Methode kann man die folgende schärfere Behauptung zeigen.

Satz II. *Gilt* (4), *dann gibt es ein System* $\varphi=\{\varphi_k(x)\}_1^\infty$ *von reellen Funktionen in* $L(0,1)$ *derart, daß*

$$\int_0^1 \sum_{k=1}^n |\varphi_k(x)\varphi_k(t)|\,dt\leqq 16\lambda_n \quad (x\in(0,1);\ n=1,2,\ldots)$$

besteht, und die Reihe (1) *in* $(0,1)$ *überall divergiert.*

Beweis des Satzes II. Für jede positive ganze Zahl i seien $I_s(i)$ $(s=\nu(i)+1,\ldots$ $\ldots,\nu(i+1))$ disjunkte Intervalle mit

$$\bigcup_{s=\nu(i)+1}^{\nu(i+1)} I_s(i)=(0,1),\quad \operatorname{mes} I_s(i)=a_s^2\Big/\sum_{k=\nu(i)+1}^{\nu(i+1)} a_k^2,$$

wenn $a_s\neq 0$, und $I_s(i)=\emptyset$, wenn $a_s=0$. Für einen Index s mit $\nu(i)<s\leqq\nu(i+1)$ und $a_s\neq 0$ setzen wir

$$\varphi_s(x)=\begin{cases} A_i/a_s, & x\in I_s(i),\\ 0 & \text{sonst};\end{cases}$$

im Falle $a_s=0$ sei $\varphi_s(x)\equiv 0$.

Sei i_0 eine positive ganze Zahl und sei $x\in(0,1)$. Dann gibt es für jede positive ganze Zahl i $(1\leqq i\leqq i_0)$ einen Index $s(x,i)$ $(\nu(i)<s(x,i)\leqq\nu(i+1))$ mit $x\in I_{s(x,i)}(i)$. Man hat dann

$$\sum_{k=1}^{\nu(i_0+1)} a_k\varphi_k(x)=\sum_{i=1}^{i_0} a_{s(x,i)}\varphi_{s(x,i)}(x).$$

Daraus, auf Grund der Definition der Funktionen $\varphi_k(x)$ folgt

$$\sum_{k=1}^{\nu(i_0+1)} a_k\varphi_k(x)=\sum_{i=1}^{i_0} A_i \quad (x\in(0,1);\ i_0=1,2,\ldots).$$

Daraus und aus (4) ergibt sich, daß die Reihe (1) in $(0,1)$ überall divergiert.

Es sei i eine positive ganze Zahl, $\nu(i)<n\leqq\nu(i+1)$ und $x\in(0,1)$. Dann gibt es einen Index $s(x,i)$ $(\nu(i)<s(x,i)\leqq\nu(i+1))$ mit $x\in I_{s(x,i)}(i)$, und so gilt

$$\int_0^1 \sum_{k=\nu(i)+1}^{n} |\varphi_k(x)\varphi_k(t)|\,dt\leqq \int_{I_{s(x,i)}(i)} |\varphi_{s(x,i)}(x)\varphi_{s(x,i)}(t)|\,dt=$$

$$=\frac{A_i^2}{a_{s(x,i)}^2}\operatorname{mes} I_{s(x,i)}(i).$$

Daraus folgt, auf Grund der Definition von A_i und $\nu(i)$

$$(6)\quad \int_0^1 \sum_{k=\nu(i)+1}^{n} |\varphi_k(x)\varphi_k(t)|\,dt \leqq 2\lambda_n \quad \big(x\in(0,1);\ \nu(i)<n\leqq \nu(i+1);\ i=1,2,\ldots\big).$$

Es sei n eine beliebige positive ganze Zahl. Dann gibt es einen Index i_0 mit $\nu(i_0)<$ $<n\leqq\nu(i_0+1)$, und gilt

$$\int_0^1 \sum_{k=1}^{n} |\varphi_k(x)\varphi_k(t)|\,dt \leqq \sum_{i=0}^{i_0-1}\int_0^1 \sum_{k=\nu(i)+1}^{\nu(i+1)} |\varphi_k(x)\varphi_k(t)|\,dt + \int_0^1 \sum_{k=\nu(i_0)+1}^{n} |\varphi_k(x)\varphi_k(t)|\,dt \leqq$$

$$\leqq 2(\lambda_{\nu(2)}+\ldots+\lambda_{\nu(i_0)}+\lambda_n) \leqq 4(2^2+\ldots+2^{i_0+1}) \leqq 16 2^{i_0} \leqq 16\lambda_{\nu(i)+1} \leqq 16\lambda_n,$$

für jedes $x\in(0,1)$, auf Grund von (6).

Damit haben wir bewiesen, daß (5) erfüllt ist.

Schriftenverzeichnis

[1] K. Tandori, Bemerkung zu einem Satz von S. Kaczmarz, *Acta Sci. Math.*, **42** (1980), 171—173.
[2] K. Tandori, Über einen Satz von Alexits und Sharma, *Acta Sci. Math.*, **42** (1980), 175—182.

BOLYAI INSTITUT
UNIVERSITÄT SZEGED
ARADI VÉRTANÚK TERE 1
6720 SZEGED, UNGARN

Acta Sci. Math. 44 (1982), 313—320

The stability of d'Alembert-type functional equations

L. SZÉKELYHIDI

In this paper we deal with the following problem: if f, g, h, k are complex valued functions on the Abelian group G with the property, that the function $(x, y) \to f(x+y)+g(x-y)-h(x)k(y)$ is bounded, what can be said about the functions f, g, h, k? Obviously, this problem is a generalization of the well-known functional equations

$$f(x+y)+f(x-y) = 2f(x)g(y), \tag{0}$$

$$f(x+y)+g(x-y) = h(x)k(y). \tag{1}$$

Special cases of this problem has been treated by many authors. The special case $k=1$ is of "additive type" and can be reduced to the problem: if $(x, y) \to f(x+y)-f(x)-f(y)$ is bounded, what can be said about f? The problem in this form is treated in [2], [4], [5], [6], [8]. The special case $g=0$ and $h=k=f$ is treated in [3], and the case $g=0$ and $h=f$ is treated in [9]. Further, the special case where $f=g=h$ and $k=2f$ is treated in [3], and the case where $f=g=h$ is treated in [10]. In this paper we completely solve the above problem.

First we make a simple observation: evidently, if f, g, h, k is a solution of the functional equation (1) and a, b are arbitrary bounded complex valued functions on G, then the functions $f+a, g+b, h, k$ solve our problem. Our main result is the following: if f, g, h, k are unbounded functions, then essentially this is the only solution of our problem.

In the sequel we shall use the following notation and terminology: C denotes the set of complex numbers. If G is a group and $M: G \to C$ is a function for which $M(x+y)=M(x)M(y)$ holds for all x, y in G, then we call M an *exponential.* The function $A: G \to C$ is called *additive*, if $A(x+y)=A(x)+A(y)$ holds whenever x, y is in G. If $F: G \to C$ is a function, then F_e and F_o denotes the even and the odd

Received July 8, 1981, and in revised form November 28, 1981.

part of F respectively, that is,

$$F_e(x) = -\frac{1}{2}(F(x)+F(-x)), \quad F_o(x) = \frac{1}{2}(F(x)-F(-x))$$

for all x in G.

In what follows we suppose, that G is a fixed Abelian group in which the mapping $x \to 2x$ is an automorphism.

We shall use the following theorem:

Theorem 1. *If $f, g\colon G \to C$ satisfy* (0), *then there are an exponential $M\colon G \to C$, an additive function $A\colon G \to C$ and α, β constants such that we have the following possibilities:*

(i) $f = 0$, *g is arbitrary,*
(ii) $f = A+\alpha$, $g = 1$,
(iii) $f = \alpha M_e + \beta M_o$, $g = M_e$.

The proof of this theorem can be obtained by the method of [1], using the results of [7].

Lemma 2. *Let $f, g, h\colon G \to C$ be functions for which the function $(x, y) \to f(x+y)-g(x)h(y)$ is bounded. Then there are an exponential $M\colon G \to C$, a bounded function $a\colon G \to C$ and α, β constants such that we have the following possibilities:*

(i) *f is bounded, h is arbitrary,* $g=0$,
(ii) *f is bounded,* $h=0$, *g is arbitrary,*
(iii) *f, g, h are bounded,*
(iv) $f = \alpha\beta M + a$, $g = \alpha M$, $h = \beta M$.

Proof. The first three cases are trivial, hence we may suppose that f, g, h are unbounded. Let $\alpha = g(0)$, $\beta = h(0)$ and $a = f - \beta g$. Obviously, a is bounded, and the identity

$$f(x+y)-g(x)h(y)-a(x+y) = \beta g(x+y)-g(x)h(y)$$

implies that $\beta \neq 0$, and the function $(x, y) \to g(x+y)-g(x)\beta^{-1}h(y)$ is bounded. By [9], it follows (iv).

Lemma 3. *Let $f, g\colon G \to C$ be functions for which the function $(x, y) \to f(x+y) + f(x-y)-2f(x)g(y)$ is bounded. Then there are an exponential $M\colon G \to C$, an additive function $A\colon G \to C$, a bounded function $a\colon G \to C$ and α, β constants such that we have the following possibilities:*

(i) $f = 0$, *g is arbitrary,*
(ii) *f, g are bounded,*
(iii) $f = A + a$, $g = 1$,
(iv) $f = \alpha M_e + \alpha M_o$, $g = M_e$.

Proof. The first two cases are trivial. We may suppose that f is unbounded. This implies that $g \neq 0$. If $g=1$, then by [8], (iii) follows. Suppose that $g \neq 1$. Let $F(x, y)=f(x+y)+f(x-y)-2f(x)g(y)$ for all x, y in G. By [10] and Theorem 1, there is an exponential $M\colon G \to C$ for which $g=M_e$, in particular g is even. Now consider the identity

$$2g(z)F(x, y) = F(x, y+z)+F(x, y-z)-F(x+y, z)-F(x-y, z),$$

which shows that either g is bounded, or $F=0$. Suppose, that g is bounded, and observe that the following identities hold:

$$(2)\quad f_e(y)g(x)-f_e(x)g(y) = \frac{1}{4}\left(F(x, y)-F(y, x)+F(-x, -y)-F(-y, -x)\right),$$

$$(3)\quad f_o(x+y)-f_o(x)g(y)-f_o(y)g(x) = \frac{1}{4}\left(F(x, -y)-F(-y, x)-F(-x, y)+F(y, -x)\right).$$

By (2) we obtain that f_e is bounded, and by (3) we see that the function $x \to f_o(x+y)-f_o(x)g(y)$ is bounded for all fixed y in G. Since f_o cannot be bounded, by [9] it follows that g is an exponential. As $g \neq 0$, we have $g(0)=1$, and for all x in G,

$$1 = g(0) = g\left(\frac{x}{2}\right)g\left(-\frac{x}{2}\right) = g\left(\frac{x}{2}\right)g\left(\frac{x}{2}\right) = g(x),$$

a contradiction. Hence g is unbounded and $F=0$, that is, (iv) follows by Theorem 1.

Theorem 4. *Let $f, g, h, k\colon G \to C$ be functions for which the function $(x, y) \to f(x+y)+g(x-y)-h(x)k(y)$ is bounded. Then there are an exponential $M\colon G \to C$, an additive function $A\colon G \to C$, bounded functions $a, b, c\colon G \to C$, and constants $\alpha, \beta, \gamma, \delta$ such that we have the following possibilities:*

(i) *f, g, h, k are bounded,*

(ii) *f, g are bounded, $h=0$, k is arbitrary,*

(iii) *f, g are bounded, h is arbitrary, $k=0$,*

(iv) *f is bounded, $g = \alpha\beta M+b$, $h = \alpha M$, $k = \beta M^{-1}$,*

(v) *$f = \alpha\beta M+a$, g is bounded, $h = \alpha M$, $k = \beta M$,*

(vi) $f = \frac{1}{2}\alpha A+a,\quad g = -\frac{1}{2}\alpha A+b,\quad h = \alpha,\quad k = A+c,$

(vii) $f = \frac{1}{2}\beta A+a,\quad g = \frac{1}{2}\beta A+b,\quad h = A+c,\quad k = \beta,$

(viii) $f = \frac{1}{4}\alpha\beta A^2+\frac{1}{2}(\alpha\delta+\beta\gamma)A+a,\quad g = -\frac{1}{4}\alpha\beta A^2+\frac{1}{2}(\alpha\delta-\beta\gamma)A+b,$
$h = \alpha A+\gamma,\quad k = \beta A+\delta,$

$$\text{(ix)}\ f=\frac{1}{2}(\alpha\gamma+\beta\delta)M_e+\frac{1}{2}(\alpha\delta+\beta\gamma)M_o+a,\quad h=\alpha M_e+\beta M_o,$$

$$g=\frac{1}{2}(\alpha\gamma-\beta\delta)M_e-\frac{1}{2}(\alpha\delta-\beta\gamma)M_o+b,\quad k=\gamma M_e+\delta M_o.$$

Proof. The first three cases are trivial, and if f or g is bounded, then by Lemma 2 we have (iv) or (v). Now we may suppose that f, g are unbounded, and $h\neq 0$, $k\neq 0$. Let $h(x_0)\neq 0$, $k(y_0)\neq 0$, and we introduce the new functions:

$$F(x)=h(x_0)^{-1}k(y_0)^{-1}f(x+x_0+y_0),\quad G(x)=h(x_0)^{-1}k(y_0)^{-1}g(x+x_0-y_0),$$

$$H(x)=h(x_0)^{-1}h(x+x_0),\quad K(x)=k(y_0)^{-1}k(x+y_0).$$

We have that F, G are unbounded, $H(0)=K(0)=1$, and the function D defined by

$$D(x,y)=F(x+y)+G(x-y)-H(x)K(y) \tag{4}$$

is bounded. First we present some simple identities concerning F, G, H, K, D, which we shall need in the sequel:

$$H(x+y)+H(x-y)-2H(x)K_e(y)= \tag{5}$$

$$=D(x,y)+D(x,-y)-D(x+y,0)-D(x-y,0),$$

$$H_o(y)K_o(x)-H_o(x)K_o(y)=\frac{1}{4}(D(x,y)-D(y,x)-D(x,-y)+ \tag{6}$$

$$+D(-y,x)+D(-x,-y)-D(-y,-x)-D(-x,y)+D(y,-x),$$

$$H(x+y)K_o(x-y)-H(x)K_o(x)+H(y)K_o(y)= \tag{7}$$

$$=\frac{1}{2}(D(x,x)-D(x,-x)+D(y,-y)-D(y,y)+D(x+y,y-x)-D(x+y,x-y)),$$

$$H_o(x+y)K_o(x-y)-H_o(x)K_o(x)+H_o(y)K_o(y)= \tag{8}$$

$$=\frac{1}{4}(D(x,x)+D(-x,-x)-D(x,-x)-D(-x,x)+D(y,-y)+D(-y,y)-$$

$$-D(y,y)-D(-y,-y)+D(x+y,y-x)+D(-x-y,x-y)-D(x+y,x-y)-$$

$$-D(-x-y,y-x)),$$

and finally, if $H_o=0$, that is, H is even, then

$$K(x+y)+K(x-y)-2K(x)H(y)=2D(y,x)-D(0,x+y)-D(0,x-y)- \tag{9}$$

$$-D\left(\frac{x+y}{2},\frac{x+y}{2}\right)+D\left(\frac{-x-y}{2},\frac{x+y}{2}\right)-D\left(\frac{y-x}{2},\frac{x-y}{2}\right)+D\left(\frac{x-y}{2},\frac{x-y}{2}\right).$$

These identities can be checked by an easy computation and they show, that the expressions on the left hand sides are bounded. Finally, we shall need the relations

$$F(x) = H\left(\frac{x}{2}\right)K\left(\frac{x}{2}\right)+D\left(\frac{x}{2}, \frac{x}{2}\right)-G(0), \tag{10}$$

$$G(x) = H\left(\frac{x}{2}\right)K\left(-\frac{x}{2}\right)+D\left(\frac{x}{2}, -\frac{x}{2}\right)-G(0).$$

Now we assume that H is bounded, and show that (vi) follows. By (5) K_e is bounded, and if H is not even, then by (6) K_o is bounded, too, which is impossible by (10). Hence H is even, and then by (9) and Lemma 3 either $K=A+a$ and $H=1$, or $K=M_e+\beta M_o$, $H=M_e$. In the latter case M_e is bounded, and by the identity $M_e(x+y)-M_e(x-y)=2M_o(x)M_o(y)$ the function M_o is bounded, too, that is, K is also bounded, which is impossible by (10). This means that $H=1$ and $K=A+a$, where $A\colon G\to C$ is additive, and $a\colon G\to C$ is bounded. By (10) and by the definition of F, G, H, K we have (vi).

Hence we may suppose in the sequel, that H is unbounded.

From (5) by Lemma 3 we have two cases. In the first case $K_e=1$, $H=A+c$, where $A\colon G\to C$ is additive and $c\colon G\to C$ is bounded. Here $A\neq 0$ and $H_o\neq 0$, hence by (6) $K_o=\alpha A+d$, where $d\colon G\to C$ is odd and bounded, and α is a constant. If $\alpha=0$, then by (6) either H_o is bounded, which is impossible, or $K_o=0$, that is $K=K_e=1$ and from (10) we obtain (vii) using the definition of F, G, H, K.

Let $\alpha\neq 0$, then we substitute H_o and K_o into (6) and we have that the function

$$(x, y) \to A(x)\left(\alpha c_o(y)-d(y)\right)-A(y)\left(d(x)-\alpha c_o(x)\right)$$

is bounded. If there is a y in G, for which $d(y)\neq\alpha c_o(y)$, then $A=0$, which is impossible. Hence $d=\alpha c_o$, and $H=A+c_o+c_e$, $K=\alpha A+\alpha c_o+1$. Substituting into (8) we have that the function

$$(x, y) \to A(x)\left(c_o(x+y)+c_o(x-y)-2c_o(x)\right)-A(y)\left(c_o(x+y)-c_o(x-y)-2c_o(y)\right)$$

is bounded. Substituting $x+y$ for x and $x-y$ for y, we have that the function

$$(x, y) \to A(x+y)c_o(x+y)-A(x-y)c_o(x-y)-A(y)c_o(2x)-A(x)c_o(2y) \tag{11}$$

is bounded. Let $p(x)=A(x)c_o(x)$ and $P(x, y)=p(x+y)-p(x-y)-A(x)c_o(2y)$, then (11) implies the boundedness of $x\to P(x, y)$ for all fixed y in G. On the other hand, the identity

$$P(x+y, z)+P(x-y, z)-P(x, y+z)+P(x, y-z) =$$

$$= A(x)\left(c_o(2y+2z)-c_o(2y-2z)-2c_o(2z)\right)$$

shows, that for all fixed y, z in G the function $x \to A(x)\big(c_o(2y+2z)-c_o(2y-2z)-$ $-2c_o(2z)\big)$ is bounded, and hence

$$c_o(2y+2z)-c_o(2y-2z) = 2c_o(2z)$$

holds for all y, z in G. Interchanging y and z, we have that c_o is additive and as it is bounded, $c_o=0$, $H=A+c_e$, $K=\alpha A+1$. Substituting into (7) we get that the function

$$(x, y) \to A(x)\big(c_e(x+y)-c_e(x)\big)-A(y)\big(c_e(x+y)-c_e(y)\big)$$

is bounded. Writing $x+y$ for x and $x-y$ for y we obtain that the function

$$(12) \qquad (x, y) \to A(x+y)c_e(x+y)-A(x-y)c_e(x-y)-2A(y)c_e(2x)$$

is bounded. Let $p(x)=A(x)c_e(x)$ and $P(x, y)=p(x+y)-p(x-y)-2A(y)c_e(2x)$, then (12) implies that P is bounded. On the other hand, the identity

$$P(x+y, z)+P(x-y, z)-P(r, y+z)+P(x, y-z) =$$
$$= -2A(z)\big(c_e(2x+2y)+c_e(2x-2y)-2c_e(2x)\big)$$

shows that the functional equation

$$c_e(2x+2y)+c_e(2x-2y) = 2c_e(2x)$$

holds. Interchanging x and y we get that c_e is constant. Since $H(0)=1$, therefore $c_e=1$ and $H=A+1$, $K=\alpha A+1$. Using (10) and the definition of F, G, H, K we obtain case (viii).

Finally, we have to return to the second case at (5), where by Lemma 3, $H=M_e+\alpha M_o$, $K_e=M_e$. Here $M\colon G\to C$ is an exponential, and α is a constant. Of course $M_o=0$ is impossible, and so (6) implies $K_o=\beta M_o+a$, where $a\colon G\to C$ is bounded and β is a constant. Hence by (10) we have for all x in G that

$$F(x) = \frac{1+\alpha\beta}{2} M_e(x)+\frac{\alpha+\beta}{2} M_o(x)+\left[M_e\left(\frac{x}{2}\right)+\alpha M_o\left(\frac{x}{2}\right)\right] a\left(\frac{x}{2}\right)+d(x),$$

$$G(x) = \frac{1-\alpha\beta}{2} M_e(x)-\frac{\alpha-\beta}{2} M_o(x)-\left[M_e\left(\frac{x}{2}\right)+\alpha M_o\left(\frac{x}{2}\right)\right] a\left(\frac{x}{2}\right)+e(x),$$

where $d, e\colon G\to C$ are bounded functions (we have used that a is obviously odd). Substituting into (4) and using that D is bounded, we have that the function

$$(13) \qquad (x, y) \to H\left(\frac{x+y}{2}\right) a\left(\frac{x+y}{2}\right)-H\left(\frac{x-y}{2}\right) a\left(\frac{x-y}{2}\right)-H(x)a(y)$$

is bounded. Let $p(x)=H\left(\frac{x}{2}\right)a\left(\frac{x}{2}\right)$ and $P(x, y)=p(x+y)-p(x-y)-H(x)a(y)$.

Then (13) implies that P is bounded. On the other hand, using that H is unbounded, we infer from the identity

$$P(x+y,z)+P(x-y,z)+P(x,y-z)-P(x,y+z) =$$
$$= H(x)\big(a(y+z)-a(y-z)-2M_e(y)a(z)\big)$$

that the functional equation

$$a(y+z)-a(y-z) = 2M_e(y)a(z)$$

holds. If $a\neq 0$, then M_e, and consequently H is bounded, which is impossible. Hence $a=0$, and we obtain case (ix). The theorem is proved.

Remark. Theorem 4 shows that for unbounded functions $f, g, h, k\colon G\to C$ the only possibility for $(x,y)\to f(x+y)+g(x-y)-h(x)k(y)$ to be bounded is that $f+a, g+b, h, k$ be a solution of (1) with some bounded functions $a, b\colon G\to C$.

Remark. The proofs of the above theorems and lemmata show that the main result can be generalized for other functional analytic function properties instead of "boundedness". More precisely, let W be a complex linear space of complex valued functions on $G\times G$ with the properties:

(i) if F belongs to W, then $(x,y)\to F(x+u, y+v)$ belongs to W,

(ii) constant functions belong to W,

(iii) if F belongs to W, then all the functions

$$(x,y)\to F(y,x),\ (x,y)\to F(x,-y),$$

$$(x,y)\to F(x+y,x-y),\ (x,y)\to F(x+y,0),\ (x,y)\to F(x-y,0)$$

$$(x,y)\to F\left(\frac{x+y}{2},\frac{x+y}{2}\right),\ (x,y)\to F\left(\frac{x+y}{2},-\frac{x+y}{2}\right),$$

$$(x,y)\to F(x,x),\ (x,y)\to F(2x,0),$$

and for all z in G, $(x,y)\to F(x,z)$ belong to W,

(iv) if for a function $f\colon G\to C$ the function $(x,y)\to f(x+y)+f(x-y)-2f(x)$ belongs to W, then there is a function $A\colon G\to C$ such that $A(x+y)+A(x-y)=$ $=2A(x)$ holds for all x, y in G, and $(x,y)\to f(x)-A(x)$ belongs to W.

Then Theorem 4 holds, if we set everywhere "belongs to W" instead of "bounded". For instance, if $W=(0)$, then we obtain from Theorem 4 the general solution of (1). As less trivial examples, "boundedness" can be replaced by "almost periodicity", or in the cases $G=R$ (the real line) or G compact Abelian, by "continuity", provided the mapping $x\to 2x$ is a homeomorphism.

References

[1] J. Aczél, *Lectures on functional equations and their applications,* Academic Press (New York—London, 1966).

[2] M. Albert, J. A. Baker, *Functions with bounded n-th differences,* unpublished manuscript.

[3] J. A. Baker, The stability of the cosine equation, *Proc. Amer. Math. Soc.,* **80** (3) (1980), 411—416.

[4] N. G. de Bruijn, Functions whose differences belong to a given class, *Nieuw Arch. Wisk.,* **23** (1951), 194—218.

[5] D. H. Hyers, On the stability of the linear functional equation, *Proc. Nat. Acad. Sci.,* **27** (1941), 222—224.

[6] D. H. Hyers, Transformations with Bounded m-th Differences, *Pacific J. Math.,* **11** (1961), 591—602.

[7] Pl. Kannappan, The functional equation $f(xy)+f(xy)^{-1}=2f(x)f(y)$ for groups, *Proc. Amer. Math. Soc.,* **19** (1968), 69—74.

[8] L. Székelyhidi, The stability of linear functional equations, *C. R. Math. Rep. Acad. Sci. Canada,* **3** (1981), No. 2, 63—67.

[9] L. Székelyhidi, On a theorem of Baker, Lawrence and Zorzitto, *Proc. Amer. Math. Soc.,* **84** (1982), 95—96.

[10] L. Székelyhidi, On a stability theorem, *C. R. Math. Rep. Acad. Sci. Canada,* **3** (1981), No. 5, 245—255.

DEPARTMENT OF MATHEMATICS
KOSSUTH LAJOS UNIVERSITY
4010 DEBRECEN, HUNGARY

Acta Sci. Math. 44 (1982), 321—328

Existence and uniqueness of random solutions of nonlinear stochastic functional integral equations

JAN TURO

1. Introduction. Stochastic or random integral equations arise quite often in the engineering, biological, chemical, and physical sciences (see, e.g., [1], [8] and [6]).

The object of the present paper is to study a nonlinear stochastic functional integral equation of the type

$$x(t,\omega)=F\Big(t,\int_0^{g(t)} f_1(t,s,x(s,\omega),\omega)\,ds, \tag{1.1}$$

$$\int_0^{g(t)} f_2(t,s,x(s,\omega),\omega)\,dw(s,\omega),\ x(h(t),\omega),\omega\Big)\overset{\text{df}}{=}(Ux)(t,\omega),$$

where

(i) $t\in R_+\overset{\text{df}}{=}[0,+\infty)$, and $\omega\in\Omega$, the supporting set of a complete probability measure space $(\Omega,\mathscr{F},P)$;

(ii) $x\colon R_+\times\Omega\to R$ is the unknown random function;

(iii) $F\colon R_+\times R^3\times\Omega\to R$ and $f_j\colon \Delta\times R\times\Omega\to R$, $j=1,2$, are given random functions, where $\Delta\overset{\text{df}}{=}\{(t,s)\colon 0\leqq s\leqq t<\infty\}$;

(iv) $g,h\colon R_+\to R_+$ are given scalar functions;

(v) $w\colon R_+\times\Omega\to R$ is a Wiener process.

The first integral of the stochastic equation (1.1) is to be understood as an ordinary Lebesgue integral, while the second integral is an Ito stochastic integral. We shall give sufficient conditions which will ensure the existence and uniqueness of a random solution, a second order stochastic process, of the above stochastic functional integral equation. The tool which we utilize to obtain these results is the comparison method. This method is based on the convergence of successive approximations produced by a comparison operator associated with the operator U. The abstract form of the comparison method was introduced by WAŻEWSKI [11] in the case of deterministic equations.

Received February 2, 1981, and in revised form September 2, 1981.

Almost all authors use the well-known Banach fixed point theorem or the concept of admissibility theory, [1], [8] and [6], in proving the existence and uniqueness of results for cases similar to equation (1.1). Unfortunately these methods involve a strong condition concerning the function F. By the comparison method this condition can be slightly weakened. Consequently in this paper conditions involving some relation between the Lipschitz constants of the function F and the estimations imposed on the functions g and h appear.

Equation (1.1) is a generalization of equations considered by MANOUGIAN, RAO and TSOKOS [6] (if $F(t, u_1, u_2, x, \omega)=h(t, x)+u_1+u_2$, $f_j(t, s, x, \omega)=k_j(t, s, \omega)\cdot\varphi_j(s, x)$ and $g(t)=t$, $h(t)=t$), TURO [10] (if $F(t, u_1, u_2, x, \omega)=F(t, u_1, x, \omega)$), GIHMAN and SKOROHOD [3], and DOOB [2], among others.

2. Preliminaries. We introduce a family $\mathscr{F}_t$, $t\in R_+$, of σ-algebras of subsets of Ω with the following properties:

(i) $\mathscr{F}_{t_1}\subset\mathscr{F}_{t_2}$, for $t_1<t_2$, $\mathscr{F}_t\subset\mathscr{F}$ $t\in R_+$;

(ii) for every t, $w(t, \omega)$ is $\mathscr{F}_t$-measurable;

(iii) for $\lambda\geqq 0$, the increments $w(t+\lambda, \omega)-w(t, \omega)$ are independent (in the probabilistic sense) of $\mathscr{F}_t$.

Definition 2.1. We shall denote by $C(R_+, L_2)$ the space of all continuous maps x: $R_+\to L_2(\Omega, \mathscr{F}_t, P)$ with the topology of uniform convergence on compacta.

It may be noted that $C(R_+, L_2)$ is a locally convex space whose topology is defined by the following family of seminorms:

$$\|x\|_n=\sup_{0\leqq t\leqq n}\{E[|x(t, \omega)|^2]\}^{1/2}$$

where E denotes the expected value of the random process.

Definition 2.2. A sequence $\{x_k\}$ of elements of the space $C(R_+, L_2)$ will be called a Cauchy sequence if for every $\varepsilon>0$ and n there exists an N such that for $k>N$ and $l>N$ we have $\|x_k-x_l\|_n<\varepsilon$.

It is clear that the space $C(R_+, L_2)$ is complete, that is, every Cauchy sequence of its elements has a limit in $C(R_+, L_2)$.

Definition 2.3. We shall call x a random solution of the stochastic functional integral equation (1.1) if $x\in C(R_+, L_2)$ and satisfies equation (1.1) P-a.e.

With respect to the functions appearing in equation (1.1) we shall assume the following:

(i) $F(t, u_1, u_2, x, \cdot)$ is $\mathscr{F}_t$-measurable for each $t\in R_+$, $u_1, u_2, x\in R$, and is continuous in t uniformly in u_1, u_2, x;

(ii) $f_j(t, s, x, \cdot)$, $j=1, 2$, are $\mathscr{F}_s$-measurable for each $(t, s)\in\Delta$, $x\in R$, and are continuous as maps from Δ into $L_2(\Omega, \mathscr{F}, P)$;

(iii) g, h: $R_+ \to R_+$ are continuous and $g(t)\leqq t$, $h(t)\leqq t$, $t\in R_+$.

3. Somme lemmas. Let us define

$$(Ku)(t) \overset{\text{df}}{=} k(t) \int_0^{g(t)} u(s)\,ds, \quad t\in R_+,$$

$$(Lu)(t) \overset{\text{df}}{=} l(t)\,u(h(t)), \quad t\in R_+.$$

Put $Su \overset{\text{df}}{=} \sum_{n=0}^{\infty} L^n u$ with the pointwise convergence of the series in R_+, where $L^n \overset{\text{df}}{=} LL^{n-1}$, $n=1, 2, \ldots$, $L^0=I$ is the identity operator in $C(R_+, R_+)$, the class of all continuous and nonnegative functions defined on R_+.

From the definition of the operator L it follows that

$$(L^n u)(t) = l_n(t)\,u(h_n(t)),$$

where

$$h_0(t) \overset{\text{df}}{=} t, \quad h_{n+1}(t) \overset{\text{df}}{=} h(h_n(t)), \quad n = 0, 1, \ldots, \quad t\in R_+,$$

$$l_0(t) \overset{\text{df}}{=} 1, \quad l_{n+1}(t) \overset{\text{df}}{=} \prod_{k=0}^{n} l(h_k(t)), \quad n = 0, 1, \ldots, \quad t\in R_+.$$

Lemma 3.1. ([9], [5]) *Assume that*

(i) $k, l, g, h, r\in C(R_+, R_+)$ *and* $g(t)$, $h(t)\in[0, t]$, $t\in R_+$;

(ii) $s=Sr<\infty$, $s^*=Sk^*<\infty$, *where* $k^*(t)\overset{\text{df}}{=}k(t)g(t)$;

(iii) $s, s^*\in C(R_+, R_+)$ *and* $\sup\limits_{R_+} \dfrac{s^*(t)}{t}<\infty$.

Then

(a) *there exists* $u_0\in C(R_+, R_+)$ *which is a unique solution of equation*

$$u = SKu+Sr \tag{3.1}$$

in the class L_{loc} *of all non-negative and locally integrable functions on* R_+;

(b) *the function* u_0 *is the unique solution of the equation*

$$u = Ku+Lu+r \tag{3.2}$$

in the class $L_{\text{loc}}(u_0)\overset{\text{df}}{=}\{u\colon u\in L_{\text{loc}}, \|u\|_0<\infty\}$, *where* $\|u\|_0\overset{\text{df}}{=}\inf\{c\colon u\leqq cu_0, c\in R_+\}$;

(c) *the function* $u=0$ *is the unique solution of the inequality*

$$u \leqq Ku+Lu \tag{3.3}$$

in the class $L_{\text{loc}}(u_0)$.

Proof. First we prove (a). We note that if $u \in L_{\mathrm{loc}}$ and is the solution of equation (3.1), then $u \in C(R_+, R_+)$. Thus we shall prove that equation (3.1) has a unique solution in $C(R_+, R_+)$. We shall obtain a solution first on an arbitrary closed, bounded interval $[0, n]$. Let $C([0, n], R)$ be the space of all continuous functions on $[0, n]$, where we introduce a norm $\|\cdot\|_*$ in the following way:

$$\|u\|_* \overset{\mathrm{df}}{=} \sup_{t\in[0,n]} e^{-\lambda t}|u(t)|, \quad \text{where} \quad \lambda > \bar{\lambda} \overset{\mathrm{df}}{=} \sup \frac{s^*(t)}{t}.$$

Now we can prove that the operator SK is a contraction in $C([0, n], R)$, i.e., $\|SK\| < 1$. Indeed, from the inequality $e^{\alpha t} - 1 \leqq \alpha e^t$ for $\alpha \in [0, 1]$, $t \in R_+$, we have

$$\|SKu\|_* \leqq \sup_{[0,n]} e^{-\lambda t} \sum_{n=0}^{\infty} l_n(t) k(h_n(t)) \int_0^{g(h_n(t))} e^{\lambda s} \sup_{s\in[0,n]} e^{-\lambda s}|u(s)|\, ds \leqq$$

$$\leqq \frac{1}{\lambda}\frac{s^*(t)}{t}\|u\|_* \leqq \frac{\bar{\lambda}}{\lambda}\|u\|_*.$$

Hence it follows that $\|SK\| < 1$. Now from the Banach fixed point theorem it follows that equation (3.1) has a unique solution $u_0 \in C([0, n], R_+)$. Since n is arbitrary, u_0 is a unique solution of equation (3.1) on R_+.

Now we prove (b). It is obvious that the function u_0 satisfies equation (3.2). Next we prove that in the class $L_{\mathrm{loc}}(u_0)$ the function u_0 is the unique solution of equation (3.2).

Indeed, if $u \in L_{\mathrm{loc}}(u_0)$ is a solution of (3.2) then by induction we get

$$\bar{u} = \sum_{i=0}^{n-1} L^i h + \sum_{i=0}^{n-1} L^i K\bar{u} + L^n \bar{u}.$$

Because $\bar{u} \in L_{\mathrm{loc}}(u_0)$, there exists $c \in R_+$ such that $\bar{u} \leqq cu_0$, hence $L^n\bar{u} \leqq cL^n u_0$. We easily find that $L^n u_0 \to 0$ since u_0 is the solution of equation (3.1). As a consequence of this $L^n\bar{u} \to 0$, and we infer that $\bar{u}$ satisfies (3.1). In view of the uniqueness proved for this equation we conclude $\bar{u} = u_0$.

Finally we prove (c). If $u \in L_{\mathrm{loc}}(u_0)$ is the solution of inequality (3.3) then we have $L^n u \to 0$, $n \to \infty$, and by induction we get

$$u \leqq \sum_{i=0}^{n-1} L^i K u + L^n u, \quad n = 0, 1, \ldots,$$

Letting $n\to\infty$ we get $u\leqq SKu$, hence we conclude that $u=0$, and the lemma is proved.

Remark 3.1. Now we give some effective conditions under which assumption (ii) of Lemma 3.1 is fulfilled.

a) If we assume that

$$(3.4)\quad k(t)\leqq \bar{k}=\text{const},\quad l(t)\leqq \bar{l}=\text{const},\quad g(t)\leqq \bar{g}t,\quad h(t)\leqq \bar{h}t,\quad \bar{g},\bar{h}\in[0,1],$$

and $r(t)\leqq\bar{r}t$, $t\in R_+$, for some $\bar{r}\in R_+$, then assumption (ii) of Lemma 3.1 is satisfied provided $\bar{l}\bar{h}<1$.

b) if $k(t)\leqq\bar{k}$, $l(t)\leqq\bar{l}t$, $g(t)\leqq\bar{g}t$, $h(t)\leqq\bar{h}t$, $r(t)\leqq\bar{r}t$, $\bar{k},\bar{l},\bar{r}\in R_+$, $\bar{g}\in[0,1]$ and $\bar{h}\in[0,1)$, $t\in R_+$, then assumption (ii) of Lemma 3.1 is satisfied.

c) Finally, if we suppose (3.4) and $r(t)\leqq\bar{r}t^p$, $t\in R_+$, for some $\bar{r},p\in R_+$, then (ii) of Lemma 3.1 is satisfied provided $\bar{l}\bar{h}^p<1$.

We construct a sequence as follows:

$$(3.5)\qquad u_{n+1}=Ku_n+Lu_n,\quad n=0,1,\ldots,$$

where u_0 is defined in Lemma 3.1.

Lemma 3.2. [4] *If the assumptions of Lemma* 3.1 *are satisfied, then*

$$(3.6)\qquad 0\leqq u_{n+1}\leqq u_n,\quad n=0,1,\ldots,$$

and $u_n\rightrightarrows 0$ *for* $n\to\infty$, *where the sign* $\rightrightarrows$ *denotes uniform convergence in any compact subset of* R_+.

Proof. Relation (3.6) we get by induction. The convergence of the sequence $\{u_n\}$ is implied by (3.6). The limit of this sequence satisfies the inequality (3.3), and by Lemma 3.1 it must be equal to zero identically. The uniform convergence of $\{u_n\}$ follows from Dini's theorem.

4. Main results. In order to prove the existence of a solution of equation (1.1), we define the sequence $\{x_n\}$ of random functions by the relations:

$$(4.1)\qquad x_{n+1}=Ux_n,\quad n=0,1,\ldots,$$

where U is defined by (1.1) and x_0 is an arbitrarily fixed element of $C(R_+,L_2)$.

We introduce the following

Assumption H. We assume that

(1) there exist functions $\bar{k}_j, \tilde{k}_j, l \in C(R_+, R_+)$, $j=1, 2$, such that

$$|F(t, u_1, u_2, x, \omega)-F(t, \bar{u}_1, \bar{u}_2, \bar{x}, \omega)| \leqq \sum_{j=1}^{2} \bar{k}_j(t)|u_j-\bar{u}_j|+l(t)|x-\bar{x}|,$$

$$|f_j(t, s, x, \omega)-f_j(t, s, \bar{x}, \omega)| \leqq \tilde{k}_j(t)|x-\bar{x}|,$$

for $t \in R_+$, $s \leqq t$, $u_j, \bar{u}_j, x, \bar{x} \in R$, $j = 1, 2$;

(2) $F(t, 0, 0, 0, \cdot) \in L_2(\Omega, \mathscr{F}_t, P)$ for each $t \in R_+$, and $f_j(t, s, 0, \cdot) \in L_2(\Omega, \mathscr{F}_s, P)$ for each $(t, s) \in \Delta$.

Remark 4.1. We note that from condition (1) of Assumption H we obtain the following estimates:

$$|F(t, u_1, u_2, x, \omega)|^2 \leqq 4\bar{k}_1^2(t)|u_1|^2+4\bar{k}_2^2(t)|u_2|^2+4l^2(t)|x|^2+4|F(t, 0, 0, 0, \omega)|^2$$

and

$$|f_j(t, s, x, \omega)|^2 \leqq 2\tilde{k}_j^2(t)|x|^2+2|f_j(t, s, 0, \omega)|^2$$

for $t \in R_+$, $s \leqq t$, $u_j, x \in R$, $j=1, 2$, $\omega \in \Omega$.

Put

$$\begin{aligned} k(t) &= 6[\bar{k}_1^2(t)\tilde{k}_1^2(t)g(t)+\bar{k}_2^2(t)\tilde{k}_2^2(t)], \\ l(t) &= 6l^2(t), \quad r(t) = 2E[|(Ux_0)(t, \omega)-x_0(t, \omega)|^2]. \end{aligned} \tag{4.2}$$

Theorem 4.1. *If Assumption* H *and assumptions* (ii) *and* (iii) *of Lemma* 3.1 *are satisfied with k, l and r defined by* (4.2), *then there exists a random solution $\bar{x} \in C(R_+, L_2)$ of equation* (1.1) *such that*

$$E[|\bar{x}(t, \omega)-x_n(t, \omega)|^2] \leqq u_n(t), \quad n = 0, 1, \ldots, \quad t \in R_+. \tag{4.3}$$

The solution $\bar{x}$ is unique in the class $L_{\text{loc}}^*(u_0) \overset{\text{df}}{=} \{x\colon x \in L_{\text{loc}}^*,\ E[|x(t, \omega)-x_0(t, \omega)|^2] \in \in L_{\text{loc}}(u_0)\}$, where L_{loc}^* is the class of all locally integrable random functions defined on R_+ with range in $L_2(\Omega, \mathscr{F}_t, P)$, and $L_{\text{loc}}(u_0)$ is defined in Lemma 3.1.

Proof. From the assumptions of the theorem, Cauchy's inequality, and the properties of the stochastic integral ([3], [7]) it follows that the integrals in equation (4.1) exist for each n (see Remark 4.1) and $x_n \in C(R_+, L_2)$, $n=0, 1, \ldots$.

To prove the existence of a solution of equation (1.1) we first prove the following estimates

$$E[|x_n(t, \omega)-x_0(t, \omega)|^2] \leqq u_0(t), \quad n = 0, 1, \ldots, \quad t \in R_+, \tag{4.4}$$

$$E[|x_{n+m}(t, \omega)-x_n(t, \omega)|^2] \leqq u_n(t), \quad n, m = 0, 1, \ldots, \quad t \in R_+. \tag{4.5}$$

It is clear that (4.4) holds for $n=0$. If we suppose that (4.4) holds for some $n>0$, then from $(x+y)^2 \leqq 2(x^2+y^2)$ and $(x+y+z)^2 \leqq 3(x^2+y^2+z^2)$, an application of Cauchy's inequality and the properties of the stochastic integral we have

$$E[|x_{n+1}(t,\omega)-x_0(t,\omega)|^2] \leqq$$

$$\leqq 2E[|(Ux_n)(t,\omega)-(Ux_0)(t,\omega)|^2]+2E[|(Ux_0)(t,\omega)-x_0(t,\omega)|^2] \leqq$$

$$\leqq 6\bar{k}_1^2(t)E\Big[\Big|\int_0^{g(t)} (f_1(t,s,x_n(s,\omega),\omega)-f_1(t,s,x_0(s,\omega),\omega))\,ds\Big|^2\Big]+$$

$$+6\bar{k}_2^2(t)E\Big[\Big|\int_0^{g(t)} (f_2(t,s,x_n(s,\omega),\omega)-f_2(t,s,x_0(s,\omega),\omega))\,dw(s,\omega)\Big|^2\Big]+$$

$$+6l^2(t)E[|x_n(h(t),\omega)-x_0(h(t),\omega)|^2]+2E[|(Ux_0)(t,\omega)-x_0(t,\omega)|^2] \leqq$$

$$\leqq (6\bar{k}_1^2(t)\tilde{k}_1^2(t)g(t)+6\bar{k}_2^2(t)\tilde{k}_2^2(t))\int_0^{g(t)} E[|x_n(s,\omega)-x_0(s,\omega)|^2]\,ds+$$

$$+6l^2(t)E[|x_n(h(t),\omega)-x_0(h(t),\omega)|^2]+2E[|U(x_0)(t,\omega)-x_0(t,\omega)|^2] \leqq$$

$$\leqq k(t)\int_0^{g(t)} u_0(s)\,ds+l(t)u_0(h(t))+r(t) = u_0(t).$$

Now (4.4) follows by induction.

It follows from (4.4) that (4.5) holds for $n=0$, $m=0,1,\dots$. Now the inequality (4.5) follows from

$$E[|x_{n+m+1}(t,\omega)-x_{n+1}(t,\omega)|^2] \leqq (Ku_n)(t)+(Lu_n)(t) = u_{n+1}(t), \quad t\in R_+,$$

and by induction.

Since $u_n \rightrightarrows 0$ for $n\to\infty$ (see Lemma 3.2) and from (4.5) it follows that $\{x_n\}$ is a Cauchy sequence (see Definition 2.2) in $C(R_+,L_2)$. Now, since $C(R_+,L_2)$ is a complete space, there exists an $\bar{x}\in C(R_+,L_2)$ such that $x_n\to\bar{x}$. If $m\to\infty$, then (4.5) yields estimation (4.3). By the estimation

$$E[|\bar{x}(t,\omega)-(U\bar{x})(t,\omega)|^2] \leqq$$

$$\leqq 2E[|\bar{x}(t,\omega)-x_n(t,\omega)|^2]+2E[|(Ux_{n-1})(t,\omega)-(U\bar{x})(t,\omega)|^2] \leqq$$

$$\leqq 4u_n(t), \quad n=0,1,\dots, \quad t\in R_+,$$

it follows that the random function $\bar{x}$ satisfies equation (1.1).

The uniqueness part of the theorem follows immediately from assertion (c) of Lemma 3.1. Indeed, if we suppose that there exists another solution $\tilde{x}$ of equation (1.1) belonging to $L_{\mathrm{loc}}^*(u_0)$ then we easily infer that $\tilde{u}(t)=E[|\bar{x}(t,\omega)-\tilde{x}(t,\omega)|^2]\in L_{\mathrm{loc}}(u_0)$, and $\tilde{u}\leqq K\tilde{u}+L\tilde{u}$. Hence and from (c) of Lemma 3.1 it follows that $E[|\bar{x}(t,\omega)-\tilde{x}(t,\omega)|^2]=0$. This completes the proof of the theorem.

Remark 4.2. By using the Banach fixed point theorem or the concept of admissibility theory it is easy to prove that there exists a unique random solution of stochastic equation (1.1) if Assumption H is fullfiled and

$$k(t)g(t)+l(t) < 1, \quad t\in R_+, \tag{4.6}$$

where k and l are defined by (4.2).

The following theorem, which follows from part c) of Remark 3.1 and Theorem 4.1, shows that condition (4.6) is more restrictive than the assumptions of Theorem 4.1.

Theorem 4.2. *If Assumption* H, *assumption* (iii) *of Lemma* 3.1 *and condition* (3.4) *are satisfied and if* $r(t)\leqq\bar{r}t^p$, $t\in R_+$, *for some* $p, \bar{r}\in R_+$, *then the assertion of Theorem* 4.1 *holds provided* $\bar{l}\bar{h}^p<1$.

References

[1] A. T. Bharucha-Reid, *Random integral equations*, Academic Press (New York, 1972).

[2] J. L. Doob, *Stochastic Processes*, Wiley (New York, 1953).

[3] I. I. Gihman and A. V. Skorohod, *Stochastic differential equations* (Russian), Izdat. Naukova Dumka (Kiev, 1968).

[4] M. Kwapisz and J. Turo, Some integral-functional equations, *Funkcialaj Ekvacioj*, **18** (1975), 107—162.

[5] M. Kwapisz and J. Turo, Existence, uniqueness and successive approximations for a class of integral-functional equations, *Aequationes Math.*, **14** (1976), 303—323.

[6] N. M. Manougian, A. N. V. Rao and C. P. Tsokos, On a nonlinear stochastic integral equation with application to control systems, *Ann. Math. Pura Appl.*, **110** (1976), 211—222.

[7] T. T. Soong, *Random differential equations in science and engineering*, Academic Press (New York, 1973).

[8] C. P. Tsokos and W. J. Padgett, *Random integral equations with applications to life sciences and engineering*, Academic Press (New York, 1974).

[9] J. Turo, Existence theory for a class of nonlinear random functional integral equations, *Ann. Math. Pura Appl.*, **121** (1979), 145—155.

[10] J. Turo, On a nonlinear random functional integral equation, *Zeszyty Nauk. Politech. Gdańsk. Mat.*, **11** (1978), 101—107.

[11] T. Ważewski, Sur une procédé de prouver la convergence des approximations successive sans utilisation des séries de comparaison, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, **8** (1960), 45—52.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF GDAŃSK
MAJAKOWSKIEGO 11—13
80—952 GDAŃSK—WRZESZCZ, POLAND

Acta Sci. Math. 44 (1982), 329—334

Attractors of systems close to autonomous ones

M. FARKAS

1. Introduction. Most of the papers on stability theory of non-autonomous systems of differential equations start with the non generic assumption that $x=0$ is a solution of the system $\dot{x}=f(t, x)$, i.e. $f(t, 0)\equiv 0$. (The reason of this is that this state of affairs is achieved in case, originally, another system is given with a known solution and the system for the variation of the solutions is formed.) Now, clearly, in the generic case the solution x of the equation $f(t, x)=0$ will depend on t and will *not* be a constant. The method of averaging helps us to get rid of the above mentioned assumption by substituting the original non-autonomous system by the averaged autonomous one (see, e.g., [4], [5]). The cases successfully dealt with by the method of averaging are the most important ones, still, these are special cases in which it is assumed that the original system is periodic, almost periodic, asymptotically almost periodic, etc. These assumptions make it possible to say something about the stable solution of the original system that emanates from the stable equilibrium of the averaged autonomous one.

In this paper these assumptions will be dismissed apart from the assumption that the system is close in a certain simple sense to an autonomous one. An asymptotically stable equilibrium of the latter system gives rise to an attractor of the original non-autonomous system. This attractor is, in general, not the integral curve of a single solution but an invariant set which is the thinner the closer the two systems are to each other. The existence of this attractor is ensured by theorems due to YOSHIZAWA [7], [8].

We are giving an explicit upper estimate of this attractor and a lower estimate to its region of attractivity. We omit here the proof of the estimates since it is similar to the proof of the theorem of paper [1]. As an example, the results are applied to van der Pol's equation under bounded perturbation in case time tends to minus infinity.

The problem treated here is connected with the problem of structurally stable ("rough"=грубые in the Soviet literature) properties of systems (see., e.g.,

Received March 18, 1981, and in revised form January 29, 1982.

Sections 18 and 19 in [2]). However, in our case the majorant function is a constant and so the perturbation need not disappear at the equilibrium point of the unperturbed system. What could be considered here as a "structurally stable property" is the perseverance of an attractor whose character may change. The problem is also connected to the concept of practical stability (see [3] Section 25) and the result can be considered as a method to estimate the "region of practical stability".

2. The attractor and the region of attractivity of a non-autonomous system close to an autonomous one. Assume that $\Omega \subset R^n$ is an open set containing the origin,

$$f \in C^0[R^+ \times \Omega, R^n], \quad f'_x \in C^0[R^+ \times \Omega, R^{n^2}], \quad g \in C^2[\Omega, R^n],$$

and for any compact $Q \subset \Omega$, $|f'_x|$ is bounded over $R^+ \times Q$ where $R^+ = [0, \infty)$ and $x = \mathrm{col}\,(x_1, \ldots, x_n) \in R^n$. Consider the systems of differential equations

$$\dot{x} = f(t, x) \tag{1}$$

and

$$\dot{x} = g(x) \tag{2}$$

where dot denotes differentiation with respect to $t \in R^+$. Assume further that there exists an $\eta > 0$ such that

$$|f(t, x) - g(x)| < \eta, \quad (t, x) \in R^+ \times \Omega. \tag{3}$$

Without loss of generality let $g(0) = 0$, and assume that the real parts of all the eigenvalues of the matrix $g'(0)$ are negative.

Under these conditions, as it is well known, one can find a positive definite quadratic form $w(x) = x^T W x$ where x^T denotes the transpose of the column vector x, such that the derivative of w with respect to system

$$\dot{y} = g'(0) y \tag{4}$$

is negative definite. Moreover, there exist constants $\varrho_1 > 0$, $\varrho_2 > 0$ such that

$$U_{\varrho_1} = \{x \in R^n : |x| < \varrho_1\} \subset \Omega, \quad \text{and} \quad \dot{w}_{(2)}(x) \leqq -\varrho_2 w(x) \quad \text{for} \quad x \in U_{\varrho_1}. \tag{5}$$

Let us denote the eigenvalues of the positive definite matrix W by λ_i, $i = 1, 2, \ldots, n$, and let $0 < \lambda_1 \leqq \lambda_2 \leqq \ldots \leqq \lambda_n$. Clearly,

$$\lambda_1 |x|^2 \leqq w(x) \leqq \lambda_n |x|^2, \quad x \in R^n. \tag{6}$$

Finally, let us introduce the notations

$$A_\eta = \{x \in R^n : w(x) \leqq 4\lambda_n^2 \eta^2 / \varrho_2^2 \lambda_1\}, \tag{7}$$

$$B = \{x \in R^n : w(x) < \lambda_1 \varrho_1^2\}. \tag{8}$$

We are now in the position to state the following

Theorem. *Under the conditions imposed upon systems* (1), (2) *and* (4), *if*

(9) $$0 < \eta < \lambda_1 \varrho_1 \varrho_2 / 2\lambda_n,$$

then the set $R^+ \times A_\eta$ *is a uniform asymptotically stable invariant set of system* (1) *and its region of attractivity contains the set* $R^+ \times B$.

The proof is similar to the proof of the Theorem in [1].

3. Van der Pol's equation under bounded perturbation. It is well known (see e.g. [6]) that for van der Pol's equation

(10) $$d^2u/d\tau^2 + m(u^2-1)\,du/d\tau + u = 0, \quad m > 0$$

the origin of the phase plane $(u, du/d\tau) = (0, 0)$ is an asymptotically stable equilibrium *in the past*, i.e., for $\tau \to -\infty$, whose region of attractivity is the open region inside the path of the single non-constant periodic solution. Substituting $t = -\tau$ equation (10) turns into

(11) $$\ddot{u} + m(1-u^2)\dot{u} + u = 0, \quad m > 0$$

where dot denotes differentiation with respect to t. For equation (11) the origin $(u, \dot{u}) = (0, 0)$ is asymptotically stable (in the future) with a bounded region of attractivity.

We are going to consider (11) under a bounded non-autonomous perturbation. First of all a Liapunov function will be constructed to the system

(12) $$\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_1 - mx_2 + mx_1^2 x_2$$

which is equivalent to (11): $x_1 = u$. The linearized system is

(13) $$\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_1 - mx_2.$$

The latter system is, clearly, asymptotically stable and, thus, it is easy to find a positive definite quadratic form whose derivative with respect to system (13) is negative definite. For instance, the quadratic form

(14) $$w(x) = \frac{m^2+2}{2m} x_1^2 + x_1 x_2 + \frac{1}{m} x_2^2$$

is positive definite and $\dot{w}_{(13)}(x) = -(x_1^2 + x_2^2)$. Moreover,

(15) $$\dot{w}_{(13)}(x) \leqq -\alpha w(x)$$

if $0 < \alpha \leqq m\left(1 - m/(m^2+4)^{1/2}\right)$. The derivative of w with respect to the system (12) is

$$\dot{w}_{(12)}(x) = -(x_1^2 + x_2^2) + x_1^2(mx_1x_2 + 2x_2^2).$$

Introducing the notation

$$\varrho_2 = m\left(1-m/(m^2+4)^{1/2}\right)-\delta \tag{16}$$

where $0<\delta<m\left(1-m/(m^2+4)^{1/2}\right)$, we are going to determine $\varrho_1>0$ so that $\dot{w}_{(12)}(x)\leqq -\varrho_2 w(x)$ should hold for $|x|<\varrho_1$. In the expression

$$-\dot{w}_{(12)}(x)-\varrho_2 w(x) = -x_1^2(mx_1x_2+2x_2^2)+\delta w(x)-[\dot{w}_{(13)}(x)+m\left(1-m/(m^2+4)^{1/2}\right)w(x)]$$

the quadratic form in square brackets is negative semidefinite in view of (15). Thus, the whole expression is non-negative provided that

$$\delta w(x)-x_1^2(mx_1x_2+2x_2^2) = \left(\delta\frac{m^2+2}{2m}-2x_2^2\right)x_1^2+(\delta-mx_1^2)\,x_1x_2+\frac{\delta}{m}x_2^2 \geqq 0. \tag{17}$$

In case $x_1x_2<0$, if

$$\delta\frac{m^2+2}{2m}-2x_2^2 \geqq 0 \text{ and } 2\left(\frac{\delta}{m}\left(\delta\frac{m^2+2}{2m}-2x_2^2\right)\right)^{1/2} \geqq \delta-mx_1^2 \geqq 0,$$

then

$$0 \leqq \left(\left(\delta\frac{m^2+2}{2m}-2x_2^2\right)^{1/2}x_1+\left(\frac{\delta}{m}\right)^{1/2}x_2\right)^2 \leqq$$

$$\leqq \left(\delta\frac{m^2+2}{2m}-2x_2^2\right)x_1^2+\frac{\delta}{m}x_2^2+(\delta-mx_1^2)x_1x_2.$$

In case $x_1x_2\geqq 0$, (17) holds provided that

$$\delta\frac{m^2+2}{2m}-2x_2^2 \geqq 0 \quad\text{and}\quad \delta-mx_1^2 \geqq 0.$$

A simple calculation yields that in both cases (17) holds if

$$x_1^2+x_1^2 \leqq \frac{\delta}{m}\min\left(1, \frac{m^2+4}{8}\right).$$

We can summarize the result in the following way. Let us define the function

$$r^2(m) = \begin{cases} (m^2+4)/8m & \text{if } 0<m\leqq 2 \\ 1/m & \text{if } 2<m. \end{cases}$$

If $|x|<r(m)\delta^{1/2}$ then $\dot{w}_{(12)}(x)\leqq -\varrho_2 w(x)$ where ϱ_2 is given by (16). Thus, our Theorem can be applied to the equation

$$\ddot{u}+m(1-u^2)\dot{u}+u = F(t, u, \dot{u}),$$

if $F, F'_u, F'_{\dot{u}}$ are continuous functions and $|F(t, u, \dot{u})|<\lambda_1\varrho_1\varrho_2/2\lambda_2$ for $t\in R^+$, $(u^2+\dot{u}^2)^{1/2}<\varrho_1$ where $\varrho_1=r(m)\delta^{1/2}$, ϱ_2 is given by (16) and $0<\lambda_1<\lambda_2$ are the easily computable eigenvalues of the quadratic form (14). Instead of giving the details in general, we are presenting a numerical example setting $m=0.20$.

Consider the equation

$$\ddot{u}+0.20(1-u^2)\dot{u}+u = F(t, u, \dot{u}). \tag{18}$$

The quadratic form (14) is now

$$w(x) = 5.1x_1^2+x_1x_2+5.0x_2^2.$$

$\varrho_2=0.18-\delta$, $r^2(0.20)=2.5$, $\varrho_1=1.6\delta^{1/2}$ and the eigenvalues of the quadratic form are $\lambda_1=4.6$, $\lambda_2=5.6$. It is assumed that F satisfies $|F(t, x_1, x_2)|<\eta$ for $t\in R^+$, $|x|<1.6\delta^{1/2}$. The value of η will be specified later. The projections of the attractor and its region of attractivity to the x-plane are, by (7) and (8),

$$A = \{x\in R^2\colon w(x) \leqq \eta^2 27/(0.18-\delta)^2\}, \quad B = \{x\in R^2\colon w(x) < 11\delta\},$$

respectively. According to (16), δ can be chosen arbitrarily between 0 and 0.18. We want to minimize the attractive set A_η and maximize its region of attractivity B at the same time. A way of doing this is to maximize the ratio

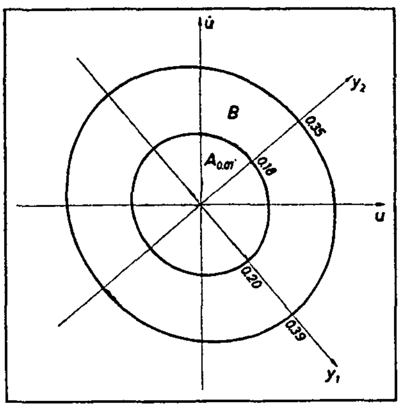
$$\frac{11\delta}{27/(0.18-\delta)^2} = 0.41\,\delta(0.18-\delta)^2.$$

One easily gets that the maximum of this ratio in the interval [0, 0.18] is achieved at $\delta=0.060$.

Substituting this value of δ into the formulae we get

$$A_\eta = \{x\in R^2\colon\ w(x) \leqq \eta^2 1900\}, \quad B = \{x\in R^2\colon\ w(x) < 0.66\}.$$

Thus, if $\eta<(0.66/1900)^{1/2}=0.019$ and $|F(t, x_1, x_2)|<\eta$ for $t\in R^+$, $|x|<0.39$, then $R^+\times A_\eta$ is a uniform asymptotically stable invariant set of the equation (18) and the set $R^+\times B$ is contained in its region of attractivity. The Figure below shows the projection of these sets into the $(u, \dot{u})=(x_1, x_2)$ plane in case $\eta=0.01$.



Figure

References

[1] M. Farkas, Attractors of systems close to periodic ones, *Nonlinear Analysis*, **5** (1981), 845—851.
[2] Н. Н. Красовский, *Некоторые задачи теории устойчивости движения*, Гос. Изд. Физ. Мат. Лит. (Москва, 1959).
[3] J. LaSalle and S. Lefschetz, *Stability by Liapunov's direct method*, Academic Press (New York, London, 1961).
[4] Ю. А. Митропольский, *Метод усреднения в нелинейной механике*, Наукова Думка (Киев, 1971).
[5] Ю. А. Митропольский, О. Б. Лыкова, *Интегральные многообразия в нелинейной механике*, Наука (Москва, 1973).
[6] M. Urabe, *Nonlinear autonomous oscillations*, Academic Press (New York—London, 1967).
[7] T. Yoshizawa, Asymptotic behavior of solutions of a system of differential equations, *Contr. Diff. Equ.*, 1 (1963), 371—387.
[8] T. Yoshizawa, *Stability theory by Liapunov's second method*, Publications of the Mathematical Society of Japan, No. 9, Math. Soc. Japan (Tokyo, 1966).

DEPARTMENT OF MATHEMATICS
BUDAPEST UNIVERSITY OF TECHNOLOGY
1521 BUDAPEST, HUNGARY

Acta Sci. Math. 44 (1982), 335—344

Polynomials over groups and a theorem of Fejér and Riesz

A. LEBOW and M. SCHREIBER

1. Introduction

A theorem of L. Fejér and F. Riesz asserts that every non-negative trigonometric polynomial is the absolute square of another trigonometric polynomial.[1]) In this note we show that the theorem does not hold in several variables.

We discovered this in the course of seeking a theorem of Fejér—Riesz type as a statement about polynomials over groups. The idea of polynomials over a group G is adequately expressed, it seems to us, by the discrete complex group algebra $\mathbf{C}[G]$ of G. This algebra is the set of complex valued functions on G of finite support, endowed with functional addition and convolution multiplication. It has an involution and a scalar product which fulfill the H^* axiom of AMBROSE [3]. In such an algebra one can define the *positivity dual* $'A$ of a subset $A\subset\mathbf{C}[G]$, this being the set of elements having non-negative scalar product with all elements of A. We consider the subset $\mathrm{S}(\mathbf{C}[G])$ of *hermitian squares* ff^* of elements of $\mathbf{C}[G]$ and, if G is abelian, the subset $\mathbf{P}(\mathbf{C}[G])$ of *positive* elements, these being the elements with non-negative Fourier transform. Interpreting these subsets for the group $\mathbf{Z}$ of integers one sees that the Fejér—Riesz theorem is equivalent to the relation $\mathbf{P}(\mathbf{C}[\mathbf{Z}])=\mathrm{S}(\mathbf{C}[\mathbf{Z}])$. Accordingly we say, for any discrete abelian group G, that the *extended* Fejér—Riesz theorem holds for G if $\mathbf{P}(\mathbf{C}[G])=\mathrm{S}(\mathbf{C}[G])$, "positive equals square".

For the class of discrete abelian groups we find, by positivity-duality and harmonic analysis, that $\mathrm{S}\subset\mathbf{P}='\mathbf{P}='\mathrm{S}$, so that $\mathbf{P}=\mathrm{S}$ if and only if $'\mathrm{S}=\mathrm{S}$, which is to say that a necessary and sufficient condition for the truth of the extended Fejér—Riesz theorem over a discrete abelian group is the self-duality of its set of hermitian squares.

For any finite group, abelian or not, we find by pure algebra (the Wedderburn theorem) that always $'\mathrm{S}=\mathrm{S}$, and therefore in particular that S is a cone. If one de-

Received May 14, and in revised form November 6, 1981.

[1]) According to Fejér's account of the matter, he conjectured the result and Riesz gave the first proof. Fejér gives that proof in his paper [1] of 1916. A more accessible reference is Pólya and Szegő [2], Sechster Abschnitt, Problem 40.

fines **P** for finite groups in terms of a natural operator-valued Fourier transform one gets also the Fejér—Riesz relation $\mathbf{P}=\mathbf{S}$.

In the investigation of the extent of validity of the self-duality $'\mathbf{S}=\mathbf{S}$, or equivalently in the abelian case the Fejér—Riesz relation $\mathbf{P}=\mathbf{S}$, the next case after **Z** to check is $\mathbf{Z}\oplus\mathbf{Z}$. Here we find, by essentially algebraic means, a class of counterexamples to the result, and this shows also that the Fejér—Riesz theorem does not hold for trigonometric polynomials in several variables. By the same methods we also find that $\mathbf{S}(\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}])$ is not a cone.

It is a pleasure to thank Dr. D. Petz (Budapest) for a number of helpful suggestions.

2. Discrete group algebra

Let G be a group, thought of as multiplicative. By the *discrete complex group algebra* $\mathbf{C}[G]$ of G one means the set of all complex valued functions $f\colon G\to\mathbf{C}$ of finite support, endowed with functional addition and convolution as multiplication. We write $f\in\mathbf{C}[G]$ as $f=\Sigma f(x)x$, regarding $f(x)\in\mathbf{C}$ as the coordinates of f relative to the elements of G as a Hamel base of $\mathbf{C}[G]$. Then $f+g=\Sigma\,(f(x)+g(x))x$, $fg=$ $=(\Sigma f(x)x)(\Sigma g(y)y)=\Sigma(\Sigma f(xy^{-1})g(y))x$, all sums being automatically finite. With the scalar product $\langle f,g\rangle=\Sigma f(x)\overline{g(x)}$ the algebra is a pre-Hilbert space. It admits the involution $f\to f^*=\Sigma\overline{f(x^{-1})}x$. This involution and the two products (scalar and algebra) are related by the H^* law of Ambrose [3],

$$\langle fg,h\rangle=\langle g,f^*h\rangle=\langle f,hg^*\rangle \tag{1}$$

where $f,g,h\in\mathbf{C}[G]$. This may be proved at once by checking it on group elements viewed as members of the algebra, observing that $x^*=x^{-1}$ for $x\in G$. $\mathbf{C}[G]$ is not a normed algebra unless G is the trivial group.

By the *positivity dual* $'A$ of a subset $A\subset\mathbf{C}[G]$ one means $\{g\in\mathbf{C}[g]\colon\langle g,f\rangle\geqq0$ for all $f\in A\}$. The positivity dual resembles in its simplest properties the commutor of a set of elements, and the notation "prime before" is intended to suggest the resemblance. In particular we have

$$A\subset B\quad\text{implies}\quad 'B\subset 'A. \tag{2}$$

We denote by $\mathbf{S}(\mathbf{C}[G])$ the set $\{ff^*\colon f\in\mathbf{C}[G]\}$ of *hermitian squares* in $\mathbf{C}[G]$. We have

$$\mathbf{S}(\mathbf{C}[G])\subset '\mathbf{S}(\mathbf{C}[G]) \tag{3}$$

by (1), as follows. If $f,g\in\mathbf{S}$, say $f=uu^*$, $g=vv^*$, then $\langle f,g\rangle=\langle uu^*,vv^*\rangle=$ $=\langle u^*,u^*vv^*\rangle=\langle u^*v,u^*v\rangle\geqq0$.

A linear operator $T\in L(H)$ on a pre-Hilbert space H is *operator-positive*, in symbols $T\geqq 0$, if $\langle T\varphi, \varphi\rangle\geqq 0$ for all $\varphi\in H$. We denote by λ_* the *left regular representation* of $\mathbf{C}[G]$ (thus $\lambda_*(f)g=fg$ for $f, g\in\mathbf{C}[G]$), by $\Lambda(\mathbf{C}[G])$ the set $\{f\colon \lambda_*(f)\geqq 0\}$ of elements of $\mathbf{C}[G]$ which go over into positive operators in the left regular representation, and by $\mathbf{PD}(\mathbf{C}[G])$ the set of *positive definite* elements of $\mathbf{C}[G]$.[2]) We have

$$\text{(4)} \qquad \mathbf{PD}(\mathbf{C}[G]) = {}'\mathbf{S}(\mathbf{C}[G]) = \Lambda(\mathbf{C}[G]),$$

as follows. For any $f, g\in\mathbf{C}[G]$, $\Sigma\Sigma f(xy^{-1})\overline{g(x)}g(y) = \Sigma\Sigma f(t)\overline{g(ty)}g(y) = \Sigma f(t)\Sigma g(ty)\overline{g(y)}=\langle f, gg^*\rangle$; and since the generic $g\in\mathbf{C}[G]$, a function of finite support, determines the generic finite subset $\{c_j\}\subset\mathbf{C}$, the equation proves that $\mathbf{PD}={}'\mathbf{S}$. Since also $\langle\lambda_*(f)g, g\rangle=\langle f, gg^*\rangle$, we have $f\in{}'\mathbf{S}$ if and only if $\lambda_*(f)\geqq 0$, or ${}'\mathbf{S}=\Lambda$, and the proof is complete.[3])

3. Discrete abelian groups

For a discrete abelian group G the following facts are well known [5]. The set of characters of G forms a compact group $\hat{G}$; each $f\in\mathbf{C}[G]$ has a Fourier transform $\hat{f}\colon G\to\mathbf{C}$ defined as $\hat{f}(\chi)=\Sigma\overline{\chi(x)}f(x)$, $\chi\in\hat{G}$; $\hat{f}$ is continuous on $\hat{G}$; $\widehat{f^*}=\overline{\hat{f}}$; $\widehat{fg}=\hat{f}\hat{g}$; and $\langle f, g\rangle=\int_G \hat{f}(\chi)\overline{\hat{g}(\chi)}\,d\chi$ for all $f, g\in\mathbf{C}[G]$.

We call *positive* those elements $f\in\mathbf{C}[G]$ such that $\hat{f}\geqq 0$, and we denote by $\mathbf{P}(\mathbf{C}[G])$ the set $\{f\colon \hat{f}\geqq 0\}$ of positive elements of $\mathbf{C}[G]$. The relation (3) has for discrete abelian groups the following refinement:

$$\text{(5)} \qquad \mathbf{S}(\mathbf{C}[G])\subset\mathbf{P}(\mathbf{C}[G])\subset{}'\mathbf{P}(\mathbf{C}[G])\subset{}'\mathbf{S}(\mathbf{C}[G]).$$

For if $f=gg^*\in\mathbf{S}$ then $\hat{f}=\hat{g}\overline{\hat{g}}\geqq 0$, so $f\in\mathbf{P}$; and if $f\in\mathbf{P}$ then for any $g\in\mathbf{P}$ we have $\langle f, g\rangle=\int_G\hat{f}\overline{\hat{g}}=\int_G\hat{f}\hat{g}\geqq 0$ so $f\in{}'\mathbf{P}$. Thus $\mathbf{S}\subset\mathbf{P}\subset{}'\mathbf{P}$. And $\mathbf{S}\subset\mathbf{P}$ entails ${}'\mathbf{P}\subset{}'\mathbf{S}$ by the duality relation (2).

Theorem 1. *For any discrete abelian group G we have* $\mathbf{P}(\mathbf{C}[G])={}'\mathbf{S}(\mathbf{C}[G])$.

Proof. We treat G as a locally compact abelian group. For such groups it is a consequence of the L^1 inversion theorem that an integrable positive definite func-

[2]) We adhere to the usual sense of this term: $f\in\mathbf{C}[G]$ is positive definite if $\Sigma\Sigma f(x_i x_j^{-1})c_i\bar{c}_j\geqq 0$ for all finite subsets $\{x_i\}\subset G$ and $\{c_i\}\subset\mathbf{C}$.

[3]) The relation $\mathbf{PD}={}'\mathbf{S}$ has a general form valid over locally compact groups. See [4], #13.4.4, page 256.

tion has a non-negative Fourier transform.[4]) This may be expressed (in an abbreviated notation) as follows:

$$L^1 \cap \mathbf{PD} \subset \mathbf{P}. \tag{6}$$

Since in $\mathbf{C}[G]$ all elements have finite support we have in fact $\mathbf{PD}(\mathbf{C}[G]) \subset \mathbf{P}(\mathbf{C}[G])$; in combination with (4) this yields the relation $'\mathbf{S}(\mathbf{C}[G]) \subset \mathbf{P}(\mathbf{C}[G])$; and this together with (5) yields the asserted equality, completing the proof.

Taken together with (4), Theorem 1 asserts that in $\mathbf{C}[G]$ for any discrete abelian group G the set of positive definite elements is equal to the set of elements with non-negative Fourier transform, $\mathbf{PD}(\mathbf{C}[G]) = \mathbf{P}(\mathbf{C}[G])$. This is not true over an arbitrary locally compact abelian group. Indeed, without conditions the statement may be vacuous (for example, over the group $\mathbf{R}$ of real numbers the function $\exp(ix)$ is positive definite but has no Fourier transform as a function).[5])

Let us interprete our apparatus for the group $\mathbf{Z}$. The characters are the maps $\chi_t \colon \mathbf{Z} \to \mathbf{T}^1$ defined as $\chi_t(n) = \exp(int)$, where $t \in [0, 2\pi]$ and $\mathbf{T}^1 = \{z \in \mathbf{C} \colon |z| = 1\}$. The element $f = \Sigma f(n) n \in \mathbf{C}[\mathbf{Z}]$ [6]) has involute $f^* = \Sigma \overline{f(-n)} n$ and Fourier transform $\hat{f}(\chi_t) = \Sigma f(n) \overline{\chi_t(n)} = \Sigma f(n) \exp(-int)$, a trigonometric polynomial. We have $\widehat{ff^*} = |\Sigma f(n) \exp(-int)|^2$. The theorem of Fejér and Riesz [1] is thus equivalent to the relation $\mathbf{P}(\mathbf{C}[\mathbf{Z}]) = \mathbf{S}(\mathbf{C}[\mathbf{Z}])$. Accordingly we say, for any discrete abelian group G, that the *extended* Fejér—Riesz theorem holds for G if $\mathbf{P}(\mathbf{C}[G]) = \mathbf{S}(\mathbf{C}[G])$. We are now in position to characterize those discrete abelian groups for which the extended Fejér—Riesz theorem holds.

Theorem 2. *For any discrete abelian group G the Fejér—Riesz relation* $\mathbf{P}(\mathbf{C}[G]) = \mathbf{S}(\mathbf{C}[G])$ *holds if and only if the set of hermitian squares in* $\mathbf{C}[G]$ *is self dual,* $\mathbf{S}(\mathbf{C}[G]) = {}'\mathbf{S}(\mathbf{C}[G])$.

Proof. By Theorem 1 the sequence (5) has the further refinement

$$\mathbf{S}(\mathbf{C}[G]) \subset \mathbf{P}(\mathbf{C}[G]) = {}'\mathbf{P}(\mathbf{C}[G]) = {}'\mathbf{S}(\mathbf{C}[G]), \tag{7}$$

whence $\mathbf{S} = \mathbf{P}$ if and only if $\mathbf{S} = {}'\mathbf{S}$, q.e.d.

The question of what groups fulfill this condition is essentially open. What little we know about it will be presented in Section 5.

[4]) [5] Corollaire 1, page 92.

[5]) With conditions a variety of statements can be made; for instance (in abbreviated notation) $L^1 \cap L^2 \cap \mathbf{PD} = L^1 \cap L^2 \cap \mathbf{P}$. That the left side is included in the right side comes immediately from (6); the opposite inclusion follows from the Plancherel theorem together with [4], # 13.4.4, page 256.

[6]) The summation is that of $\mathbf{C}[\mathbf{Z}]$, not that of $\mathbf{Z}$.

4. Finite groups

Our discussion of discrete abelian groups did not yield an extended Fejér—Riesz theorem for that class and in particular left open the question whether the hermitian squares form a cone. In contrast to this we find for finite groups, abelian or not, that the set of hermitian squares is a cone, and moreover in the strong sense of being self dual.[7] We have for this result a proof belonging to pure algebra.

Theorem 3. *For any finite group G we have* $\mathbf{S}(\mathbf{C}[G]) = {}'\mathbf{S}(\mathbf{C}[G])$.

Proof. It is well known that $\mathbf{C}[G]$ is the finite algebra-direct sum of minimal two-sided ideals, each of which is principal, generated by a uniquely determined idempotent, and that these idempotents are pairwise mutually annihilating, central, and hermitian. That is to say, if u_j are the generators of the minimal two-sided ideals $\mathbf{C}[G]u_j$ then $u_j^2 = u_j^* = u_j$, $u_j u_k = \delta_{jk} u_k$, and $u_j f = f u_j$ for all $f \in \mathbf{C}[G]$. Each $f \in \mathbf{C}[G]$ has the unique decomposition $f = \Sigma f u_j$ into its components $f u_j \in \mathbf{C}[G]u_j$, whence $ff^* = \Sigma f u_j \Sigma f^* u_k = \Sigma f f^* u_n$, which is to say $\mathbf{S}(\mathbf{C}[G]) \subset \oplus_j \mathbf{S}(\mathbf{C}[G]u_j)$. Conversely, if $h = \Sigma g_j g_j^*$ with $g_j \in \mathbf{C}[G]u_j$ then $g_j = g_j u_j$ and $h = \Sigma (g_j u_j)^* g_j u_j = \Sigma g_j^* u_j \cdot \Sigma g_k u_k = gg^* \in \mathbf{S}(\mathbf{C}[G])$, where $g = \Sigma g_j u_j$. Therefore

$$\mathbf{S}(\mathbf{C}[G]) = \oplus_j \mathbf{S}(\mathbf{C}[G]u_j). \tag{8}$$

Let $f \in {}'\mathbf{S}(\mathbf{C}[G])$ be given. Since for any $g \in \mathbf{C}[G]$ we have $g^* g u_j \in \mathbf{S}(\mathbf{C}[G])$ (because $g^* g u_j = g^* g u_j u_j^* = (g u_j)^* g u_j$), we have $\langle f, g^* g u_j \rangle \geqq 0$ for all j. But $\langle f, g^* g u_j \rangle = \langle f u_j, g^* g u_j \rangle$, so $f u_j \in {}'\mathbf{S}(\mathbf{C}[G]u_j)$, whence ${}'\mathbf{S}(\mathbf{C}[G]) \subset \oplus_j {}'\mathbf{S}(\mathbf{C}[G]u_j)$. Conversely, if $h = \Sigma g_j$ with $g_j \in {}'\mathbf{S}(\mathbf{C}[G]u_j)$ then for any $f = \Sigma f^* f u_j \in \mathbf{S}(\mathbf{C}[G])$ we have $\langle h, f \rangle = \Sigma \langle g_j, f^* f u_j \rangle \geqq 0$, $h \in {}'\mathbf{S}(\mathbf{C}[G])$, so that

$${}'\mathbf{S}(\mathbf{C}[G]) = \oplus_j {}'\mathbf{S}(\mathbf{C}[G]u_j). \tag{9}$$

Since the minimal ideals $\mathbf{C}[G]u_j$ are simple as rings each one is algebra-isomorphic (by the Wedderburn theorem) to the full algebra $L(H_j)$ of all linear transformations on a finite dimensional Hilberst space H_j. The left regular representation λ_* is faithful, and so maps each ideal $\mathbf{C}[G]u_j$ algebra-isomorphically onto a subalgebra of $L(\mathbf{C}[G])$. Since $\mathbf{C}[G]u_j$ is a full ring, so therefore is $\lambda_*(\mathbf{C}[G]u_j)$.

For H finite dimensional the algebra $L(H)$, though not a group algebra, has the involution $T \to T^*$ defined by the operator adjoint, and it has the scalar product $\langle T, S \rangle = \operatorname{trace}(TS^*)$, the so-called *trace inner product*, which trivially fulfills the Ambrose law (1). We may therefore define positivity duality and the sets $\mathbf{S}, {}'\mathbf{S}$ over $L(H)$.

[7]) Evidently the positivity dual of any set is a cone.

Lemma. *For any finite dimensional Hilbert space H we have* $\mathbf{S}(L(H)) = {}'\mathbf{S}(L(H))$.

Proof. We have $\mathbf{S}(L(H)) \subset {}'\mathbf{S}(L(H))$ by (3) since that result depends only upon (1). For the opposite inclusion let $c \in {}'\mathbf{S}(L(H))$ be given. Pick arbitrarily a unit vector $v \in H$, extend the set $\{v\}$ to an orthonormal basis $\{e^1 = v, e^2, \ldots, e^d\}$ of H, and let p be the orthogonal projection onto the subspace spanned by e^1. Then $\langle cv, v \rangle_H = \langle ce^1, e^1 \rangle_H = \Sigma \langle cpe^i, e^i \rangle_H = \text{trace}\,(cp) = \langle c, p \rangle = \langle c, pp^* \rangle \geqq 0$, whence c is a positive hermitian operator. If b is its positive square root then $c = bb^* \in \mathbf{S}(L(H))$ and the lemma is proved.

Returning to the proof of Theorem 3, we claim that the isomorphism λ_* is an essentially H^*-map in the sense that for all $f, g \in \mathbf{C}[G]$

$$\text{trace}\,(\lambda_*(f) \cdot \lambda_*(g)) = \#(G)\langle f, g \rangle. \tag{10}$$

For $\langle f, g \rangle = \langle fg^*, e \rangle = (fg^*)(e)$, and since $\lambda_*(f)x = fx = \Sigma f(s)sx = \Sigma f(tx^{-1})t$, whence $\langle \lambda_*(f)x, t \rangle = f(tx^{-1})$ for $x, t \in G$, we have also $\text{trace}\,(\lambda_*(f)) = \#(G)f(e)$. As evidently $\lambda_*(g)^* = \lambda_*(g^*)$ we have finally $\text{trace}\,(\lambda_*(f) \cdot \lambda_*(g)) = \text{trace}\,(\lambda_*(fg^*)) = \#(G)(fg^*)(e) = \#(G)\langle f, g \rangle$, which is (10). By the lemma we therefore conclude that $\mathbf{S}(\mathbf{C}[G]u_j) = {}'\mathbf{S}(\mathbf{C}[G]u_j)$ for all j, and tracing this back through (9) and (8) we reach the assertion of the theorem, q.e.d.

We turn now to the question whether one can define "positive" over finite groups in such a way as to substantiate the Fejér—Riesz relation. By (4) the elements of $'\mathbf{S}$ go over to positive operators in the regular representation. If a definition of $\mathbf{P}$ consistent with this fact can be formulated, then automatically one will have $'\mathbf{S} = \mathbf{P}$, and also automatically, by Theorem 3, the Fejér—Riesz relation $\mathbf{P} = \mathbf{S}$. The following considerations lead to such a formulation.

Definition 1. By the *unitary dual object* $\hat{G}_u$ of a finite group G we mean the set of all equivalence classes of irreducible unitary complex representations of G.[8])

Let $[X]$ denote the similarity class of the operator X or the equivalence class of the representation X, as context requires. For any representation ϱ we write ϱ_* for the extension of ϱ to the discrete group algebra $\mathbf{C}[G]$.

Definition 2. The *Fourier transform* $\hat{f}$ of $f \in \mathbf{C}[G]$, G finite, is the map of $\hat{G}_u$ to similarity classes of operators given by

$$\hat{f}([\varrho]) = [\varrho_*(f)] \tag{11}$$

for $[\varrho] \in \hat{G}_u$.

[8]) This is a variant of a procedure discussed in [6] without attribution.

Operator positivity is of course a unitary invariant, and it is known that equivalent irreducible unitary representations are in fact unitarily equivalent.[9]) It follows that, for any $[\varrho]\in\hat{G}_u$ and $f\in\mathbf{C}[G]$, $\alpha_*(f)\geqq 0$ for a single $\alpha\in[\varrho]$ if and only if $\alpha_*(f)\geqq\geqq 0$ for all $\alpha\in[\varrho]$. This makes possible the following definition.

Definition 3. For $f\in\mathbf{C}[G]$ and $\xi\in\hat{G}_u$ we say $\hat{f}$ is *positive at* ξ, and write $\hat{f}(\xi)\geqq 0$, if $\alpha_*(f)\geqq 0$ for any, hence all, $\alpha\in\xi$; we say $\hat{f}$ is *positive*, and write $\hat{f}\geqq 0$, if $\hat{f}(\xi)\geqq 0$ for all $\xi\in\hat{G}_u$.

Having formulated this concept of positive transform we now say, as in the previous case, that $f\in\mathbf{C}[G]$ for G finite is *positive* if $\hat{f}\geqq 0$, and we denote by $\mathbf{P}(\mathbf{C}[G])$, as before, the set $\{f\colon \hat{f}\geqq 0\}$ of such positive elements. The consistency with (4) of this definition of $\mathbf{P}$ follows from the reducibility of the left regular representation of G, as we now show.

Theorem 4. *For any finite group G we have* $\mathbf{P}(\mathbf{C}[G])={}'\mathbf{S}(\mathbf{C}[G])$.

Proof. Let λ denote the left regular representation of G. It is well known that every irreducible unitary representation of G is (equivalent to) a direct summand of λ with multiplicity equal to its degree.[10]) Let $\lambda^{(j)}$ be the irreducible subrepresentations of λ, and d_j their degrees. Then $\lambda\cong\oplus d_j\lambda^{(j)}$. For the extension λ_* of λ to $\mathbf{C}[G]$, which is of course nothing but the left regular representation of $\mathbf{C}[G]$[11]), we then have $\lambda_*(f)\cong\oplus d_j\lambda_*^{(j)}(f)$ for all $f\in\mathbf{C}[G]$. Now $f\in\mathbf{P}(\mathbf{C}[G])$ if and only if $\lambda_*^{(j)}(f)\geqq 0$ for all j, hence if and only if $\lambda_*(f)\geqq 0$, which is to say if and only if $f\in\mathbf{A}(\mathbf{C}[G])$. The proof is now completed by an appeal to (4).

Corollary. *For any finite group G we have the Fejér—Riesz relation* $\mathbf{P}(\mathbf{C}[G])==\mathbf{S}(\mathbf{C}[G])$.

5. The group $\mathbf{Z}\oplus\mathbf{Z}$

For finite groups we have $\mathbf{S}={}'\mathbf{S}$ as a matter of pure algebra, for discrete abelian groups generally we have $\mathbf{S}\subset\mathbf{P}={}'\mathbf{P}={}'\mathbf{S}$, and for $\mathbf{Z}$ in particular we have the self duality $\mathbf{S}={}'\mathbf{S}$, this being an equivalent formulation over discrete abelian groups of the Fejér—Riesz relation $\mathbf{S}=\mathbf{P}$. One naturally inquires into the extent of validity of this self duality, or equivalently, of the validity of the extended Fejér—Riesz theorem. In this inquiry the next case to check after $G=\mathbf{Z}$ is $G=\mathbf{Z}\oplus\mathbf{Z}$. We find that for $\mathbf{Z}\oplus\mathbf{Z}$ the extended theorem fails, $\mathbf{P}\neq\mathbf{S}$. As we shall see in a moment, this will

[9]) See for instance [7], (3.2), page 19.

[10]) [8], page 1—18.

[11]) In agreement with our previous use of the symbol λ_*.

show that the Fejér—Riesz theorem fails for trigonometric polynomials in several variables.

We will demonstrate this by exhibiting a class of counterexamples. For this purpose it proves convenient to employ Laurent polynomials, as follows. Over **Z** we may express the generic $f=\Sigma f(n)n\in \mathbf{C}[\mathbf{Z}]$ as the Laurent polynomial $f(z)=\Sigma f(n)z^n$ in the complex variable z. Evidently the addition and multiplication of Laurent polynomials duplicate the corresponding operations in $\mathbf{C}[\mathbf{Z}]$, or, in algebraic language, the set of Laurent polynomials is **C**-algebra isomorphic to $\mathbf{C}[\mathbf{Z}]$. If we put $f^*(z)=$ $=\Sigma \overline{f(n)}z^{-n}$ and use the obvious scalar product then the isomorphism preserves the Ambrose law (1) as well. In the Laurent version the Fourier transform $\hat{f}$ of $f\in\mathbf{C}[\mathbf{Z}]$ is the restriction to $\mathbf{T}^1$ of the corresponding Laurent polynomial. Over $\mathbf{Z}\oplus\mathbf{Z}$ we proceed analogously. We have elements $f=\Sigma f(n, m)(n\oplus m)\in\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}]$ and their Fourier transforms $\hat{f}(\chi_t, \chi_s)=\Sigma f(n, m)\exp(-int)\exp(-ins)$, trigonometric polynomials in two variables, which we may view as the restrictions to $\mathbf{T}^2$ of the Laurent polynomials $f(z, w)=\Sigma f(n, m)z^n w^m$ in two complex variables. Our discussion of **P** and **S** over $\mathbf{Z}\oplus\mathbf{Z}$ will thus also be a treatment of the Fejér—Riesz theorem in two variables. We note for reference that the involution in two variables reads $f^*(z, w)=$ $=\Sigma\overline{f(n, m)}z^{-n}w^{-m}$.

If $f\in\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}]$ is a hermitian square, so that $f(z, w)=h(z, w)h^*(z, w)$ for some Laurent polynomial $h(z, w)$, then we claim that without loss of generality $h(z, w)$ may be assumed to be *analytic* in z and w, and to have non-zero coefficient of z^0 when written as a polynomial in z, which is to say $h(z, w)=\sum_0^n a_k(w)z^k$ with $a_k(w)\in$ $\in\mathbf{C}[w]$[12]) and $a_0(w)\neq 0$. For by definition of the involution we have $h^*(z, w)=$ $=\bar{h}(z^{-1}, w^{-1})$, the bar denoting the conjugation of all constants. Therefore the lowest negative powers (negative exponents of greatest absolute value) of z and w occurring in $h(z, w)$ are the negatives of the highest (positive) powers of z and w occurring in $h^*(z, w)$; hence by factoring out of $h(z, w)$ the lowest negative powers of z and w, and out of $h^*(z, w)$ the highest powers, we shall have cancellation. Therefore we can substitute for $h(z, w)$ an analytic polynomial. If $a_s(w)$, $s>0$, is the non-zero coefficient of least index in $h(z, w)$, now assumed analytic, we can factor z^s out of $h(z, w)$ and z^{-s} out of $h^*(z, w)$ and cancel them, arriving thus at new polynomials with non-zero coefficients of z^0 ("constant terms").

Example. The element $f\in\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}]$ whose Laurent polynomial is $f(z, w)=$ $=(1/4)\{z^2+z^{-2}+w^2+w^{-2}+4\}$ is a hermitian square because $\hat{f}(\chi_t, \chi_s)=$

[12]) We adhere to the standard notations $F[t]$ for the ring of polynomials and $F(t)$ for the field of rational functions over the field F. There is no conflict with the notation $\mathbf{C}[G]$ where G is a group.

$= 1+(1/2)\cos(2t)+(1/2)\cos(2s) = \cos^2(t)+\cos^2(s) = |\cos(t)+i\cdot\cos(s)|^2$. From this factorization we have $f(z,w)=(1/2)\{(z+z^{-1})+i(w+w^{-1})\}\cdot(1/2)\{(z+z^{-1})-i(w+w^{-1})\}$, which is to say $f(z,w)=g(z,w)g^*(z,w)$ with $g(z,w)=(1/2)\{(z+z^{-1})+i(w+w^{-1})\}$. Factoring out the lowest negative powers we have

$$g(z,w) = z^{-1}w^{-1}(1/2)\{(z^2w+w)+i(w^2z+z)\} = z^{-1}w^{-1}\left\{\frac{w}{2}+\frac{i}{2}(w^2+1)z+\frac{w}{2}z^2\right\}.$$

Thus $f=hh^*$ where $h\in\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}]$ has the Laurent polynomial

$$h(z,w) = \frac{w}{2}+\frac{i}{2}(w^2+1)z+\frac{w}{2}z^2.$$

By means of the foregoing reduction one might hope to be able to characterize the hermitian squares in a purely algebraic way. But the system of non-linear equations one would have to discuss has so far proved intractable, and we are forced to curcumvent this difficulty by the following special arguments, which enable us to proceed a little farther.

Let $p(z,w)\in\mathbf{C}(w)[z]$ be given, and suppose its degree in z is 2. We then have $p(z,w)=p_0(w)z^2+p_1(w)z+p_2(w)$ with $p_j(w)\in\mathbf{C}[w]$. In general the equation $p(z,w)=0$ defines two branches $r_\pm=\{-p_1\pm\sqrt{(p_1^2-4p_0p_2)}\}/2p_0$; $r_\pm$ are algebraic over $\mathbf{C}(w)$, and $\mathbf{C}(w,r_+,r_-)$ is the splitting field of $p(z,w)$. Therefore $p(z,w)$ is reducible in $\mathbf{C}(w)[z]$ if and only if one, hence both, of $r_\pm$ are in $\mathbf{C}(w)$.

Consider now

(12) $$f(z,w) = (z+z^{-1})(w+w^{-1})+c, \quad c\in\mathbf{R}.$$

We rewrite this as $f(z,w)=z^{-1}\{(w+w^{-1})z^2+cz+(w+w^{-1})\}$ and put $p(z,w)=(w+w^{-1})z^2+cz+(w+w^{-1})$. The equation $p(z,w)=0$ determines the branches $r_\pm=\{-c\pm\sqrt{(c^2-4(w+w^{-1})^2)}\}/2(w+w^{-1})$. To ascertain the character of these functions we examine the radical $\sqrt{(c^2-4(w+w^{-1})^2)}=\sqrt{(c-2(w+w^{-1}))}\cdot\sqrt{(c+2(w+w^{-1}))}$. We have $c\pm2(w+w^{-1})=0$ if and only if $w^2\pm(c/2)w+1=0$, which is to say $w=\{\pm c\mp\sqrt{(c^2-16)}\}/4$. Thus except for $c=\pm4$ the functions $r_\pm$ both have branch points at these values of w, so that $r_\pm(w)\notin\mathbf{C}(w)$ and $p(z,w)$ is irreducible in $\mathbf{C}(w)[z]$ for $c\neq\pm4$.

If f were a hermitian square, so that $f(z,w)=h(z,w)h^*(z,w)$ with $h(z,w)$ analytic, $h(z,w)=\sum_0^n a_k(w)z^k$, $a_k(w)\in\mathbf{C}[w]$, $a_0(w)\neq0$, then precisely because the constant term is not zero the highest power of z occurring in $h(z,w)h^*(z,w)$ is the highest power of z occurring in $h(z,w)$. But since $f=hh^*$ this is the highest power of z occurring in $f(z,w)$, namely 1. Therefore $h(z,w)$ must have the form $\alpha(w)(z-\beta(w))$, with $\alpha(w)\in\mathbf{C}[w]$ and $\alpha(w)\beta(w)\in\mathbf{C}[w]$. Since $\beta(w)$ is at worst rational we have

$\beta(w)\in\mathbf{C}(w)$. By definition of the involution we have

$$h^*(z,w) = \bar{\alpha}(w^{-1})(z^{-1}-\bar{\beta}(w^{-1})) = \bar{\alpha}(w^{-1})z^{-1}(1-z\bar{\beta}(w^{-1})) =$$

$$= (-1)z^{-1}\bar{\alpha}(w^{-1})\bar{\beta}(w^{-1})(z-1/\bar{\beta}(w^{-1})).$$

Substituting this into our relation $f(z,w)=h(z,w)h^*(z,w)$ and recalling that $zf(z,w)=p(z,w)$ we get $p(z,w)=(-1)\alpha(w)\bar{\alpha}(w^{-1})\bar{\beta}(w^{-1})(z-\alpha(w))(z-1/\bar{\beta}(w^{-1}))$, a factorization of $p(z,w)$ in $\mathbf{C}(w)[z]$. But we have just observed, in the previous paragraph, that $p(z,w)$ is irreducible in $\mathbf{C}(w)[z]$ if $c\neq\pm 4$. Hence there can be no factorization of the form $f=hh^*$ if $c\neq\pm 4$. Since $\hat{f}(\chi_t,\chi_s)=2\cos(t)\cdot 2\cos(s)+c$ we have $f\in\mathbf{P}(\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}])$ for $c>4$. We have established.

Theorem 5. *For each real $c>4$ the Laurent polynomial $f(z,w)=(z+z^{-1})\cdot(w+w^{-1})+c$ defines an element of $\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}]$ which is positive but is not a hermitian square.*

By the same methods we have the following further result.

Theorem 6. *The set $\mathbf{S}(\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}])$ of hermitian squares over $\mathbf{Z}\oplus\mathbf{Z}$ is not a cone.*

Proof. With $h(z,w)=z+w$ put $g=hh^*$, $f(z,w)=g(z,w)+c$ with $c\in\mathbf{R}$, and $p(z,w)=z^{-1}f(z,w)$. One checks that there exists $0<c_0\in\mathbf{R}$ such that $p(z,w)$ is irreducible in $\mathbf{C}(w)[z]$. Hence $f_0(z,w)=g(z,w)+c_0$ does not correspond to a hermitian square in $\mathbf{C}[\mathbf{Z}\oplus\mathbf{Z}]$ even though f is the sum $f=hh^*+(\sqrt{c_0})(\sqrt{c_0})^*$ of hermitian squares.

Bibliography

[1] L. Fejér, Über trigonometrische Polynome, *J. reine angew. Math.*, **146** (1916), 55—82.

[2] G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus des Analysis*, second edition, Springer-Verlag (Berlin, 1954).

[3] W. Ambrose, Structure theorem for a special class of Banach algebras, *Trans. Amer. Math. Soc.*, **57** (1945), 364—386.

[4] J. Dixmier, *Les C*-algèbres et leurs représentations*, Gauthier-Villars (Paris, 1964).

[5] H. Cartan et R. Godement, Analyse harmonique et théorie de la dualité dans les groupes abeliens localement compacts, *Ann. Ecole Norm. Sup.*, 1947, Fasc. 1, pp. 79—99.

[6] K. I. Gross, On the evolution of non-commutative harmonic analysis, *Amer. Math. Monthly*, **85**, No. 7 (1978), 525—547.

[7] W. Feit, *Characters of finite groups*, Benjamin (New York, 1967).

[8] J. P. Serre, *Représentations linéares des groupes finis*, Hermann (Paris, 1967).

(M. SCH.)
THE ROCKEFELLER UNIVERSITY
1230 YORK AVENUE,
NEW YORK, N.Y. 10021, USA

(A. L.)
YESHIVA UNIVERSITY
2495 AMSTERDAM AVENUE
NEW YORK, N.Y. 10024, USA

Acta Sci. Math. 44 (1982), 345—357

Doubly stochastic, unitary, unimodular, and complex orthogonal power embeddings

ROBERT C. THOMPSON and CHI-CHIN TUNG KUO

1. Introduction

We shall say that a finite matrix A is embedded in a larger finite matrix M if A is the leading principal submatrix of M, and we write $M\supset A$ or $A\subset M$. If $M^i\supset A^i$ for $i=1, 2, \ldots, k$, we say A is power embedded in M to exponent k. We also say that M is a dilation of A. Many years ago, in connection with unitary dilation theory for Hilbert space operators, E. EGERVÁRY [2] studied power embeddings of a contraction A into a unitary M. The objective of this note is to sharpen Egerváry's result and also to obtain analogous power embedding theorems into a doubly stochastic matrix, or into an integral unimodular matrix, or into a complex orthogonal matrix. The fact that more or less analogous theorems are obtainable suggests that various other parts of the presently existing rather extensive unitary dilation theory for infinite dimensional operators is capable of expansion in various directions. See, for example [1] and [5].

In each of our cases, the dilation M will turn out to exist if and only if it has at least $k\delta$ more rows then A, where δ is a measure of how far A is itself from the doubly stochastic, unimodular, unitary, or orthogonal state.

2. Doubly stochastic power embeddings

Let A be an $\alpha\times\alpha$ matrix with nonnegative entries. We consider whether it is possible to find a power embedding of A to exponent k into a doubly stochastic matrix M. (Thus M is to have real nonnegative entries with row and column sums equal to one.) Here k is fixed and specified in advance, and we wish also to know the size of the smallest dilation M.

Of course, a doubly stochastic dilation of A could exist only if A is a contraction in the sense of nonnegative matrices, that is, has each row and column sum at most one. In this case we say that A is a double stochastic contraction. So assume that A is

Received July 27, 1981.

a double stochastic contraction, set $A = [a_{ij}]_{1 \leqq i, j \leqq \alpha}$, and let

$$d = \alpha - \sum_{i,j=1}^{\alpha} a_{ij}$$

measure the doubly stochastic deficiency of A. (Plainly $d \geqq 0$ with equality if and only if A is doubly stochastic.) Now take δ to be the least integer satisfying $\delta \geqq d$. This quantity δ is an integral measure of the doubly stochastic deficiency of A, with $\delta = 0$ precisely when contraction A is doubly stochastic.

Theorem 1. *Let the nonnegative matrix A be a doubly stochastic contraction, and define the integral doubly stochastic deficiency δ of A as above. Then A possesses a doubly stochastic power embedding M to exponent k, i.e.,*

$$M^i \supset A^i \quad \text{for} \quad 1 \leqq i \leqq k, \tag{1}$$

with M having μ more rows than A, if and only if $\mu \geqq k\delta$.

The proof requires several lemmas.

Lemma 1. *Suppose that A is not doubly stochastic, that M is, and that* (1) *holds. Then, after a permutation similarity preserving A, M takes the form $M = [M_{ij}]_{0 \leqq i, j \leqq k}$ with*

(a) $M_{00} = A$, *and the other diagonal blocks M_{ii} square,*
(b) $M_{ij} = 0$ *whenever* $j \geqq i+2$,
(c) $M_{i0} = 0$ *for all* i, $1 \leqq i < k$,
(d) *Each row of each $M_{i,i+1}$ is nonzero, for all i with $2 \leqq i \leqq k$. (Read $M_{k,k+1}$ as M_{k0}.)*

Proof. The proof is by induction on k, the case $k=1$ being trivial. Suppose the result established for k, and now assume that $M^i \supset A^i$ for $i = k+1$ also. Perform a permutation similarity on M, permuting only rows (and columns) that pass through M_{11}. Note that $M_{12} \neq 0$, since $M_{12} = 0$ forces M_{11} to be doubly stochastic, hence $M_{01} = 0$, and therefore forces $M_{00} = A$ to be doubly stochastic, a contradiction. (When $k=1$, M_{12} is M_{10}.) We chose our permutation similarity so that the nonzero rows in M_{12} are the last rows, i.e.,

$$M_{12} = \begin{bmatrix} M'_{13} \\ M'_{23} \end{bmatrix}$$

with $M'_{13} = 0$ and with each row in M'_{23} nonzero. (Conceivably, block M'_{23} is vacuous, and when $k=1$ subscript 3 is read as 0.) Partition and renumber the blocks in

M in accord with this pattern:

$$M_{00} = M'_{00}, \qquad M_{01} = [M'_{01}, M'_{02}], \qquad M_{0j} = M'_{0,j+1},$$

$$M_{10} = \begin{bmatrix} M'_{10} \\ M'_{20} \end{bmatrix}, \quad M_{11} = \begin{bmatrix} M'_{11}, M'_{12} \\ M'_{21}, M'_{22} \end{bmatrix}, \qquad M_{1j} = \begin{bmatrix} M'_{1,j+1} \\ M'_{2,j+1} \end{bmatrix},$$

$$M_{i0} = M'_{i+1,0}, \quad M_{i1} = [M'_{i+1,1}, M'_{i+1,2}], \quad M_{ij} = M'_{i+1,j+1},$$

with square blocks M'_{11}, M'_{22}. Then M partitions as $M=[M'_{ij}]_{0\leqq i,j\leqq k+1}$ with $M'_{00}=A$. At this moment, it is conceivable that block row 1 is absent. We show that $M'_{02}=0$. The leading block in M^{k+1} is

$$A^{k+1}+M'_{02}M'_{23}\dots M'_{k,k+1}M'_{k+1,0}.$$

By hypothesis this equals A^{k+1}. Since $M'_{23}, \dots, M'_{k+1,0}$ each has all rows non-zero, we deduce that $M'_{02}=0$. If the block row labeled $i=1$ were absent, so would be block column i, hence $M'_{00}=A$ would be doubly stochastic, a contradiction. This completes the induction step.

Lemma 2. *Suppose that doubly stochastic* $M=[M_{ij}]$ *has the block form described in Lemma* 1, *where* M_{ij} *is* $n_i \times n_j$. *Then* $d \leqq n_i$ *for* $1 \leqq i \leqq k$.

Proof. If Q is a matrix, σQ will denote the sum of the entries of Q. Because M is doubly stochastic and has leading block row $A, M_{01}, 0, 0, \dots, 0$, we get $\sigma M_{01}=d$. Fix p, $1 \leqq p \leqq k$. Then, by columns,

$$\sigma[M_{ij}]_{1\leqq i\leqq k, 1\leqq j\leqq p} = n_1+\dots+n_p-d,$$

hence

$$\sigma[M_{ij}]_{1\leqq i\leqq p, 1\leqq j\leqq p} \leqq n_1+\dots+n_p-d,$$

therefore,

$$\sigma[M_{ij}]_{1\leqq i\leqq p, 1\leqq j\leqq p+1}-\sigma M_{p,p+1} \leqq n_1+\dots+n_p-d.$$

But

$$\sigma[M_{ij}]_{1\leqq i\leqq p, 1\leqq j\leqq p+1} = \sigma[M_{ij}]_{1\leqq i\leqq p, 0\leqq j\leqq k} = n_1+\dots+n_p.$$

Therefore $d \leqq \sigma M_{p,p+1} \leqq \sigma M_{p0}+\sigma M_{p1}+\dots+\sigma M_{p,k}=n_p$. Hence $d \leqq n_p$ as desired; $1 \leqq p \leqq k$. (Where necessary, read subscripts modulo $k+1$.)

Lemma 3. *If A possesses a doubly stochastic power embedding M, to exponent k, with M having μ more rows than A, then* $\mu \geqq k\delta$.

Proof. $\mu=n_1+\dots+n_k$ and $d \leqq n_i$ for each i. Since n_i is integral, $\delta \leqq n_i$ for all i, therefore $\mu \geqq k\delta$.

Lemma 4. *Given* $\alpha\times\alpha$ *contraction* A, *there exists an* $(\alpha+\delta)\times(\alpha+\delta)$ *doubly stochastic matrix*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Here δ *is as before,* $\delta \geqq d$.

Proof. We apply Ky Fan's criterion [4, see also 3] for the solvability of a mixed system of linear equalities and inequalities. Let the row sums of A be r_i, and the column sums be c_j; $1 \leqq i, j \leqq \alpha$. The condition to be satisfied is that the matrix

$$\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} \tag{2}$$

have row sums $1-r_i$ in the top block row, and 1 in the other block row, plus a corresponding column statement. These are the equalities to be considered. The inequalities are that the entries of B, C, D are to be nonnegative. We treat these entries as unknowns. If (p, q) is a position in B, C, or D, we introduce real dummy variables u_p, v_q and a real nonnegative dummy variable w_{pq}. Form a column vector u from the u_p, and a row vector v from the v_q. Form also a column vector $c=[1-r_1, 1-r_2, \ldots, 1-r_\alpha, 1, 1, \ldots, 1]^T$ displaying the proposed row sums in (2), and a row vector $r=[1-c_1, 1-c_2, \ldots, 1-c_\alpha, 1, 1, \ldots, 1]$ displaying the proposed column sums. Ky Fan's test for the solvability of our mixed system of equalities and inequalities amounts to this: We must show that the conditions

$$u_p+v_q+w_{pq} = 0 \tag{3}$$

(u_p, v_q real, $w_{pq} \geqq 0$) for all (p, q) belonging to blocks B, C, D imply

$$(u, c)+(r, v) \leqq 0, \tag{4}$$

where $(\cdot, \cdot)$ is the standard inner product. This is easily done. Let U_0 be the maximum entry among $u_1, \ldots, u_\alpha$ and U_1 the maximum entry among $u_{\alpha+1}, \ldots, u_{\alpha+\delta}$. Similarly let V_0 be the maximum entry among $v_1, \ldots, v_\alpha$, and V_1 the maximum entry among $v_{\alpha+1}, \ldots, v_{\alpha+\delta}$. Noting that $\sum_1^\alpha (1-r_i) = \sum_1^\alpha (1-c_j) = d$, we get

$$(u, c)+(r, v) \leqq U_0 d + U_1\delta + V_0 d + V_1\delta =$$

$$= (U_0+V_1)d + (U_1+V_0)d + (U_1+V_1)(\delta-d) \leqq 0,$$

owing to (3) and $d \leqq \delta$.

Proof of Theorem 1. If A is doubly stochastic the necessity of $\mu \geqq k\delta$ is trivial since $\delta=0$. If A is not doubly stochastic, $\mu \geqq k\delta$ follows from Lemma 3. Conversely, let μ be any integer satisfying $\mu \geqq k\delta$. Construct the blocks B, C, D

described by Lemma 4. Now take M to be the direct sum of

$$\begin{bmatrix} A & B & 0 & 0 & \dots & 0 \\ 0 & 0 & I & 0 & \dots & 0 \\ . & . & . & . & \dots & . \\ 0 & 0 & 0 & 0 & \dots & I \\ C & D & 0 & 0 & \dots & 0 \end{bmatrix}$$

and a $(\mu-k\delta)$-square identity matrix. The identity matrices in the above block are $\delta\times\delta$. Then M is doubly stochastic and $M^i\supset A^i$ for $1\leqq i\leqq k$.

3. Unitary power embeddings

Our objective in this section is to sharpen Egerváry's theorem on unitary power embeddings. Let matrix A have complex elements. To be embeddible at all in a unitary matrix, each singular value of A must be $\leqq 1$, i.e., A must be a contraction. So assume that A is a contraction. The unitary deficiency δ of contraction A is now defined as the number (with multiplicity) of singular values of A strictly less than one. Thus $\delta=0$ if and only if A is unitary, and, in general for contraction A, δ is the rank of $I-AA^*$, also the rank of $I-A^*A$.

Theorem 2. *Let complex matrix A be a contraction, and define the unitary deficiency δ of A as above. Then A possesses a unitary power embedding M to exponent k, with M having μ more rows than A, if and only if $\mu\geqq k\delta$.*

When $k=1$, this Theorem is an easy special case of a known result [6] on singular values.

Lemma 5. *Suppose that A is not unitary, that M is, and that* (1) *holds with $k\geqq 2$. Then, after a unitary similarity preserving A, the matrix M takes the form $M=[M_{ij}]_{0\leqq i,j\leqq k}$ with*

(a) *$M_{00}=A$ and the other diagonal blocks square;*

(b) *each block is zero, except perhaps for M_{00}, M_{01}, M_{11}, M_{k0}, M_{k1}, and $M_{i,i+1}$ for $1\leqq i<k$;*

(c) *blocks $M_{23}, M_{34}, \dots, M_{k-1,k}$ are each unitary and $\delta\times\delta$;*

(d) *block M_{12} has δ columns and at least δ rows; block M_{k0} has δ rows, all linearly independent.*

Proof. We begin with $M=\begin{bmatrix} A & B \\ C & D\end{bmatrix}$. After a block diagonal similarity by a unitary matrix of the form diag (I, W), we preserve A and change C to WC. Choosing the last rows of W to be an orthonormal basis for the row space of C^*, we convert C to a matrix in which the first rows are zero and the last are linearly independent. So repartition M as

$$M = \begin{bmatrix} M_{00} & M_{01} & M_{02} \\ 0 & M_{11} & M_{12} \\ M_{20} & M_{21} & M_{22} \end{bmatrix}$$

with square diagonal blocks, $M_{00}=A$, and linearly independent rows in M_{20}. From $M^*M=I$ we get $M_{20}^*M_{20}=I-A^*A$. Since $I-A^*A$ has rank δ, this also is the rank of $M_{20}^*M_{20}$ and therefore of M_{20}. Hence M_{20} has δ rows. The leading block in M^2 is $A^2+M_{02}M_{20}$. Hence $M_{02}M_{20}=0$, and as M_{20} has independent rows, $M_{02}=0$. Orthogonality of columns then forces $M_{22}=0$. Since M_{22} is $\delta\times\delta$ and since $M_{02}=0$, M_{12} must have independent columns. It follows that M_{12} has at least δ rows. This completes the proof for $k=2$.

Suppose the result established for k, and now assume $M^{k+1}\supset A^{k+1}$. After a block diagonal unitary similarity preserving A, we may make the last rows of M_{12} independent and the remaining rows zero. Since M_{12} has δ columns, necessarily independent, this lower block in M_{12} is $\delta\times\delta$ and unitary. Now repartition as in the proof of Lemma 1. Then M_{23}' is $\delta\times\delta$ and M_{12}' possibly is vacuous. We must show that $M_{02}'=0$, M_{12}' has δ columns and at least δ rows, $M_{21}'=0$, $M_{22}'=0$, $M_{k+1,2}'=0$.

For simplicity, drop primes. Since $M^{k+1}\supset A^{k+1}$, we have

$$M_{02}M_{23}\ldots M_{k,k+1}M_{k+1,0} = 0.$$

Linear independence of rows in $M_{23}, \ldots, M_{k+1,0}$ forces $M_{02}=0$. Orthogonality of columns forces $M_{21}=0$, $M_{22}=0$, $M_{k+1,2}=0$. If M_{12} were absent, M_{00} would be forced to be a direct summand of M, hence unitary. Therefore M_{12} is present, and as M_{22} is $\delta\times\delta$, M_{12} has δ columns, necessarily independent. Therefore it has at least δ rows.

Proof of Theorem 2. Suppose that a power embedding of A into a unitary matrix M exists, to exponent k. We wish to show that the number μ of additional rows in M satisfies $\mu\geqq k\delta$. If A is already unitary this is evident. Suppose A to be not unitary. If $k=1$ we have

$$M = \begin{bmatrix} A & M_{01} \\ M_{10} & M_{11} \end{bmatrix}$$

where $M_{10}^*M_{10}=I-A^*A$ has rank δ; therefore M_{10} has at least δ rows, hence $\mu\geqq\delta$. Now suppose $k\geqq 2$. Then M may be put in the form described in Lemma 5, with blocks $M_{12}, M_{23}, \ldots, M_{k0}$ each having at least δ rows. Therefore $\mu\geqq k\delta$.

Turning to the converse, since both $I-AA^*$ and $I-A^*A$ have rank δ, nonsingular matrices X and Y exist such that

$$I-AA^* = X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*, \quad I-A^*A = Y^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix}Y.$$

Set

$$B = X\begin{bmatrix} I_\delta \\ 0 \end{bmatrix}, \quad C = -[0, I_\delta]Y, \quad D = [0, I_\delta]YA^*X^{*-1}\begin{bmatrix} I_\delta \\ 0 \end{bmatrix},$$

and form the matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{5}$$

We claim that this matrix is unitary. Certainly $AA^*+BB^*=I_\alpha$. Next, note that

$$X^{-1}AY^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} = X^{-1}A(I-A^*A)Y^{-1} = X^{-1}(I-AA^*)AY^{-1} = \tag{6}$$

$$= \begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix} X^*AY^{-1}.$$

The definition of D shows that

$$\begin{bmatrix} D^* \\ 0 \end{bmatrix} = X^{-1}AY^*\begin{bmatrix} 0 \\ I_\delta \end{bmatrix}, \quad \text{and hence} \quad X\begin{bmatrix} I_\delta \\ 0 \end{bmatrix}D^* = AY^*\begin{bmatrix} 0 \\ I_\delta \end{bmatrix}.$$

This equation is the same as $AC^*+BD^*=0$.

Finally, we show that $CC^*+DD^*=I_\delta$. We have, using (6) at one point,

$$\begin{bmatrix} 0 & 0 \\ 0 & CC^* \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} Y\cdot Y^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} = Y^{*-1}(I-A^*A)\cdot(I-A^*A)Y^{-1} =$$

$$= Y^{*-1}(I-2A^*A+A^*(AA^*)A)Y^{-1} =$$

$$= Y^{*-1}\left(I-2A^*A+A^*\left\{I-X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*\right\}A\right)Y^{-1} =$$

$$= Y^{*-1}\left(I-A^*A-A^*X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*A\right)Y^{-1} =$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - Y^{*-1}A^*X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*AY^{-1} =$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} YA^*X^{-1*}\cdot X^{-1}AY^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ D & 0 \end{bmatrix}\begin{bmatrix} 0 & D^* \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & DD^* \end{bmatrix}.$$

Hence $CC^*+DD^*=I_\delta$, as claimed.

Therefore matrix (5) is unitary, and has δ more rows than A. This settles the case $k=1$. For $k>1$, we use these blocks B, C, D together with the construction in the proof of Theorem 1. This completes the proof of Theorem 2.

Let us note two additional facts: (i) If contraction A has real entries, this proof produces a real dilation M; (ii) If contraction A has quaternion entries, this proof produces a symplectic dilation M, i.e., a unitary M with quaternion entries.

It should also be noted that essentially this proof is already in the literaure; see [1].

4. Unimodular power embeddings

Now let $\alpha\times\alpha$ matrix A have entries from a commutative principal ideal domain R. We wish to embed A to exponent k into a unimodular matrix M. We take the unimodular deficiency δ of A to be the number of non-unit invariant factors of A. Then $\delta=0$ if and only if A is unimodular. We adopt the convention that any matrix A over R now is to be viewed as a contraction. (This is not quite as unnatural as it seems: if we were to permit matrices with entries from the field of fractions of R, by p-adic theory the contractions would be just those with entries in R.)

Theorem 3. *Let matrix A have entries in the principal ideal domain R. Then A possesses a unimodular power embedding M to exponent k, with M having μ more rows than A, if and only if $\mu\geqq k\delta$, where δ is the unimodular deficiency of A defined above.*

When $k=1$, this theorem is a special case of a known result [7] on invariant factors.

Lemma 6. *Assume that A is not unimodular, and that A is power embedded in unimodular M to exponent k. Then, after a unimodular similarity preserving A, M takes the form $M=[M_{ij}]_{0\leqq i,j\leqq k}$, where*

(a) *$M_{00}=A$ and each diagonal block M_{ii} is square,*

(b) *$M_{ij}=0$ if $j\geqq i+2$,*

(c) *$M_{i0}=0$ for $1\leqq i<k$,*

(d) *Each block $M_{12}, M_{23}, \ldots, M_{k0}$ has at least δ rows, with $M_{23}, \ldots, M_{k0}$ each having all rows independent and M_{12} at least δ independent rows.*

Proof. By induction on k. For $k=1$ we need only show that M_{10} has at least δ independent rows. After a block diagonal unimodular similarity of M preserving A, no generality is lost if M_{10} is cast into Hermite form. If it has fewer than δ nonzero rows, by a column Laplace expansion $\det M$ is a linear combination of $\alpha\times\alpha$ minors formed from the first α columns of M, each minor using at least $\alpha-\delta+1$ rows from M_{00}. Thus $\det M$ is a linear combination of $(\alpha-\delta+1)$-square minors from M_{00}. If s is the first nonunit invariant factor of $A=M_{00}$, each of these minors is divisible by s, so that s is a factor of $\det M$. This is impossible.

Now assume the result for k. By a block diagonal unimodular similarity of M

preserving A, we may cast M_{12} into row Hermite form, with the nonzero (and independent) rows last. Now repartition M as before. Then $M^{k+1} \supset A^{k+1}$ implies $M_{02}M_{23} \dots M_{k,k+1}M_{k+1,0}=0$ and hence (by independence of rows) $M_{02}=0$. Let the blocks be $n_i \times n_j$. We must show that $\delta \leqq n_1$ and that M_{12} has at least δ independent rows. To see that $\delta \leqq n_1$, expand det M down its first $n_0=\alpha$ columns, with complementary minors coming from the last $n_1+\dots+n_{k+1}$ columns. If a minor uses x rows from M_{00}, and if this minor is not to be divisible by s we must have $x \leqq n_0-\delta$. Also, it must use n_0-x rows from $M_{k+1,0}$.

Expand the complementary minor down the columns running through M_{01}. A nonzero minor in this expansion must use all the rows in M_{01} that were not used in M_{00}, and perhaps some rows from $M_{11}, \dots, M_{k+1,1}$. The complement of this minor uses all the columns and some of the rows of

$$\begin{bmatrix} M_{12} & 0 & 0 & \dots & 0 \\ M_{22} & M_{23} & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ M_{k2} & M_{k3} & \cdot & \dots & M_{k,k+1} \\ M_{k+1,2} & M_{k+1,3} & \cdot & \dots & M_{k+1,k+1} \end{bmatrix}.$$

There are $n_1+\dots+n_{k+1}-(n_0-x)$ rows to select from to produce a nonzero minor, and $n_2+\dots+n_{k+1}$ must be used. Consequently if there is to be a term in the expansion of det M not divisible by s, we must have $n_2+\dots+n_{k+1} \leqq n_1+\dots+n_{k+1}-(n_0-x)$. Thus $\delta \leqq n_0-x \leqq n_1$. If there were not δ independent rows in M_{12}, we could cast M_{12} into row Hermite form, and repeat the last argument with a smaller matrix M_{12} having less than δ rows.

Proof of Theorem 3. We first show that $\mu \geqq k\delta$. If A is unimodular this is clear. If not, M partitions into $n_i \times n_j$ blocks, $0 \leqq i,j \leqq k$, with $\delta \leqq n_1, \delta \leqq n_2, \dots, \delta \leqq n_k$. Hence $\mu \geqq k\delta$.

Conversely, we first produce a unimodular matrix of size $(\alpha+\delta)\times(\alpha+\delta)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Let $A=U \operatorname{diag}(I, S)V$, where U and V are unimodular, and S is diagonal with the nonunit invariant factors of A as diagonal elements. (The Smith form of A.) Here S is $\delta \times \delta$. Then let

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & I_\delta \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & S & I_\delta \\ 0 & I_\delta & 0 \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & I_\delta \end{bmatrix}.$$

This matrix is plainly unimodular and has δ more rows and columns than A. This settles the case $k=1$, and $k>1$ is now treated as in the proof of Theorem 1.

5. Complex orthogonal power embeddings

Let $\alpha\times\alpha$ matrix A now have complex entries. The closeness of A to (complex) orthogonality will be measured by the rank of δ of $I-AA^{\mathrm{T}}$. This also is the rank of $I-A^{\mathrm{T}}A$, because: the rank of $I-AA^{\mathrm{T}}$ is α minus the number of elementary divisors of AA^{T} belonging to eigenvalue 1. Since the elementary divisors of a matrix product AB belonging to a nonzero eigenvalue are also those of BA (Flander's Theorem), $A^{\mathrm{T}}A$ must have the same elementary divisors for 1 as AA^{T}. Of course, $\delta=0$ if and only if A is already orthogonal.

Theorem 4. *Let matrix A have complex entries. Then A possesses a complex orthogonal power embedding into M, to exponent k, with M having μ more rows than A, if and only if $\mu\geqq k\delta$, where δ is the orthogonal deficiency of A defined above.*

The proof of sufficiency is entirely analogous to the sufficiency proof in Theorem 2, changing * to $^{\mathrm{T}}$. Only the necessity needs proof. First we treat the case $k=1$. Let

$$M=\begin{bmatrix} A & M_{01} \\ M_{10} & M_{11}\end{bmatrix}.$$

Orthogonality demands that $M_{01}M_{01}^{\mathrm{T}}=I-AA^{\mathrm{T}}$, and hence M_{01} must have rank at least δ. Therefore it has at least δ columns.

The following lemma will be required below.

Lemma 7. *Let S be a $k\times n$ complex matrix with $SS^{\mathrm{T}}=I_k$. Then an $n\times n$ orthogonal matrix O exists with S as the last k rows.*

Proof. Plainly S has rank k, hence $k\leqq n$ and it has a $k\times k$ nonsingular submatrix. Let P be a permutation matrix such that SP has its initial $k\times k$ submatrix nonsingular. Set $SP=[S_1, S_2]$, with S_1 invertible. Now take

$$O=\begin{bmatrix} X & Y \\ S_1 & S_2\end{bmatrix}P^{-1}$$

with $X=-YS_2^{\mathrm{T}}S_1^{-1\mathrm{T}}$. Then $XS_1^{\mathrm{T}}+YS_2^{\mathrm{T}}=0$, for any choice of Y. We require $XX^{\mathrm{T}}+YY^{\mathrm{T}}=I_{n-k}$ and this amounts to

$$Y[S_2^{\mathrm{T}}S_1^{-1\mathrm{T}}S_1^{-1}S_2+I_{n-k}]Y^{\mathrm{T}}=I_{n-k}. \tag{7}$$

Now $I_k=S_1S_1^{\mathrm{T}}+S_2S_2^{\mathrm{T}}=S_1[I_k+(S_1^{-1}S_2)(S_1^{-1}S_2)^{\mathrm{T}}]S_1^{\mathrm{T}}$. Hence -1 is not an eigenvalue of $(S_1^{-1}S_2)(S_1^{-1}S_2)^{\mathrm{T}}$, and therefore not of $(S_1^{-1}S_2)^{\mathrm{T}}(S_1^{-1}S_2)$. Thus $S_2^{\mathrm{T}}S_1^{-1\mathrm{T}}S_1^{-1}S_2+I_{n-k}$ is nonsingular, and hence (7) can be satisfied by some choice of Y. (Note that Y is square.) For this choice of Y, O is orthogonal, and the lemma is proved.

Now we handle the case $k=2$. Rename the blocks in M as

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Then $C^{T}C = I - A^{T}A = Y^{T}\operatorname{diag}(0, I_\delta)Y$ for some invertible Y. Hence

$$(CY^{-1})^{T}(CY^{-1}) = \operatorname{diag}(0, I_\delta).$$

Let S be the $\mu\times\delta$ matrix comprising the last δ columns of CY^{-1}. Then $S^{T}S = I_\delta$, and by the lemma an orthogonal O exists of the form

$$O = \begin{bmatrix} R^{T} \\ S^{T} \end{bmatrix}.$$

Then $OCY^{-1} = \begin{bmatrix} Z & 0 \\ 0 & I_\delta \end{bmatrix}$, for some Z, and so $OC = \begin{bmatrix} ZY_1 \\ Y_2 \end{bmatrix}$, where Y_1, Y_2 arise from a partitioning of Y as $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$. Then

$$C^{T}C = Y_1^{T}Z^{T}ZY_1 + Y_2^{T}Y_2 = (ZY_1)^{T}(ZY_1) + Y^{T}\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix}Y = (ZY_1)^{T}(ZY_1) + C^{T}C.$$

Therefore $(ZY_1)^{T}(ZY_1) = 0$ and $Y_2^{T}Y_2 = C^{T}C$. Moreover Y_2 is $\delta\times\alpha$ with independent rows. Perform an orthogonal similarity on M by $\operatorname{diag}(I_\alpha, O)$. After this similarity, we partition M as

$$M = \begin{bmatrix} A & M_{01} & M_{02} \\ M_{10} & M_{11} & M_{12} \\ M_{20} & M_{21} & M_{22} \end{bmatrix}$$

with square diagonal blocks, $M_{20} = Y_2$ is $\delta\times\alpha$ with independent rows, and $M_{10} = ZY_1$, so that $M_{10}^{T}M_{10} = 0$. If we pass to PMP^{T} with a block diagonal permutation matrix, we perform a permutation similarity on A and arrange that the initial $\delta\times\delta$ block in M_{20} is nonsingular, and still have $M_{10}^{T}M_{10} = 0$.

We proceed to simplify the form of M. Let T be an as yet unspecified $\alpha\times\delta$ matrix. Observe that $M_{10}TT^{T}M_{10}^{T}$ is nilpotent (its square is zero), and therefore $I + M_{10}TT^{T}M_{10}^{T}$ is invertible. Choose a nonsingular X such that

$$X(I + M_{10}TT^{T}M_{10}^{T})X^{T} = I,$$

then set $Y = -XM_{10}T$. Both X and Y depend on T. Now set

$$O = \begin{bmatrix} X & Y \\ T^{T}M_{10}^{T} & I_\delta \end{bmatrix}. \tag{8}$$

This matrix is orthogonal. Applying to M a block diagonal orthogonal similarity by diag (I_α, O), we get in the lower left positions

$$O\begin{bmatrix} M_{10} \\ M_{20} \end{bmatrix} = \begin{bmatrix} XM_{10}(I_\alpha - TM_{20}) \\ M_{20} \end{bmatrix}.$$

We now choose T, therefore also X and Y. Since the leading $\delta\times\delta$ submatrix in M_{20} is nonsingular, we may choose T so that

$$TM_{20} = \begin{bmatrix} I_\delta & \cdot \\ 0 & 0 \end{bmatrix}.$$

For this choice of T, the upper block in (8) has the form $[0, \cdot]$, where the 0 has δ columns. That is, after an orthogonal similarity preserving the structure of M, we may take M in the form

$$M = \begin{bmatrix} A & M_{01} & M_{02} \\ [0, M''_{10}] & M_{11} & M_{12} \\ [M'_{20}, M''_{20}] & M_{21} & M_{22} \end{bmatrix}$$

where M'_{20} is $\delta\times\delta$ and invertible. (The whole purpose of this reduction was to get a nonsingular block in M_{20} beneath a zero block in M_{10}.)

Now invoke the condition $M^2 \supset A^2$. This yields

$$M_{01}[0, M''_{10}] + M_{02}[M'_{20}, M''_{20}] = 0.$$

Because M'_{20} is invertible, we get $M_{02}=0$. And now, because M is orthogonal,

$$[0, M''_{10}]^{\mathrm{T}} M_{12} + [M'_{20}, M''_{20}]^{\mathrm{T}} M_{22} = 0,$$

yielding $M'^{\mathrm{T}}_{20} M_{22}=0$, whence $M_{22}=0$. Also, we now have $AA^{\mathrm{T}} + M_{01}M_{01}^{\mathrm{T}} = I_\alpha$, whence $M_{01}M_{01}^{\mathrm{T}}$ has rank δ, and thus M_{01} has at least δ columns. Hence M has at least 2δ more rows than A.

We have $M_{12}^{\mathrm{T}} M_{12} = I_\delta$, hence an orthogonal O exists with

$$O = \begin{bmatrix} Z \\ M_{12}^T \end{bmatrix},$$

for some Z. Then $OM_{12} = \begin{bmatrix} 0 \\ I_\delta \end{bmatrix}$, and a block diagonal orthogonal similarity of M by diag (I, O, I) preserves the block structure and converts M_{12} to $\begin{bmatrix} 0 \\ I_\delta \end{bmatrix}$. Repartitioning we now get

$$M = \begin{bmatrix} A & M_{01} & M_{02} & 0 \\ [0, M''_{10}] & M_{11} & M_{12} & 0 \\ [0, M''_{20}] & M_{21} & M_{22} & I_\delta \\ [M'_{30}, M''_{30}] & M_{31} & M_{32} & 0 \end{bmatrix}$$

with M'_{30} $\delta\times\delta$ and nonsingular. Orthogonality implies M''_{20}, M_{21}, M_{22} are all zero.

We now continue by induction, analogous to the proof of Theorem 2. For example, $M^3 \supset A^3$ now implies $M_{02}=0$ (using $M^2 \supset A^2$), whence $M_{32}=0$, and a splitting of M_{12} can be obtained, etc. This completes the proof of Theorem 4.

Comment. Each theorem above is of the following type: Given a semigroup G of matrices of specified size $n \times n$, and a fixed matrix A, how large must n be so that M exists in G with $M^i \supset A^i$ for $i=1, \ldots, k$. The same question can be formulated for other semigroups. For example, if G is the full linear group, then M must have at least $k\delta$ more rows than A, where δ is the nullity of A. We omit the proof.

References

[1] J. Bognár, *Indefinite inner product spaces*, Ergebnisse der Mathematik, Band 78, Springer-Verlag (Berlin—Heidelberg—New York, 1974).

[2] E. Egerváry, On the contractive linear transformations of n-dimensional vector space, *Acta Sci. Math.*, **15** (1954), 178—182.

[3] Ky Fan, On systems of linear inequalities, in: *Linear inequalities and related systems* (H. Kuhn and A. W. Tucker editors), Annals of Math. Studies, no. 38 (Princeton, 1956); 99—156.

[4] Ky Fan, Two applications of a consistency condition for systems of linear inequalities, *Linear Algebra and Applications*, **11** (1975), 171—180.

[5] B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space*, North Holland (Amsterdam, 1970).

[6] R. C. Thompson, Principal submatrices IX: Introducing inequalities for singular values, *Linear Algebra and Applications*, **5** (1972), 1—12.

[7] R. C. Thompson, Interlacing inequalities for invariant factors, *Linear Algebra and Applications*, **24** (1970), 1—31.

ALGEBRA INSTITUTE, MATHEMATICS DEPARTMENT
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CALIFORNIA 93106, USA

Acta Sci. Math. 44 (1982), 359—364

Spectral results for some Hausdorff matrices

B. E. RHOADES and N. K. SHARMA

Let $BV[0, 1]$ denote the Banach space of functions f of bounded variation over the interval $[0, 1]$, with $\|f\|=V(f)$, the variation of f over $[0, 1]$. In addition, each f is assumed to satisfy $f(t)=[f(t+0)+f(t-0)]/2$ at each point of discontinuity for $0<t<1$, and $f(0+)=f(0)=0$. $\mathfrak{A}$ will denote the subspace of $BV[0, 1]$ consisting of absolutely continuous functions. Let $\mathscr{H}$ denote the Banach algebra of multiplicative conservative Hausdorff matrices. For $f\in BV[0, 1]$, let H_f denote the Hausdorff matrix corresponding to the moment sequence $\{\mu_n\}$, defined by $\mu_n=\int_0^1 t^n\,df$, $n=0, 1, 2, \ldots,$ and define $T_f(z)=\int_0^1 t^{-1+1/z}\,df$ for each $z\in\bar{D}-\{0\}$, where $D=\{z\colon |z-1/2|<1/2\}$.

The Euler matrix of order q, written (E, q), is the Hausdorff matrix corresponding to

$$\varphi_q(t)=\begin{cases}0, & 0\leqq t<q,\\ 1, & q\leqq t\leqq 1,\end{cases}\qquad 0<q<1.$$

The corresponding function for I, the identity matrix is

$$\varphi_1(t)=\begin{cases}0, & 0\leqq t<1,\\ 1, & t=1.\end{cases}$$

M will denote the Cesaro matrix of order 1, corresponding to $f(t)=t$, $0\leqq t\leqq 1$.

In [11] it was shown that $\mathscr{H}$ can be identified with $BV[0, 1]$, and that $\mathfrak{A}$ can be identified with $L^1[0, 1]$, the Banach space of absolutely integrable functions on $[0, 1]$. In [5] it was shown that $\|H_f\|=V(f)$. For additional properties of Hausdorff matrices the reader may consult [3].

For $f, g\in BV[0, 1]$, define

$$(f*g)(t)=f(t)g(1)-\int_t^1 g(u)\,df(t/u),\quad 0\leqq t\leqq 1. \tag{1}$$

Received August 12, 1981.

The multiplication defined by (1) is a reformulation, and an extension to multiplicative Hausdorff matrices, of the formula developed in [2, p. 196].

Theorem 1. *Let* $f, g \in BV[0, 1]$. *Then*
(a) $H_{f*g} = H_f \cdot H_g$,
(b) $\|f*g\| \leqq \|f\| \, \|g\|$, *and*
(c) $f*g = g*f \in BV[0, 1]$.

Since $f, g \in BV[0, 1]$, there exist moment sequences $\{b_n\}$, $\{c_n\}$ defined by $b_n = \int_0^1 t^n \, dg$, $c_n = \int_0^1 t^n \, df$, $n = 0, 1, 2, \ldots$. From Theorem 210 of [3], the sequence $\{a_n\}$, defined by $a_n = b_n c_n$, is also a moment sequence. Therefore, there exists a mass function $h \in BV[0, 1]$ such that $a_n = \int_0^1 t^n \, dh$. To prove (a), it remains to show that $h = f*g$.

If one defines $b(z) = \int_0^1 t^z \, dg(t)$, $c(z) = \int_0^1 t^z \, df(z)$, then $b(z)$ and $c(z)$ are analytic for $\operatorname{Re} z > 0$ and continuous for $\operatorname{Re} z \geqq 0$. Since $a_n = b_n c_n$ for all nonnegative integers n, it follows that $a(z)$, defined by $a(z) = b(z)c(z)$, has the representation $a(z) = \int_0^1 t^z \, dh(t)$.

By defining $u = e^{-t}$, $v = e^{-s}$, $h(u) = h(1) - A(t)$, $g(u) = g(1) - B(t)$, and $f(u) = f(1) - C(t)$, where A, B, and C are as defined on pages 200—201 of [2], the equation

$$\int_0^1 t^z \, dh(t) = \int_0^1 t^z \, dg(t) \cdot \int_0^1 t^z \, df(t)$$

becomes

$$\int_0^\infty e^{-tz} \, dA(t) = \int_0^\infty e^{-tz} \, dB(t) \cdot \int_0^\infty e^{-tz} \, dC(t).$$

Using the multiplication theorem derived on page 201 of [2], it follows that $\int_0^\infty e^{-zt} \, dA(t) = z \int_{-\infty}^\infty e^{-zt} \, dA(t)$, where

$$A(t) = \int_0^t B(t-s) \, dC(s). \tag{2}$$

Substituting the values of A, B, C, u, v into (2), and noting that $B(t-s) = g(1) - g(u/v)$, yields

$$h(1) - h(u) = \int_1^u [g(1) - g(u/v)](-df(v)) = -g(1) \int_1^u df(v) + \int_1^u g(u/v) \, df(v) =$$

$$= g(1) \int_u^1 df(v) - \int_u^1 g(u/v) \, df(v).$$

Therefore

$$h(u) = h(1)-g(1)[f(1)-f(u)]+\int_u^1 g(u/v)\,df(v) =$$

$$= h(1)-g(1)f(1)+f(u)g(1)-\int_u^1 g(t)\,df(u/t),$$

and (1) is established, provided it can be shown that $h(1)=g(1)f(1)$. But this is easy. Since $a_n=b_nc_n$ for all n, $a_0=b_0c_0$; i.e.,

$$\int_0^1 dh(t) = \int_0^1 dg(t)\cdot\int_0^1 df(t),$$

so that $h(1)-h(0)=\big(g(1)-g(0)\big)\big(f(1)-f(0)\big)$. But $h(0)=g(0)=f(0)=0$, so that $h(1)=g(1)f(1)$.

To prove (b), $\|f*g\|=\|H_{f*g}\|=\|H_f\cdot H_g\|\leqq\|H_f\|\cdot\|H_g\|=\|f\|\cdot\|g\|$.

Since multiplication of Hausdorff matrices is commutative, (c) follows immediately from (a) and (b).

Definition (1) is useful, not only for computing mass functions for products of moment sequences, but also is useful as a tool for computing the spectra for particular Hausdorff matrices, as the following theorem illustrates.

Theorem 2. *The spectrum of* M, $\sigma(M)=\{z:\ |z-1/2|\leqq 1/2\}$.

Suppose there exists a mass function $f\in BV[0,1]$ such that $\big((t-\lambda\varphi_1)*f\big)(t)=$ $=\varphi_1(t)$ for some complex number λ. Then, for $0<t<1$, $tf(1)=t\int_t^1 \big(f(u)/u^2\big)\,du=$ $=\lambda f(t)$. Hence

$$f(1)+\int_t^1 \frac{f(u)}{u^2}\,du = \lambda\frac{f(t)}{t},$$

which implies

$$-\frac{f(t)}{t^2} = \lambda\left[-\frac{f(t)}{t^2}+\frac{f'(t)}{t}\right]$$

a.e. in (0,1). For $\lambda\neq 0$ the above equation takes the form $f'(t)/f(t)=(\lambda-1)/\lambda t$, which has the solution $f(t)=At^{(\lambda-1)/\lambda}$ for some constant A. For $\operatorname{Re}\big((\lambda-1)/\lambda\big)>0$, $f\in BV[0,1]$ and, for $\operatorname{Re}\big((\lambda-1)/\lambda\big)<0$, $f\notin BV[0,1]$. Since $\operatorname{Re}\big((\lambda-1)/\lambda\big)>0$ is equivalent to $\lambda\notin\bar{D}$, and since the spectrum is always closed, the theorem is proved.

Remarks. 1. Theorem 2 is not a new result, but the proof is new. A different proof appears in [4]. Theorem 2 can also be established using the techniques employed in Theorem 4 of [6]. Since M is also a weighted mean method, Theorem 2 is a special case of Theorem 1 of [1].

2. Using the same technique as in Theorem 2, it can be shown that $\sigma(E, q) = \{z: |z| \leqq 1\}$, a result established by a different method in Theorem 3 of [9]. Also, if $f(t)=t^k$, $k>0$, then $\sigma(H_f)=\bar{D}$.

Theorem 3. *Let $f \in BV[0, 1]$ such that $f(t)/t \in L^1[0, 1]$. Then $f(t)*t \in \mathfrak{A}$.*

From (1), $f(t)*t = tf(1)+h(t)$, where $h(t)=t\int_t^1 (f(u)/u^2)\, du$. Then, a.e. on [0, 1],

$$h'(t) = \int_t^1 \frac{f(u)}{u^2}\, du - \frac{f(t)}{t},$$

and

$$\int_0^1 |h'(t)|\, dt \leqq \int_0^1 \int_t^1 \frac{|f(u)|}{u^2}\, du\, dt + \int_0^1 \frac{|f(t)|}{t}\, dt.$$

Interchanging the order of integration in the first integral yields the second integral, so $h' \in L^1[0, 1]$. Thus, $h \in L^1[0, 1]$ and the theorem follows.

It follows from [7] that the set of conservative Hausdorff matrices is a maximal commutative Banach subalgebra of the algebra of conservative matrices. Consequently, the spectrum of any member of $\mathscr{H}$ is determined by the set of multiplicative linear functionals defined in this subalgebra. (See, e.g. [12, p. 264].) The remainder of this paper is devoted to a study of these functionals, and in extending some of the results of [10].

Theorem 4. *Let χ be any multiplicative linear functional defined in $\mathscr{H}$ such that $\chi(M)\neq 0$. Then $\chi(T_f M)=T_f(\chi(M))$ for each $f \in BV[0, 1]$.*

Without loss of generality it may be assumed that f is nondecreasing. Define $f_\delta(t)=0$ for $0 \leqq t < \delta$, $f_\delta(t)=f(t)$ for $\delta \leqq t \leqq 1$. Then $f_\delta \in BV[0, 1]$ and $\|f-f_\delta\| \leqq f(\delta)$. Since $f(t)\to 0$ as $t\to 0$, $\|f-f_\delta\|\to 0$ as $\delta\to 0$. Therefore $\|H_f-H_{f_\delta}\|\to 0$ as $\delta\to 0$. Since $|T_f(z)-T_{f_\delta}(z)| \leqq \|H_f-H_{f_\delta}\|$ for each $z \in \bar{D}-\{0\}$, $T_{f_\delta}(z)\to T_f(z)$ as $\delta\to 0$.

Define $\chi(M)=z$, $z\in\bar{D}-\{0\}$. Since clearly $f(t)/t \in L^1[0, 1]$, $f_\delta(t)*t\in\mathfrak{A}$ by Theorem 3. As in the proof of Corollary 4 of [10], $\chi(MH_{f_\delta})=zT_{f_\delta}(z)$. Since $z\neq 0$, $\chi(H_{f_\delta})=T_{f_\delta}(z)$. Taking the limit as $\delta\to 0$ yields $\chi(H_f)=T_f(z)$, and the proof is finished, since $H_f=T_f(M)$.

Remark 3. In Remark 3 of [10] it was shown that, for $z\in D$, $\chi(M)=z$ implies $\chi(T_f M)=T_f(\chi(M))$. The above theorem has extended this result to $\bar{D}-\{0\}$. That Theorem 4 cannot be extended to $\bar{D}$ will be shown in the example following Remark 4.

Define χ_0 by $\chi_0(H_f)=\lim \mu_n$, where $\mu_n=\int_0^1 t^n\, df$. Then χ_0 is a nonzero multiplicative linear functional on $\mathscr{H}$ with the property that $\chi_0(E, q)=0$ for each $0<q<1$.

Theorem 5. *Let χ be any nonzero multiplicative linear functional on $\mathscr{H}$. Then $\chi(E, q)=0$ for some $0<q<1$ if and only if $\chi=\chi_0$.*

If $\chi=\chi_0$ then clearly $\chi(E, q)=0$. Suppose $\chi(E, q)=0$ for some $0<q<1$. Let q_1 satisfy $0<q_1<q$. Then $(E, q_1)=(E, q)(E, q_1/q)$ so that $\chi(E, q_1)=\chi(E, q)\cdot\chi(E, q_1/q)=0$. Suppose $q<q_2<1$. Since $\lim q_2^n=0$, choose any n such that $q_2^n<q$. Then $\chi(E, q_2^n)=0$, which implies $\chi(E, q_2)=0$.

Now let $f\in BV[0, 1]$ such that $f(1-0)=f(1)$. Consider the function g defined by $g(t, q)=f(qt)*\varphi_q(t)$, $0<q<1$. Using (1) it is easy to verify that $g(t, q)=f(t)$ for $0\leqq t<q$, and $g(t, q)=f(q)$ for $q\leqq t\leqq 1$. Therefore $\|g(t, q)-f(t)\|\to 0$ as $q\to 1$. Since $\chi(E, q)=0$ for each $0<q<1$, $\chi(H_{g(t,q)})=0$ for each $0<q<1$, and hence $\chi(H_f)=0$.

Any $h\in BV[0, 1]$ can be written in the form $h(t)=f(t)+\lambda\varphi_1(t)$, where $\lambda=\lim\int_0^1 t^n\,df$, and $f(t)=h(t)-\lambda\varphi_1(t)$. Then $\chi(H_h)=\lambda$, so that $\chi=\chi_0$.

Theorem 6. *Let $f\in\mathfrak{A}$, a any constant. Then*

$$\sigma(T_f(M)+a(E, q)) = \overline{\{T_f(z)+aq^{-1+1/z}:\ z\in\bar{D}-\{0\}\}}.$$

Let χ be any multiplicative linear functional on $\mathscr{H}$. Suppose $\chi(M)=z$ for $z\in\bar{D}-\{0\}$. From Theorem 4, $\chi(T_f(M)+a(E, q))=T_f(z)+aq^{-1+1/z}\in\sigma(T_f(M)+a(E, q))$. If $\chi(M)=0$ and $\chi(E, q)=r\neq 0$, then there exists a $z\in\bar{D}-\{0\}$ such that $r=q^{-1+1/z}$. Define $\alpha=\alpha(n)=z\log q/(\log q+2n z\pi i)$ for n any integer. Then $q^{-1+1/\alpha}=r$, and $\alpha\to 0$ as $n\to\infty$. Since $f\in\mathfrak{A}$, $T_f(\alpha)\to 0$ as $n\to\infty$. Thus, for all n sufficiently large, $\alpha\in\bar{D}-\{0\}$ and $T_f(\alpha)+aq^{-1+1/\alpha}\in\sigma(T_f(M)+a(E, q))$. The result now follows since the spectrum is closed.

Remark 4. A different proof of Theorem 6, for functions of regular bounded variation, appears in [8].

The following example shows that $\chi(M)=z$, $z=0$ need not imply $\chi(T_f M)=T_f(\chi(M))$. Consider $M+(E, q)\in\mathscr{H}$. From Remark 2, $-1\in\sigma(E, q)$. Define $\{z_n\}$ by $z_n=\log q/(\log q+(2n-1)\pi i)$. Then $z_n\in\bar{D}-\{0\}$ for each n, $z_n\to 0$, and, the multiplicative linear functional on $\mathscr{H}$ defined by $\chi(M)=z_n$ also satisfies $\chi(E, q)=q^{-1+1/z_n}=-1$ for each n. Therefore, by Theorem 6, $z_n-1\in\sigma(M+(E, q))$ for each n. Since the spectrum is closed, $-1\in\sigma(M+(E, q))$. Hence there exists a χ_1 such that $\chi_1(M+(E, q))=-1$. It will now be shown that $\chi_1(M)=0$. Suppose not. Then $\chi_1(M)=z\neq 0$. From Theorem 6, $-1=\chi_1(M+(E, q))=z+q^{-1+1/z}$, or

$q^{-1+1/z} = -(1+z)$ for some $z \in \bar{D} - \{0\}$, which is impossible, since $0 < q < 1$. Therefore $\chi_1(M) = 0$ and $\chi_1(E, q) = -1$.

Thus Theorem 4 cannot be extended to $\bar{D}$. This example also shows that $\chi(M) = 0$ does not imply $\chi(E, q) = 0$ even though the converse is true from Theorem 5.

References

[1] Frank P. Cass and B. E. Rhoades, Mercerian theorems via spectral theory, *Pacific J. Math.*, **73** (1977), 63—71.

[2] H. L. Garabedian, E. Hille and H. S. Wall, Formulation of the Hausdorff inclusion problem, *Duke Math. J.*, **8** (1941), 193—213.

[3] G. H. Hardy, *Divergent Series*, Clarendon Press (Oxford, 1949).

[4] W. A. Hurwitz and L. L. Silverman, On the consistency and equivalence of certain definitions of summability, *Trans. Amer. Math. Soc.*, **18** (1917), 1—20.

[5] A. Jakimovski, B. E. Rhoades and J. Tzimbalario, Hausdorff matrices as bounded operators over l^p, *Math. Z.*, **138** (1974), 173—181.

[6] B. E. Rhoades, Spectra of some Hausdorff operators, *Acta Sci. Math.*, **32** (1971), 91—100.

[7] —, Commutants of some Hausdorff matrices, *Pacific J. Math.*, **42** (1972), 715—719; **49** (1973), 617—619.

[8] W. W. Rogosinski, On Hausdorff's method of summability, *Proc. Cambridge Phil. Soc.*, **38** (1942), 162—192.

[9] N. K. Sharma, Spectra of conservative matrices, *Proc. Amer. Math. Soc.*, **35** (1972), 515—518.

[10] —, Isolated points of the spectra of conservative matrices, *Proc. Amer. Math. Soc.*, **51** (1975), 1—5.

[11] —, Banach algebras of conservative Hausdorff matrices, *Indiana Univ. Math. J.*, **25** (1976), 1035—1038.

[12] A. Wilansky, *Functional Analysis*, Blaisdell (New York—Toronto—London, 1964).

(B. E. R.)
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405, USA

(N. K. S.)
HIGHER INSTITUTE OF ENGINEERING
POST BOX 31930
HOON, LIBYA

Acta Sci. Math. 44 (1982), 365—369

A note on boundedly complete decomposition of a Banach space

P. K. JAIN, KHALIL AHMAD and S. M. MASKEY

1. Introduction. Let E be a Banach space. A sequence (M_i) of subspaces of E is said to be a *decomposition* of E if each $x \in E$ can uniquely be expressed as $x = \sum_{i=1}^{\infty} x_i$, where $x_i \in M_i$ for each i, and convergence is with respect to the norm on E. The uniqueness implies the existence of (not necessarily continuous) associated projections P_i of E onto M_i such that $P_i P_j = \delta_{ij} P_j$, where $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$, and we write $P_i(x) = x_i$. If each P_i is continuous, the decomposition is called a *Schauder decomposition* and we write it as (M_i, P_i). A decomposition (M_i) is called *boundedly complete* if the relation $\sup_{1 \leqq n < \infty} \left\| \sum_{i=1}^{n} x_i \right\| < \infty$ implies that $\sum_{i=1}^{\infty} x_i$ converges, where $x_i \in M_i$ for each i.

The study of decomposition of a Banach space was initiated in the work of Grinblyum [3] and developed further in [2, 9, 10, 11, 12]. The purpose of the present note is to give certain sufficient conditions for a decomposition to be boundedly complete.

2. In this section, we state and prove a lemma, on which we rely heavily when proving our main results.

Lemma. *Let (M_i) be a Schauder decomposition of E. Then the following statements are equivalent:*

(A) *For each number $\lambda > 0$ there exists a number $r_\lambda > 0$ such that*

$$\left\| \sum_{i=1}^{n} x_i \right\| = 1, \quad \left\| \sum_{i=n+1}^{\infty} x_i \right\| \geqq \lambda \quad \textit{imply} \quad \left\| \sum_{i=1}^{\infty} x_i \right\| \geqq 1 + r_\lambda$$

($x_i \in M_i$ for each i).

Received August 20, 1981.

(B) *For every* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that*

$$\left\|\sum_{i=1}^{n} x_i\right\| > 1-\delta, \quad \left\|\sum_{i=1}^{\infty} x_i\right\| = 1 \quad \textit{imply} \quad \left\|\sum_{i=n+1}^{\infty} x_i\right\| \leqq \varepsilon$$

($x_i \in M_i$ for each i).

Proof. (A)⇒(B). Suppose (A) holds and (B) is not true, then there exists an $\varepsilon > 0$ such that for every $\delta > 0$,

$$\left\|\sum_{i=1}^{n} x_i\right\| > 1-\delta, \quad \left\|\sum_{i=1}^{\infty} x_i\right\| = 1, \quad \left\|\sum_{i=n+1}^{\infty} x_i\right\| > \varepsilon \quad (x_i \in M_i).$$

Then, for $\lambda = \frac{\varepsilon}{K}$, where K is a constant appearing in Grinblyum's K-condition for Schauder decomposition (see [8], p. 93), there exists no $r_\lambda > 0$ so as to satisfy (A). Indeed, let $r_\lambda > 0$ be arbitrary, $\delta = r_\lambda/(1+r_\lambda)$ and $y_i = x_i \Big/ \left\|\sum_{j=1}^{n} x_j\right\|$. Then

$$\left\|\sum_{i=1}^{n} y_i\right\| = 1, \quad \left\|\sum_{i=n+1}^{\infty} y_i\right\| \geqq \varepsilon/K \left\|\sum_{j=1}^{\infty} x_j\right\| = \lambda,$$

$$\left\|\sum_{i=1}^{\infty} y_i\right\| = 1 \Big/ \left\|\sum_{j=1}^{n} x_j\right\| < \frac{1}{1-\delta} = 1+r_\lambda.$$

This is a contradiction and hence (A) implies (B).

(B)⇒(A). Assume that (A) is not true, i.e. there exists a $\lambda > 0$ such that for every $r_\lambda > 0$,

$$\left\|\sum_{i=1}^{n} x_i\right\| = 1, \quad \left\|\sum_{i=n+1}^{\infty} x_i\right\| \geqq \lambda, \quad \left\|\sum_{i=1}^{\infty} x_i\right\| < 1+r_\lambda.$$

Then, for $\varepsilon = \lambda(1-\eta)$ with $0 < \eta < 1$ arbitrary, there exists no $\delta > 0$ so as to satisfy (B). Indeed, let $\delta > 0$ be arbitrary with $\delta \leqq \eta$. Let $r_\lambda = \delta/(1-\delta)$ and $y_i = x_i \Big/ \left\|\sum_{j=1}^{\infty} x_j\right\|$. Therefore

$$\left\|\sum_{i=1}^{n} y_i\right\| = 1 \Big/ \left\|\sum_{j=1}^{\infty} x_j\right\| > \frac{1}{1+r_\lambda} = 1-\delta, \quad \left\|\sum_{i=1}^{\infty} y_i\right\| = 1,$$

$$\left\|\sum_{i=n+1}^{\infty} y_i\right\| \geqq \lambda \Big/ \left\|\sum_{j=1}^{\infty} x_j\right\| > \frac{\lambda}{1+r_\lambda} \geqq \varepsilon,$$

which is a contradiction, hence (B) implies (A).

Note. The statements (A) and (B) in the lemma will be referred to as properties A and B, respectively.

3. Main results

Theorem 3.1. *Let (M_i) be a Schauder decomposition of E. If (M_i) satisfies property A (or B), then (M_i) is boundedly complete. The converse may not be true.*

Proof. Suppose $\sup_{1 \leqq n < \infty} \left\| \sum_{i=1}^{n} x_i \right\| = \alpha < \infty$ with $x_i \in M_i$. Let $y_n = \sum_{i=1}^{n} x_i$, choose a sequence (n_k) of positive integers such that $\lim_{k \to \infty} \|y_{n_k}\| = \overline{\lim}_{n \to \infty} \|y_n\| = \beta$ (say). If $\beta = 0$, then $\sum_{i=1}^{\infty} x_i$ converges (to zero). If $\beta \neq 0$, we shall show that (y_{n_k}) is a Cauchy sequence. In fact, otherwise there would exist a $\delta > 0$ and subsequences $(y_{n_{k_j}})$, $(y_{n_{l_j}})$, of (y_{n_k}) with $n_{k_j} > n_{l_j}$ $(j = 1, 2, \ldots)$ such that

$$\|y_{n_{k_j}} - y_{n_{l_j}}\| \geqq \delta \quad (j = 1, 2, \ldots).$$

Then, since

$$\left\| \frac{y_{n_{k_j}} - y_{n_{l_j}}}{\|y_{n_{l_j}}\|} \right\| \geqq \frac{\delta}{\alpha} = \lambda > 0,$$

we would have by property A

$$\left\| \frac{y_{n_{l_j}}}{\|y_{n_{l_j}}\|} + \frac{y_{n_{k_j}} - y_{n_{l_j}}}{\|y_{n_{l_j}}\|} \right\| \geqq 1 + r_\lambda,$$

hence

$$\|y_{n_{k_j}}\| \geqq \|y_{n_{l_j}}\| (1 + r_\lambda).$$

Thus

$$\beta = \lim_{j \to \infty} \|y_{n_{k_j}}\| \geqq \lim_{j \to \infty} \|y_{n_{l_j}}\| (1 + r_\lambda) = \beta(1 + r_\lambda),$$

which is impossible since $\beta \neq 0$. Consequently, (y_{n_k}) is a Cauchy sequence. Hence $\lim_{k \to \infty} y_{n_k} = x \in E$. Therefore, (M_i) being a Schauder decomposition,

$$x = \lim_{k \to \infty} y_{n_k} = \lim_{k \to \infty} \sum_{i=1}^{n_k} x_i = \sum_{i=1}^{\infty} x_i.$$

This shows that $\sum_{i=1}^{\infty} x_i$ converges, whenever

$$\sup_{1 \leqq n < \infty} \left\| \sum_{i=1}^{n} x_i \right\| < \infty.$$

For the converse, consider the following counter-example which would complete the proof of the theorem.

Example 3.2. Let $(\chi, \| \ \|)$ be a Banach space. Define

$$l_1(\chi) = \left\{ (x_i) \colon \ x_i \in \chi, \ \sum_{i=1}^{\infty} \|x_i\| < \infty \right\},$$

the norm on $l_1(\chi)$ being given by

$$\|(x_i)\|^* = \sum_{i=1}^{\infty} \|x_i\|.$$

Further, let us assume the Banach space χ to be such that the topological dual of the space $l_1(\chi)$ is its respective cross dual (see [6], Table 3.29, and [5]). Now, we observe that (N_i) with $N_i=\{\delta_i^{x_i}:\ x_i\in\chi\}$, where $\delta_i^{x_i}$ means the sequence $(0, 0, \ldots, x_i, 0, \ldots)$ i.e. the i-th entry in $\delta_i^{x_i}$ is x_i and all others are zero, forms a Schauder decomposition (see [4], p. 290, and [8], p. 95) of $l_1(\chi)$. Now, we define

$$\bar{N}_1 = \{\delta_1^{\frac{x}{2}}+\delta_2^{\frac{x}{2}}: x\in\chi\},\quad \bar{N}_2 = \{\delta_1^{-\frac{x}{2}}+\delta_2^{\frac{x}{2}}: x\in\chi\},\quad \bar{N}_i = N_i, \quad \text{for} \quad i \neq 1, 2.$$

Then $(\bar{N}_i)$ is a boundedly complete decomposition, but does not satisfy property A.

Remark. Properties A and B are not invariant under an isomorphism of the space E onto another space E_1. Hence they are not isomorphic properties since (N_i) forms a boundedly complete decomposition, equivalent to $(\bar{N}_i)$, of E which satisfies property A.

Definition 3.3. A Schauder decomposition (M_i) is said to be *monotone* if $\left\|\sum_{i=1}^{n} x_i\right\| \leqq \left\|\sum_{i=1}^{n+1} x_i\right\|$, for all n, where x_i is an arbitrary element of M_i.

Definition 3.4. A Banach space is *uniformly convex* if for every $\varepsilon>0$ there exists $\delta>0$ such that whenever $\|x\|, \|y\| \leqq 1$, and $\|x-y\| \geqq \varepsilon$, then $\|(x+y)/2\| \leqq \leqq 1-\delta$.

Now, we give sufficient conditions for a decomposition to satisfy property A (or B).

Theorem 3.5. *If (M_i) is a monotone decomposition of a uniformly convex space E, then (M_i) satisfies property* B *(hence property* A*).*

Proof. Let property B be not true. Then for any given $\varepsilon>0$ and $\delta>0$ (in particular, we choose $\delta>0$ of definition 3.4), there exists a sequence (x_i), $x_i\in M_i$ such that

$$\left\|\sum_{i=1}^{\infty} x_i\right\| > 1-\delta, \quad \left\|\sum_{i=1}^{\infty} x_i\right\| = 1 \quad \text{and} \quad \left\|\sum_{i=n+1}^{\infty} x_i\right\| > \varepsilon.$$

Therefore, monotonicity of (M_i) implies

$$1-\delta < \left\|\sum_{i=1}^{n} x_i\right\| \leqq \left\|\sum_{i=1}^{\infty} x_i\right\| = 1.$$

Let $x=\sum_{i=1}^{n} x_i$ and $y=\sum_{i=1}^{\infty} x_i$. Then $\|x\|, \|y\| \leqq 1$, and

$$\|x-y\| = \left\| \sum_{i=n+1}^{\infty} x_i \right\| > \varepsilon,$$

and so $\|(x+y)/2\| \leqq 1-\delta$. Hence

$$1-\delta \geqq \|(x+y)/2\| = \left\| \sum_{i=1}^{n} x_i + \frac{1}{2} \sum_{i=n+1}^{\infty} x_i \right\| \geqq \left\| \sum_{i=1}^{n} x_i \right\| > 1-\delta,$$

which is a contradiction.

Corollary 3.6. *If (M_i) is a monotone decomposition of a uniformly convex space E, then (M_i) is boundedly complete.*

References

[1] P. Casazza, On a geometric condition related to boundedly complete bases and normal structure in Banach spaces, *Proc. Amer. Math. Soc.*, **36** (1972), 443—447.

[2] W. J. Davis, Schauder decompositions in Banach spaces, *Bull. Amer. Math. Soc.*, **74** (1968), 1083—1085.

[3] M. M. Grinblyum, On the representation of a space of type B in the form of the direct sum of the subspaces, *Dokl. Akad. Nauk SSSR (N. S.)*, **70** (1950), 749—752.

[4] M. Gupta, P. K. Kamthan and K. L. N. Rao, Generalised Köthe sequence spaces and decompositions, *Ann. Mat. Pura Appl.* (4), **113** (1977), 287—301.

[5] M. Gupta, P. K. Kamthan and J. Patterson, Duals of generalised sequence spaces, *J. Math. Anal. Appl.*, **82** (1981), 152—168.

[6] P. K. Kamthan and M. Gupta, *Sequence spaces and series*, Marcel Dekker, Inc. (New York, 1981).

[7] A. R. Lovaglia, Locally uniformly convex Banach spaces, *Trans. Amer. Math. Soc.*, **78** (1955), 225—238.

[8] J. T. Marti, *Introduction to the theory of bases*, Springer-Verlag (Berlin, 1969).

[9] W. H. Ruckle, The infinite sum of closed subspaces of an *F*-space, *Duke. Math. J.*, **31** (1964), 543—554.

[10] B. L. Sanders, On a generalisation of the Schauder basis concept, Dissertation, Florida State University, 1962.

[11] B. L. Sanders, Decompositions and reflexivity in Banach spaces, *Proc. Amer. Math. Soc.*, **16** (1965), 204—208.

[12] B. L. Sanders, On the existence of [Schauder] decompositions in Banach spaces, *Proc. Amer. Math. Soc.*, **16** (1965), 987—990.

[13] I. Singer, *Bases in Banach spaces I*, Springer-Verlag (Berlin, 1970).

(P. K. J. and S. M. M.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF DELHI
DELHI, INDIA

(K. A.)
DEPARTMENT OF MATHEMATICS
JAMIA MILLIA ISLAMIA
NEW DELHI — 110025, INDIA

Bibliographie

Analytical Methods in Probability Theory. Proceedings, Oberwolfach, Germany, 1980. Edited by D. Dugué, E. Lukács and V. K. Rohatgi (Lecture Notes in Mathematics, 861), X+183 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

This volume contains most of the papers read at the conference "Analytical Methods in Probability Theory", Oberwolfach, June 9—14, 1980. Characterizations of distributions (unimodal, Poisson, Gamma, unimodality of infinitely divisible distributions) are investigated in nine papers. Ten papers deal with asymptotic properties of stochastic processes and their applications in statistics (tests for exponentiality and independence, local limit theorem for sample extremes, local time and invariance, rate of convergence in the central limit theorem, weak convergence of point processes).

Lajos Horváth (Szeged)

S. Burris—H. P. Sankappanavar, A Course in Universal Algebra (Graduate Texts in Mathematics, vol. 78), XVI+276 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1981.

The book gives a high-level course in modern methods and results of universal algebra. It is divided into 5 numbered chapters, contains a large bibliography, author and subject index and a list of some open problems and applications.

Chapter I contains the neccessary definitions and theorems from lattice theory. Chapter II (The elements of universal algebra) describes the most important concepts and notions. Here we can find the most commonly used methods to construct algebraic structures. Chapter III discusses several topics, e. g., how universal algebra can be related to combinatorics and to regular languages. Chapter IV — starting from the notion of a Boolean algebra — presents results of the last years. It deals with primal algebras, quasi-primal algebras, functionally complete algebras and some interesting classes of varieties. Chapter V discusses connections between universal algebra and model theory. Each chapter is divided into sections, and each section ends with references and exercises.

The book is very elegantly and clearly written and can be recommended for all students and mathematicians interested in universal algebra.

László Hannák (Budapest)

W. K. Bühler, Gauss. A Biographical Study, VIII+208 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

Im Jahre 1977 wurde der 200. Geburtstag von Gauss gewürdigt; 1980 sein 125. Todestag. Die Lebenszeit des Princeps mathematicorum liegt also nicht allzu weit zurück; dennoch wird es mit jedem Jahr schwieriger, eine Gauss-Biography zu schreiben. Wie der Verfasser selbst schreibt, enthält sein Werk kaum Dinge, die dem Spezialisten nicht bereits bekannt sein dürften: fast alle

Informationen können in bereits veröffentlichten Quellen gefunden werden; darüberhinaus stehen die Gesammelten Abhandlungen von Gauss einigermaßen vollständig zur Verfügung.

Ziel des Buches ist eine Darstellung des Lebens und Wirkens von Carl Friedrich Gauss, der — sogar mit den Maßstäben unserer schnellebigen Zeit — in einer Periode außerordentlicher politischer und sozialer Veränderungen lebte. Der Verfasser vertritt die Auffassung, daß Gauss nicht in die intellektuelle Szene dieser Zeit "paßte" und daß sein Lebensweg ein ungewöhnlich geradliniger gewesen ist.

Mit seinem Werk wendet sich der Autor an Mathematiker und andere Wissenschaftler unserer Zeit, wobei aber die Wissenschaftshistoriker und die Psychologen, welche "die Skalps großer Männer sammeln" ausgeschlossen werden. Der Verfasser ist — wie er in seinem Vorwort selbst zugibt — bei seiner Darstellungsweise gleichzeitig bescheiden und unbescheiden. Bescheiden ist er deswegen, weil er nicht den Versuch unternimmt, das "Leben von Gauss" in definitiver Weise niederzuschreiben. Als unbescheiden, wenn nicht gar grandios, sieht er seinen Versuch an, aus dem Leben und Wirken von Gauss diejenigen Gesichtspunkte wenigstens teilweise hervorzuheben, die einerseits von zeitgemäßem Interesse sind und sich andererseits an einen nicht primär historische motivierten Leser werden.

Dieses Vorhaben des Verfassers ist als gelungen einzuschätzen; sein Werk ist eine anspruchsvolle Lektüre, die zu lesen es sich lohnt.

Abschließend ein Überblick über den inhaltlichen Aufbau des Buches: Kindheit und Jugend, 1777—1795; Die zeitgenössische politische und soziale Lage; Studentenzeit in Göttingen, 1795—1798; Zahlentheoretische Arbeiten; Der Einfluss von Gauss' arithmetischen Arbeiten; Rückkehr nach Braunschweig. Dissertation. Die Umlaufbahn der Ceres; Heirat, spätere Jahre in Braunschweig; Die politische Szene in Deutschland, 1789—1848; Familienleben, Umzug nach Göttingen; Tod von Johanna und zweite Ehe. Die ersten Jahre als Professor in Göttingen; Der Stil von Gauss; Astronomische Arbeiten. Elliptische Funktionen; Modulare Formen. Die Hypergeometrische Funktion; Geodäsie und Geometry; Der Ruf nach Berlin und Gauss' soziale Rolle. Das Ende der zweiten Ehe; Physik; Persönliche Interessen nach dem Tode der zweiten Frau; die Göttinger Sieben; Die Methode der kleinsten Quadrate; Numerische Arbeiten; Die Jahre 1838—1855; Gauss' Tod.

In drei Appendizes werden die Organisation der Gesammelten Werke von Gauss, ein Überblick über die Sekundärliteratur und ein Index der Arbeiten von Gauss gegeben. Jedes Kapitel ist von zahlreichen erläuternden Fußnoten begleitet, die zu einem noch besseren Verständnis beitragen.

Manfred Stern (Halle)

J. T. Cannon—S. Dostrovsky, The Evolution of Dynamics: Vibration Theory from 1687 to 1742 (Studies in the History of Mathematics and Physical Sciences, 6), IX+184 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1981.

Very few works have produced such a tremendous effect on the development of theoretical physics and matematics as Newton's Principia published in 1687. The book deals with the history of vibration theory during the half century that followed this date. Reading the book one has the feeling that the originality and brilliancy of ideas of Newton, Taylor, Euler and the Bernoulli family can really be appreciated only by studying their original works and correspondence. The reader obtains considerable help to do this from the authors who make the contents of these works readily accessible and make clear the historical connections among many of the pertinent ideas and concepts such as isochronism, simultaneous crossing of the axis, pendulum condition. These concepts were used for the study of the problems of floating bodies, hanging chains, resonating beams, the vibrating ring and the vibrating string pertaining to dynamics in many degrees of freedom.

The nicely presented book is concluded by the facsimile of Daniel Bernoulli's papers on the hanging chain and the linked pendulum with translations.

L. Hatvani (Szeged)

J. Carr, Applications of Centre Manifold Theory (Applied Mathematical Sciences, 35), XII+142 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1981.

What is the centre manifold theory in the title of the book? In the theory of differential equations the case when the linear part of the right-hand side has pure imaginary eigenvalues and the real parts of the remaining spectrum points are bounded above by a negative constant is called critical. If the proper subspace corresponding to the pure imaginary part of the spectrum is of finite dimension the centre manifold theory guarantees the existence of a locally invariant manifold for the system which is tangent to this finite dimensional subspace at the origin. This gives a machinery to reduce the dimension of the system under investigation in the critical case.

The book is based on a series of lectures given in the Lefshetz Center for Dynamical Systems in the Division of Applied Mathematics at Brown University during the academic year 1978—79. This is what may cause that, as in a good lecture, this introductory book first gives a full account of the key ideas and methods in a simpler but interesting in itself case, illustrates them by examples, and then proceeds toward the necessary generalizations. In the first two chapters the basic theorems of the theory are formulated, illuminated and proved. Chapters 3—5 are devoted to applications such as the Hopf bifurcation theory and its application to a singular perturbation problem which arises in biology. In Chapter 6 the theory is extended to a class of infinite dimensional problems. Finally, the use of the extension in partial differential equations is illustrated by means of some simple examples.

We can recommend these notes both to users of mathematics and to mathematicians interested in bifurcation and stability theory and their applications.

L. Hatvani (Szeged)

The Chern Symposium 1979. Proceedings of the International Symposium on Differential Geometry in Honor of S.-S. Chern, held in Berkeley, California, June 1979. Edited by W. Y. Hsiang, S. Kobayashi, I. M. Singer, A. Weinstein, I. Wolf, H.-H. Wu, VII+259 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1980.

This book contains 12 articles reflecting the connection of modern differential geometry with other subjects in mathematics and physics.

Some of the papers are intended to exploit the influence of such differential geometric notions as fiber bundle, characteristic classes etc, which originated from the work of S.-S. Chern, to modern physics, especially to general relativity and gauge theory (M. P. Atiyah: Real and Complex Geometry in Four Dimensions; Raoul Bott: Equivariant Morse Theory and the Yang—Mills Equations on Riemann Surfaces; Chen Ning Yang: Fibre Bundles and the Physics of the Magnetic Monopole; Shing-Tung Yau: The Total Mass and the Topology of an Asymptotically Flat Space-Time.) Three papers are devoted to the study of problems in algebraic geometry. (Eugenio Calabi: Isometric Families of Kahler Structures; Mark Green and Philip Griffiths: Two Applications of Algebraic Geometry to Entire Holomorphic Mappings; F. Hirzebruch: The Canonical Map for Certain Hilbert

Modular Surfaces.) Two papers are discussing classical problems in geometry (Nicolaas H. Kuiper: Tight Embeddings and Maps. Submanifolds of Geometrical Class Three in E^N; Robert Osserman: Minimal surfaces, Gauss Maps, Total Curvature, Eigenvalue Estimates, and Stability). Two papers are devoted to the interaction of differential geometry and functional analysis (J. Moser: Geometry of Quadrics and Spectral Theory; Louis Nirenberg: Remarks on Nonlinear Problems). One paper is related to homology theory (Wu Wen-tsün: de Rham-Sullivan Measure of Spaces and Its Calculability).

The reader can trace in these papers the influence of differential geometric ideas in the development of mathematical and physical sciences. The book is worth studying for everyone working in differential geometry or in related topics.

Péter T. Nagy (Szeged)

M. Csörgő—P. Révész, Strong Approximations in Probability and Statistics, 284 pages, Akadémiai Kiadó, Budapest and Academic Press, New York—San Francisco—London, 1981.

When writing about a book of which the brother of the reviewer is one of the authors, the topic of which is in a broad sense almost identical to and constitutes a starting point for a large part of recent research of the reviewer and, thirdly, in which the reviewer's earlier work is cited to such a large extent (perhaps because of the first connection?), there is an unavoidable danger of lack of objectivity. Exposing himself freely to such a charge, but greatly economizing with praising adjectives, the reviewer feels that this is a significant monograph with effects that will influence much wider circles than that what is broadly termed nowadays as the "Hungarian school" of probability and statistics.

Chapter 2 (Strong approximations of partial sums of independent identically distributed random variables by Wiener processes) and Chapter 4 (Strong approximation of empirical processes by Gaussian processes) deal with the basic issues indicated in their titles. Both chapters describe the history of the corresponding developments, concentrating on the evolution of the involved ideas. The first breakthrough of this evolution, after Strassen's and Kiefer's work with the Skorohod embedding technique, was the two authors' invention of the quantile transformation technique, formerly used also by Bártfai, by which they disproved a conjecture of Strassen concerning the approximation of partial sums and were able to extend Kiefer's approximation of the empirical process. Up to these points both chapters are complete. Then the descriptions of the final breakthrough follow, respectively, the celebrated Komlós—Major—Tusnády approximations. The proofs of the latter results are not given in detail, save for a partial result for the empirical process. The fourth chapter also containes the authors' approximation for the quantile process.

When the final approximation results by Komlós, Major and Tusnády emerged, another basic question of the possible applications, other than convergence rates for the distribution of functionals and the law of the iterated logarithm, became important. These depend on what the corresponding properties, to be inherited by strong invariance, of the approximating (uni and multivariate) Wiener, Brownian bridge and Kiefer processes are within the best rates of approximation. This question lead to the two authors' important series of "How big and small" papers concerning the almost sure size of the increments of these processes. A comprehensive account of these results, at least in the univariate and in certain bivariate cases, is given in Chapter 1 together with new constructions of the corresponding processes. This 67 pages first chapter is in itself a significant contribution to the literature of stochastics.

Chapters 3 and 5 utilize the results of Chapter 1 for partial sum and empirical processes, respectively, via the strong approximations in Chapters 2 and 4. The sixth chapter deals with applications of the basic approximation results for further empirical processes such as density and regression estimators and the empirical characteristic function, while Chapter 7 with those for randomly indiced partial sum and empirical processes. Each chapter ends with a section of supplementary remarks concerning various side developments.

It is no risk to predict that this book will be a frequent reference in research papers for a longer time to come. It is also appropriate as a textbook for a one-year graduate course.

Sándor Csörgő (Szeged)

L. R. Foulds, Optimization Techniques. An introduction (Undergraduate Texts in Mathematics), XI+502 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

This book provides an introduction to the main optimization techniques which are presently in use. Chapter 1 contains a brief introduction to the basic terminology of the theory of optimization. Chapter 2 is concerned with linear programming (LP). After describing the basic problem of LP, the author presents different forms of the simplex algorithm such as the primal simplex method, the "big M" method and the two-phase method. Relationships are introduced between dual and primal problems. We can find a short part about the postoptimal analysis. Some special LP problems (transportation problem, assignment problem) are studied and different solutions of them are described. Chapter 3 introduces the techniques developed to solve large scale LP problems (Revised Simplex Method, Dantzig—Wolfe decomposition, dual-primal algorithm). There is a short overview of the parametric programming. Chapter 4 discusses the integer programming problem and its solutions by different methods as enumerative techniques, and cutting plane methods. New formulations and models are presented. Chapter 5 is concerned with network optimization, emphasizing the shortest path problem, the minimal spanning tree problem and different flow problems. Chapter 6 contains a short review about some known dynamic programming problems and their solutions. Chapter 7 — classical optimization — is an introduction to nonlinear programming, which takes place in Chapter 8. This part of the book consists of several methods to solve unconstrained nonlinear problems, and among the methods concerned with the solution of constrained problems, the author discusses some efficient strategies such as the gradient projection method, the penalty function method and linear approximations.

For unexperienced readers, there is an Appendix with linear algebra and basic calculus. The book contains a large number of excercises and their solutions.

This work evolved out from the experience of teaching the material to finishing undergraduates and beginning graduates. So all chapters contain a "real-life" problem to solve, which is modified depending on the aims of the chapter. All problems are examined not from theoretical but from practical point of view, so algorithms are simply explained with illustrative numerical examples. This book is easy to read, it is a very useful material for education.

G. Galambos (Szeged)

From A to Z. Proceedings of a symposium in honour of A. C. Zaanen. Edited by C. B. Huijsmans, M. A. Kaashoek, W. A. J. Luxemburg, W. K. Vietsch (Mathematical Centre Tracts, 149) VII+130 pages, Mathematisch Centrum, Amsterdam, 1982.

The symposium was held on July 5—6, 1982, at the University of Leiden, on the occasion of Professor A. C. Zaanen's retirement.

There were three invited lectures (H. H. Schaefer: Some recent results on positive groups and semi-groups; F. Smithies: The background to Cauchy's definition of the integral; B. Sz.-Nagy:

Some lattice properties of the space L^2), and nine lectures by former Ph. D. students of Prof. Zaanen (J. L. Grobler: Orlicz spaces — a survey of certain aspects; C. B. Huijsmans: Orthomorphisms; K. de Jonge: Embeddings of Riesz subspaces with an application to mathematical statistics; M. A. Kaashoek: Symmetrizable operators and minimal factorisation; W. A. J. Luxemburg: Orthomorphisms and the Radon-Nikodym theorem revisited; P. Maritz: On the Radon-Nikodym theorem; B. de Pagter: Duality in the theory of Banach lattices; A. R. Schep: Integral operators; W. K. Vietsch: Compact operators). A Curriculum vitae of A. C. Zaanen, a list of his publications, and a list of all his former Ph. D. students close the handy and beautifully presented small volume.

Professor Zaanen's infatigable and successful mathematical activity, which, we hope, he will be able to continue by the same energy and precision through many years to come (he is actually working on his book *Riesz spaces. II*) has won for him a high appreciation from the mathematical community. All who had the chance to meet him or to contact him at least by correspondence, enjoyed his attractive and suggestive, but utmost modest personality: no wonder that his excellent former pupils, although spread out by now on four continents, are still so much attached to him. His fine personality is mirrored also by his nice photo at the beginning of this volume.

As a survey of various modern areas of mathematical analysis, this book also appeals to, and will be an interesting and useful reading for, mathematicians far beyond the circle of those closely attached to the person of Professor Zaanen.

Béla Sz.-Nagy (Szeged)

Geometry and Differential Geometry. Proceedings, Haifa, Israel 1979. Edited by R. Artzy and I. Vaisman (Lecture Notes in Mathematics 792) VI+444 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1980.

The present book contains the text of the lectures presented at a Conference on Geometry and Differential Geometry, which was held at the University of Haifa, Israel, March 18—23, 1979. The conference was divided into two sections, namely Geometry and Differential Geometry, and in both sections the subject matters covered a broad range, and many of the aspects of modern research in the field were discussed. Altogether 42 papers are published in the book, thus it is impossible to give a complete list of the authors' works. The reader can find several papers from the fields of synthetic and axiomatic geometry, theory of matroids, Riemannian manifods, Lie algebras and Lie groups, conformal geometry, foliations and fibrations etc.

The book is structurally well arranged and the single papers are of high-level affording good reading.

Z. I. Szabó (Szeged)

B. D. Hassard—N. D. Kazarinoff—Y.-H. Wan, Theory and Applications of Hopf Bifurcation (London Mathematical Society Lecture Note Series, 41), VIII+311 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1981.

The adventurous history of James Watt's centrifugal governor provides an interesting example of Hopf bifurcation. This device was invented by Watt in about 1782 for the purpose of controlling steam engines. Centrifugal governors worked well for roughly a century after their introduction. Later on, however, when a great number of machines were constructed having different physical parameters, curious behaviour was observed in a large percentage of them: some governors would "hunt" for the right operating speed before settling down; others would oscillate, never attaining a constant angular velocity. To eliminate "hunting", viscous dampers (dashpots) were added on to the governors. It was proved that there is a minimum amount of damping which must be present in the system in order to guarantee stability. As the damping is decreased past a critical value, the stationary solution becomes unstable; but a stable periodic solution appears and takes its place as the long-term behaviour of the system.

In general, the "Hopf Bifurcation" describes this phenomenon: the birth of a family of oscillations as a controlling parameter is varied.

The authors first present the mathematical theory of Hopf bifurcation (Ch. 1). The reader can find the bifurcation formulae for computing the form of the oscillations, their amplitudes, their periods, and their stability or lack of it. A "Recipe-Summary" also is given making easier to apply the formulae in a particular case. The amount of algebraic manipulation, necessary for analytical evaluation of bifurcation formulae, increases rapidly with the number of state coordinates. For the avoidance of this difficulty a numerical algorithm is presented in FORTRAN codes for the required calculation (Ch. 3); moveorer, a set of computer programs is provided on microfiche, which enables anyone with moderate FORTRAN ability to run Hopf bifurcation computations.

The applications illustrated by the numerous examples worked out are divided into groups according as the basic model is an ordinary differential equation (Ch. 2), differential-difference and integro-differential equation (Ch. 4), or partial differential equation (Ch. 5). In these chapters the reader can find interesting examples such as Watt's steam-engine governor, the Hodgkin—Huxley model nerve conduction equations, the Brusselator, and Dowell's panel flutter model.

Summing up, this book will be very useful for all scientists in whose fields bifurcation phenomena occur.

L. Hatvani (Szeged)

T. Hida, Brownian Motion (Applications of Mathematics, Vol. 11), XVI+325 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1980.

Mentioning the name of Brownian motion provokes a lively reaction from most mathematicians, and their attitude reflects mostly memories of their former professor of probability. But before one would decide to punish a sometime teacher by neglecting his subject, it is worth of weighing who else will be hit by this decision. We warn: The Brownian motion is not a privilege of probabilists! On the contrary, by its internal wealth it forms an everywhere dense set in mathematics. Even if it is not a primary object of someone's investigation, the connaissance of Brownian motion can offer a different, often equivalent view-point to the matter or at least it can provide the scientist with essential, non-trivial examples.

The author of the present book had K. Yosida and K. Ito as professors, so it wasn't difficult for him to get engaged to Brownian motion. But for him the fundamental structure is less a random process with its sample paths (as for Ito) or a semi-group of operators on a Banach space (as for Yosida) but rather a measure on a space of generalized functions. Taking advantage of the possibilities offered by Gaussian measures, he presents essentially a Hilbert space theory of Brownian motion and of white noise. Fourier transforms, orthogonal expansions and infinite-dimensional transformation groups play key role in the book. At the same time as he investigates one of the infinitely many aspects of Brownian motion the author presents a new, interesting functional analytic theory.

Besides being indispensable for specialists of stochastic processes, the book is recommended to any mathematician feeling enough force to get acquined with a new non-trivial field. The reviewer would like to call special attention of collegues bearing a deep sympathy towards Hilbert space methods. Brownian motion can not only give much more motivation to mathematicians as say quantum mechanics, but being a natural, internal creature of mathematics with its infinitely many interrelations, it can promote cohesion and mutual understanding in the divergent family of mathematicians of today.

D. Vermes (Szeged)

J. Kevorkian—J. D. Cole, Perturbation Methods in Applied Mathematics (Applied Mathematical Sciences, 34), X+558 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1981.

This is a revised and updated version, including a substantial portion of new material, of J. D. Cole's book under the same title published by Ginn-Blaisdell in 1968.

Perturbation methods are very often used by applied mathematicians and physicians when attempting to solve physical problems. In essence, a perturbation procedure consists of constructing the solution for a problem involving a small parameter ε, either in the differential equation or the boundary conditions or both, when the solution is known for the limiting case $\varepsilon=0$.

Traditional regular perturbation problems, for example, the problem of calculating the perturbed eigenvalues and eigenfunctions of a selfadjoint differential operator, are omitted; they are discussed in most texts on differential equations. Rather, the present book concentrates on the so-called singular perturbation problems. Such problems are, among others, layer type problems and cumulative perturbation problems.

In a layer type problem the small parameter multiplies a term in the differential equation which becomes large in a thin layer. Often this is the highest derivative in the differential equation and the $\varepsilon=0$ approximation is therefore governed by a lower order equation which cannot satisfy all the initial or boundary conditions prescribed. In a cumulative perturbation problem the small parameter multiplies a term which never becomes large. However, its cumulative effect becomes important for large values of the independent variable.

The book consists of five chapters, and ends with Bibliography, Author and Subject Index. No particular attempt is made to have a complete list of references.

Chapter 1 contains some background on asymptotic expansions. Chapter 2 gives a deeper exposition of limit process expansions through a sequence of examples for ordinary differential equations. Chapter 3 is devoted to cumulative perturbation problems using the so-called multiple variable expansion procedure. Applications to nonlinear oscillations, celestial mechanics are discussed in detail. In Chapter 4 the procedures of the preceding chapters are applied to partial differential equations. Finally, Chapter 5 deals with typical examples from fiuid mechanics: linearized and transonic aerodynamics, shallow water theory etc.

The book is written from the point of view of an applied mathematician. Sometimes less attention is paid to mathematical rigour. Instead, physical reasoning is often used as an aid to understanding a problem and to formulating an appropriate approximation procedure.

To sum up, this well-written book contains a unified account on the methods of current researches in Perturbation Theory. The authors present the "state of the art" in a systematic manner. The book under review will certainly serve as a textbook both for advanced undergraduate students and for practicing scientists dealing with truly complicated problems of mathematical physics.

F. Móricz (Szeged)

Oldrich Kowalski, Generalized Symmetric Spaces (Lecture Notes in Mathematics, 805) XII+187 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1980.

In the last 15 years the theory of "s-manifolds", which is a very succesful generalization of symmetric spaces, has been largely built up. This type of spaces is defined similarly as symmetric spaces by postulating the existence of a geodesic symmetry at each point of the space without requiring the involutiveness of these symmetries. The author of this Lecture Note is one of the developers of the theory of these spaces. He presents in this book a self-contained treatment of the geometric theory of Riemannian and affine "s-manifolds". Must of the results contained in this lecture note were avaible earlier only in journal articles, and some of them are published here the first time. The reader is supposed to be familiar with the basic notions and methods of modern differential geometry and Lie group theory.

The list of the chapter headings gives a glimpse of the content: Generalized symmetric Riemannian spaces, Reductive spaces, Differentiable s-manifolds, Locally regular s-manifolds, Operations with s-manifolds, Distinguished s-structures on generalized symmetric spaces, The classification of generalized symmetric Riemannian spaces of low dimensions, The classification of generalized affine symmetric spaces in low dimensions.

The Lecture Note is very clearly and well written. It is recommended to everyone interested in the geometric theory of spaces with a transformation group.

Péter T. Nagy (Szeged)

Measure Theory. Proceedings, Oberwolfach 1979. Edited by D. Közlow (Lecture Notes in Mathematics, 794), XV+573 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1980.

Judging from its proceedings, this must have been a significant conference on measure theory. There are 10 papers on what is classified as general measure theory, 8 on measurable selections and 4 on liftings. Two papers deal with differentiation of measures and integrals and 5 are on vector and group valued measures and their probabilistic applications. Stochastic analysis and various abstract probabilistic topics are the theme for 8 more articles, L^p-spaces and related topics are dealt with in 2 papers. Integral representations and transforms figure in 3 papers and there are 4 papers classified miscellaneous. This comes to altogether 46 articles, and the collection is closed by a record of the problem session of the conference with 8 problems from 7 authors.

Sándor Csörgő (Szeged)

Richard M. Meyer, Essential Mathematics for Applied Fields (Universitext), XVI+555 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1979.

The purpose of this work is to provide a wide spectrum of "essential" mathematics for workers in the variety of applied fields. Much of the material covered here either too widely scattered or too advanced as presented in the literature to those who need it. The treatment in this book requires only the calculus through differential equation. There is no need for measure theoretic background.

This bulky volume consists of twenty sections, arranged into six units. The first unit comprises Basic Real Analysis beginning with the notions of sets, (single and multiple) sequences, series and functions, and ending with some Abelian and Tauberian theorems. The second unit is devoted to (the 1- and n-dimensional) Riemann—Stieltjes Integration. The third unit treats Finite Calculus, i.e., the theory of finite differences and difference equations. The fourth unit contains Basic Complex Analysis, including Laurent's theorem and expansion, and residue theorem. The fifth unit gives an account of Applied Linear Algebra, among others, of the generalized inverse and characteristic roots. Finally, the sixth unit collects Miscellaneous things such as convex sets and functions, max-min problems and some basic inequalities.

Each unit develops its topic rigorously based upon material previously established. Throughout the text are found solved Examples and Exercises requiring solution, both being essential parts of the development. Complete hints or answers are provided for the exercises. There are References to additional and related material, too.

This self-contained textbook is warmly recommended to everyone who is going to get acquainted with the background of applied mathematics.

F. Móricz (Szeged)

Physics in One Dimension. Proceedings, Fribourg 1980. Edited by J. Bernasconi, and T. Schneider (Springer Series in Solid-State Sciences, Vol. 23) IX+368 pages. Springer-Verlag, New York—Heidelberg—Berlin, 1981.

Mathematics in one dimension is a commonly accepted thing, but physics is essentially three-dimensional. Nevertheless in certain cases some problems of 3D mathematical physics can be reduced to one dimension and then solved in a simpler way. Ideas of this type of transformations are described by the introductory lecture of this volume, which contains the Proceedings of an International Conference on Physics in One Dimension, held in Fribourg in 1980. The rapid development of experimental physics of long polymer chains in the last decade has shown however that one dimensional physics itself is more than speculation.

This book is a clever mixing of recent theoretical and experimental works in this field. Part II is devoted to the theory of soliton type excitations, Part III deals with magnetic properties. Several points of views on solitons in the simplest polymer, the polyacetilene are outlined in Part IV. The last three parts contain papers on metallic conductivity, disorder, localization, excitons and other interesting questions relevant to one dimensional systems. The wide range and the great number of the authors yield a good overview on the whole subject which will attract an even growing interest in the near future.

M. G. Benedict (Szeged)

J. H. Pollard, A Handbook of Numerical and Statistical Techniques with Examples Mainly from the Life Sciences, XII+349 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1979.

This is a well and carefully compiled easy-to-use handbook of basic numerical and statistical techniques designed to aid the practising statistician in solving day-to-day problems on a small programmable desk calculator or mini-computer.

Part I (Basic numerical techniques) consists of seven chapters. After listing the modest amount of prerequisites from calculus and linear algebra and the usual techniques of reducing truncation and round-off errors, it describes methods for smoothing data, for numerical integration and differentation, interpolation and some related topics.

Part II (Basic statistical techniques) consists of equally seven chapters but covers, of course, nearly the 60% of the text. After supplying the elements of probabilistic and statistical theory very briefly, the essential characteristics of the most commonly used continuous and discrete distributions are given. Following then a short chapter on fitting a Pearson curve, two rich chapters come on hypothesis testing and estimation. The final one is on random numbers, data transformation and on the simplest techniques with randomly of deterministically censored samples. Part III is fully devoted to the method of least squares in (simple and multiple) linear, curvilinear and non-linear regression.

There are 15 tables, 105 references and good author and subject indices. Each of the amazingly large number of techniques covered is demonstrated by at least one numerical example according to the title and there is a wealth of good computational and programming advices. All in all, this book is way above the usual level of the many books on "basic statistics" published nowadays. Experimental scientists, particularly those in life sciences, will find it very useful.

Sándor Csörgő (Szeged)

Probability in Banach Spaces III. Proceedings, Medford, USA, 1980. Edited by. A. Beck (Lecture Notes in Mathematics, 860), VI+329 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1981.

Four papers of this volume give surveys on the developments of probability theory in infinite-dimensional vector spaces in the last two-three years. Generalized domains of attraction, empirical processes, central limit theorems and inequalites in Banach spaces are studied. Twenty research papers present various directions of this subject, well-known results of finite- dimensional spaces are extended to the case of Banach spaces (martingales, strong, and weak laws of large numbers, laws of iterated logarithm, selfdecomposability and stability of measures, stability of linear and quadratic forms). Properties of Gaussian measures in function spaces are also investigated.

Lajos Horváth (Szeged)

Robert D. Richtmyer, Principles of Advanced Mathematical Physics, Vol. II. (Texts and Monographs in Physics) XI+322 pages, Springer-Verlag,New York—Heidelberg—Berlin, 1981.

There are many good books on the mathematical methods of physics, and it seems quite difficult to select just the principles out of this enormous subject, but this book of Richtmyer has successfully solved this problem, enabling the student or the non-expert research worker to grasp some modern views on contemporary mathematical physics.

This second volume begins with chapter 18, which together with the subsequent four, cover the traditional material of applications of group theory in physics. On the other hand, the next three chapters approach continuous groups from a less conventional viewpoint. It is the concept of manifold that leads us to Lie groups and then through chapters 26—28 to the apparatus of general relativity.

A topic, which has never been considered so far in books of this kind, is outlined in the last three chapters. Starting from the problem of hydrodynamical turbulence, the first steps of the theory of bifurcations, attractors and chaos are presented.

The main virtue of this book is that it gives much information about newly developed mathematical concepts on a language, adopted usually in books on physics. There are a lot of examples and several problems, challenging the conscientious reader.

M. G. Benedict (Szeged)

Charles E. Rickart, Natural Function Algebras (Universitext) XIII+240 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1979.

The term "function algebra" in the title refers to a uniformly closed algebra of complex-valued continuous functions defined on a compact Hausdorff space. Such Banach algebras have been intensively studied recently. Since the most important examples of these algebras are built up from analytic functions, the majority of the papers in question was earlier dominated by problems of analyticity. The present author is concerned, however, with another facet of the subject based on the observation that very general algebras of continuous functions may exhibit certain properties that are reminiscent of analyticity. The most striking one of them is a local maximum principle proved by Hugo Rossi in 1960. This deep result plays a key role throughout the discussion. Although the main body of the material presented here was published in a series of papers during the last 10 or 15 years, the book contains numerous improvements on old results as well as unpublished results.

To be more definite, the classical holomorphy theory, based on the n-dimensional complex space C^n, is ultimately determined by the algebra $\mathscr{P}$ of all polynomials in C^n. In the abstract situation the space C^n is replaced by a Hausdorff space Σ and the algebra $\mathscr{P}$ by a given algebra $\mathscr{A}$ of continuous complex-valued functions on Σ. In order to obtain interesting results, one must impose some rather general conditions on the pair $[\Sigma, \mathscr{A}]$. In the first place, $\mathscr{A}$ is assumed to determine the topology of Σ in the sense that the given topology is equivalent to the weakest one under which the elements of $\mathscr{A}$ are continuous. Secondly, it is assumed that every homomorphism of $\mathscr{A}$ onto the complex field C, which is continuous relative to the compact-open topology in $\mathscr{A}$, is a point evaluation in the space Σ. Then $\mathscr{A}$ determines an "$\mathscr{A}$-holomorphy" theory based on Σ roughly analogous to the way $\mathscr{P}$ determines the classical theory.

This fairly general setting makes it possible to establish a variety of results, many of which are full or partial generalizations of results in Several Complex Variables. The $\mathscr{A}$-holomorphy theory might also be considered as an approach to the Infinite Dimensional Holomorphy theory. The latter subject, which already has an extensive literature, involves the study of functions on infinite dimensional linear topological spaces.

The material is divided into fourteen chapters: 1. The category of pairs, 2. Convexity and naturality, 3. The Šilov boundary and local maximum principle, 4. Holomorphic functions, 5. Maximum properties of holomorphic functions, 6. Subharmonic functions, 7. Varieties, 8. Holomorphic and subharmonic convexity, 9. $[\Sigma, \mathscr{A}]$-domains, 10. Holomorphic extensions of $[\Sigma, \mathscr{A}]$-domains, 11. Holomorphy theory for dual pairs of vector spaces, 12. $\langle E, F\rangle$-domains of holomorphy, 13. Dual pair theory applied to $[\Sigma, \mathscr{A}]$-domains, 14. Holomorphic extensions of Δ-domains. The text is supplemented with Bibliography containing 75 items, Index of Symbols, and General Index.

This self-contained textbook is addressed to graduate students and warmly recommended to everyone who wants to keep pace with up-to-date developments in Holomorphy Theory and Banach Algebras.

F. Móricz (Szeged)

Séminaire de Probabilités XIV, 1978/79. Edité par J. Azéma et M. Yor (Lecture Notes in Mathematics, 784), VII+546 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1980.

This is the first volume of this series of seminar notes after that the old Strasbourg seminar has been "decentralised to Paris". Nevertheless the authors and the names in the titles of the papers are basically the same as before, completed by some distinguished overseas visitors (10 of the 49 papers are in English). Thus, with a very few exceptions, the topic continues to be what is broadly described the general theory of stochastic processes. Although, naturally, some of the shorter communications might prove important, it is interesting to note that if we take out the 14 longer or middle size papers then the average lenght of the remaining 35 notes is 6 typed pages. Many of them (commenting, correcting, or exposing already existing developments by other or the same authors) could have never been published elsewhere. This is of course the very feature of these seminar notes which reflect a vivid but somewhat closed research activity.

Sándor Csörgő (Szeged)

Solitons. Edited by R. K. Bullough and P. J. Caudrey (Topics in Current Physics, Vol. 17) XVIII + 389 pages, Springer-Verlag, New York—Heidelberg—Berlin, 1980.

Frontiers of the two closely related sciences, mathematics and physics have moved away quite a distance from each other, after the mathematical foundations of quantum mechanics have been established.

Now again some important questions of physics have become of fundamental interest in mathematics and vice versa. Namely the problem of solving nonlinear partial differential equations has arisen in many fields of physics and this "white land" then attracted the mathematicians too.

The soliton is a very special solution of a nonlinear wave equation, it propagates e.g. like a single, bell shaped wavelet, preserving its shape during the propagation and even after "interaction" with similar "objects". It behaves much like a physical particle, and that is why it has attracted most attention among other possible solutions. Some of them were known for a long time, but a systematic study has become possible only after the discovery of the so-called inverse scattering method in the late sixties and early seventies. This method reduces the nonlinear problem to the solution of a system of linear integral equations by a process resembling the Fourier transformation. Besides the editors' review on the history and the present status of this field, the volume contains 11 invited studies on the subject, all of them written by the establishers of the theory of solitons. Almost the half of the contributions deals with the inverse scattering method, namely the works of Newell, Zakharov, Wadati, Faddeev, Calgero and Degasperis. A direct method, transforming the nonlinear equation into a bilinear one is outlined by Hirota. Physical aspects are treated in the contributions of Lamb and Maclaughlin and also of Bullough, Caudrey and Gibbs. The famous Toda lattice is dealt by Toda. Novikov investigates equations with periodic boundary conditions. Possible quantization procedures are treated by Luther.

Each work is clearly written, emphasizing the underlying principles and giving many examples.

This book is highly recomended to all who work in the field of nonlinear analysis and partial differential equations, and without doubt, it must be found in every physics library.

M. G. Benedict (Szeged)

Superspace and Supergravity. Proceedings of the Nuffield Workshop, Cambridge, June 16—July 12, 1980. Edited by S. W. Hawking and M. Rocek, XII + 527 pages, Cambridge University Press, London—New York—New Rochelle—Melbourne—Sydney, 1981.

10—15 years ago the considerable part of physicists was convinced on the "in principle" impossiblity of the unified field theory. But in the last decade the research has been boomed in this direction. In contrast with the classical unified theories of Weyl, Einstein, Cartan, Calusa, Klein and others the new theories are not restricted only to the gravitation and electromagnetism and are quantized. The possibility of building of unified theories of this type turned to be promising by the large development of mathematical tools as differential geometry and topology, functional analysis and representation theory.

The new unifying theory bears the name of supergravity theory: the classical space-time of relativity theory has been extended by commuting and anticommuting variables to a higher dimensional manifold, the so called superspace. The symmetry groups have a graded structure corresponding to the bundle structure of the superspace, these are the "supersymmetries", their infinitesimal versions are the "superalgebras".

The present book contains the proceedings of the Workshop on Supergravity held in the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge from

16 June to 12 July, 1980, and supported by the Nuffield Foundation. The aim of this meeting was to give a survey on the present state of the very recent and rapidly developing supergravity theory.

This book is a collection of lectures devided into six parts: Introduction to supergravity, Quantization, Extended supergravity, N = 8 supergravity, Kähler spaces and supersymmetry, Other aspects of supergravity.

The supergravity theory is far from being worked out. One of the outstanding problems at the present time is "The construction of a complete formulation of extended supergravity with auxiliary fields, preferably in superspace" formulated by the aditors in the introduction.

The book gives for the reader a very recent survey on the present state of the very interesting subject of supergravity theory.

Péter T. Nagy (Szeged)

Michio Suzuki, Group Theory I (Grundlehren der mathematischen Wissenschaften, 247), XIV + 434 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982.

The theory of finite groups developed extremely rapidly during the past twenty five years or so, reaching its zenith in 1981, when the full classification of finite simple groups was completed. The book under review is the first part of a two-volume work first published in Japanese by Iwanami Shoten in 1977—78, and now translated into English. For the translation several mistakes of the original version were corrected, and, more importantly, to reflect the progress achieved in the meantime, a few paragraphs were added to the survey of finite simple groups in Chapter 3, and a few items were added to the bibliography as well.

According to the Preface, one of the main aims of the author was to present an introduction to the theory of finite simple groups. Of course, taking into account the tremendous diversity of the branches of group theory involved, it is understandable that even such important topics as permutation groups, or representation theory could not be discussed in detail.

Volume I contains three of the six chapters constituting the book. In Chapter 1 the basic ideas of group theory are discussed, including permutation groups, operator groups, semidirect products, and general linear groups. Chapter 2 is devoted to the most fundamental theorems and methods of group theory, for example, some results on p-groups, Sylow's Theorems, Schreier's Refinement Theorem, the Krull—Remak—Schmidt Theorem, the fundamental theorem on finitely generated abelian groups, Schur's Lemma, cohomology theory, the Schur-Zassenhaus Theorem, and wreath products. Chapter 3 starts the discussion of more specific branches of group theory with considering several particular classes of groups, namely torsion-free abelian groups, symmetric and alternating groups, linear groups, and Coxeter groups. The volume ends with a survey of finite simple groups, and the proof of Dickson's theorem on the subgroups of 2-dimensional special linear groups over finite fields.

As a rule, each section is followed by a collection of interesting exercises, most of them supplemented with a "Hint" to a solution. The exercises mainly serve to introduce the reader to important concepts and theorems not discussed in the text. However, the book can also be read without solving a single exercise, as no reference is made in the text to results of earlier exercises.

The book is written in a very clear style. The only prerequisite for its reading is some basic knowledge in linear algebra (matrices, determinants) and elementary number theory. This excellent book is warmly recommended to students intending to specialize in group theory, as well as to other non-specialists who are interested in the recent advances of group theory.

Ágnes Szendrei (Szeged)

W. Törnig, Numerische Mathematik für Ingenieure und Physiker, Band 2: Eigenwertprobleme und numerische Methoden der Analysis, XIII+350 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1979.

The development of large-scale computers have formed a basis for algorithmic constructions and extensive mathematical experiments in many areas of science and technology, thereby attracting a new generation of scientists to problems of numerical mathematics. This is the second volume of a textbook. The first volume published also in 1979 contains the following three parts. Part I: Auxiliary results, Computation of zeros of a function; Part II: Solution of linear systems of equations; Part III: Solution of nonlinear systems of equations.

This volume consists of four parts. Part IV: The eigenvalue problem for matrices. It contains, among others, the iteration procedure of Mises, the inverse iteration, the algorithms of Jacobi and Givens, and the LR algorithm. Part V: Interpolation, approximation and numerical integration. It treats the usual interpolating polynomials, approximation by orthogonal polynomials as well as the cubic spline interpolation, the well-known quadrature and some cubature formulae, the Romberg integration procedure etc. Part VI: Numerical solution of ordinary differential equations. It deals with methods for initial value problems (mainly one-step methods, in particular Runge-Kutta ones), for boundary and eigenvalue problems (difference methods, variational and Ritz methods). The notions of consistency and convergence are discussed in detail,but multi-step methods (in particular, predictor-corrector ones) are not presented. Part VII: Numerical solution of partial differential equations. One section contains the method of finite differences for the numerical solution of initial value and initial-boundary value problems of hyperbolic and parabolic differential equations. Another section is devoted to hyperbolic systems of first order and the last section to the boundary value problem of elliptic differential equations of second order, including the method of differences, the Ritz method and the method of finite elements.

There are examples throughout the text. Each section ends with exercises and some of them with a FORTRAN programme. The reader can find references to additional and related material, too.

This accurately written textbook is warmly recommended to every undergraduate student in applied mathematics who is going to acquire a firm basis of Numerical Analysis.

F. Móricz (Szeged)

G. Whyburn—E. Duda, Dynamic topology (Undergraduate Texts in Mathematics), Springer-Verlag, New York—Heidelberg—Berlin, 1979.

Everybody who learned mathematics is convinced that the knowledge acquired by means of active learning is the deepest and most fruitful. This can be carried out by developing one's own proofs for the theorems. Both learning and teaching a subject in this way are conditioned upon reducing to individual steps of the proofs. In this book the reader finds an excellent realization of this method in topology. As we can learn from J. L. Kelly's foreword, G. Whyburn was a master of this manner of theaching. The book is based on a set of his notes, which have been completed and arranged by E. Duda. Each of the short sections, which can serve as the subject-matter of one lesson, consists of a preparing part including the necessary definitions, exercises and, finally, their solutions. The problems in the first sections are rather simple but later they become more complicated. The first exercies of the book is: "Prove that the union of the elements of any countable collection of countable sets results in a countable set", and the last one is proving the Jordan Curve Theorem. The materials of the sections are collected from the field of dynamic topology, developed origonally by G. Whyburn. It concerns those topological objects which center around the function concept. The book is concluded by Whyburn's article entitled "Dynamic Topology" which appeared in 1970 in the American Mathematical Monthly.

J. L. Kelley writes in the Foreword: "Constructing a proof for a good known theorem is next best to finding and proving a good theorem. It gives one hope of eventually creating mathematics". Therefore, we recommend this book to every student and teacher interested in analysis, especially in topology.

L. Hatvani (Szeged)

Livres reçus par la rédaction

S. S. Abhyankar, Weighted expansions for canonical desingularization (Lecture Notes in Mathematics, Vol. 910), VII+236 pages, Springer-Verlag, Berlin—Heidelberg— NewYork, 1982. — DM 25,—.

Analytic Theory of Continued Fractions. Proceedings, Loen, Norway 1981. Edited by W. B. Jones, W. J. Thron, and H. Waadeland (Lecture Notes in Mathematics, Vol. 932), VI+240 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 28,—.

D. M. Arnold, Finite rank torsion free Abelian groups and rings (Lecture Notes in Mathematics, Vol. 931), VII+191 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 24,—.

Automorphic Forms, Representation Theory and Arithmetic. Papers presented at the Bombay Colloquium 1979 (Tata Institute Studies in Mathematics), VII+355 pages, Tata Institute of Fundamental Research, Bombay—Springer-Verlag, Berlin—Heidelberg—New York, 1981. — DM 39,—.

J. Back—D. J. Newman, Complex analysis (Undergraduate Texts in Mathematics), X+244 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 54,—.

P. Berthelot—L. Breen—W. Messing, Théorie de Dieudonné cristalline II (Lecture Notes in Mathematics, Vol. 930), XI+261 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 33,50.

J. M. Bismut—L. Gross—K. Krickeberg, Ecole d'été de Probabilités de Saint-Flour X—1980. Edité par P. L. Hennequin (Lecture Notes in Mathematics, Vol. 929), X+313 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 39,—.

P. M. C. de Boer, Price effects in input-output relations: A theoretical and empirical study for the Netherlands 1949—1967 (Lecture Notes in Economics and Mathematical Systems, Vol. 201), X+140 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 24,—.

R. Bott—L. W. Tu, Differential forms in algebraic topology (Graduate Texts in Mathematics, Vol. 82), XIV+331 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 74,—.

Brauer Groups in Ring Theory and Algebraic Geometry. Proceedings, University of Antwerp, Belgium, August 17—28, 1981. Edited by F. van Oystaeyen and A. Verschoren (Lecture Notes in Mathematics, Vol. 918), VIII+300 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 34,50.

Categorical Aspects of Topology and Analysis. Proceedings of an International Conference held at Carleton University, Ottawa, August 11—15, 1981. Edited by B. Banaschewski (Lecture Notes in Mathematics, Vol. 915), XI+385 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 39,—.

K. L. Chung, Lectures from Markov processes to Brownian motion (Grundlehren der mathematischen Wissenschaften, Band 249), VIII+239 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 79,—.

E. Combet, Intégrales exponentielles: Développements asymptotiques, propriétés lagrangiennes (Lecture Notes in Mathematics, Vol. 937), VIII+114 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 19,80.

Convex Analysis and Optimization. Edited by J.-P. Aubin and R. B. Vinter (Research Notes in Mathematics, 57), VI+210 pages, Pitman Advanced Publ., Boston—London—Melbourne, 1982. — £9.95.

I. P. Cornfeld—S. V. Fomin, Ergodic theory (Grundlehren der mathematischen Wissenschaften, Band 245), X+486 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 118,—.

B. Dacorogna, Weak continuity and weak lower semicontinuity of non-linear functionals (Lecture Notes in Mathematics, Vol. 922), V+120 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 18,—.

J. Dassow, Completeness problems in the structural theory of automata (Mathematical Research — Mathematische Forschung, 7), 148 pages, Akademie-Verlag, Berlin, 1981. — 20,— M.

Differential Geometric Methods in Mathematical Physics. Proceedings of an International Conference held at the Technical University of Clausthal, FRG, July 23—25, 1980. Edited by H. D. Doebner, J. I. Anderson, H. R. Petry (Lecture Notes in Mathematics, Vol. 905), VI+309 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 34,50.

K. Donner, Extension of positive operators and Korovkin theorems (Lecture Notes in Mathematics, Vol. 904), XII+181 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 21,50.

E. B. Dynkin, Markov processes and related problems of analysis. Selected papers, VI+312 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1982. — £ 14.00.

T. J. Enright, Lectures on representations of complex semi-simple Lie groups (Tata Institute Lectures on Mathematics), III+91 pages, Tata Institute of Fundamental Research, Bombay—Springer-Verlag, Berlin—Heidelberg—New York, 1981. — DM 16,—.

Evaluating Mathematical Programming Techniques. Proceedings of a Conference held at the National Bureau of Standards Boulder, Colorado, January 5—6, 1981. Edited by J. M. Mulvey (Lecture Notes in Economics and Mathematical Systems, Vol. 199), XI+379 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 52,—.

Y. Z. Flicker, The trace formula and base change for GL(3) (Lecture Notes in Mathematics, Vol. 927), XII+219 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 28,—.

Functional Analysis in Markov Processes. Proceedings of the International Workshop held at Katata, Japan, August 21—26, 1981 and of the International Conference held at Kyoto, Japan, August 27—29, 1981. Edited by M. Fukushima (Lecture Notes in Mathematics, Vol. 923), V+307 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 34,50.

A. Gardiner, Infinite processes background to analysis, IX+306 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 74,—.

Geometric Techniques in Gauge Theories. Proceedings of the Fifth Scheveningen Conference on Differential Equations, The Netherlands, August 23—28, 1981. Edited by R. Martini and E. M. de Jager (Lecture Notes in Mathematics, Vol. 926), IX+219 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM. 25,—.

K. U. Grusa, Zweidimensionale, interpolierende Lg-Splines und ihre Anwendungen (Lecture Notes in Mathematics, Vol. 916), VIII+238 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 25,—.

Harmonic Analysis. Proceedings of a Conference held at the University of Minnesota, Minneapolis, April 20—30, 1981. Edited by F. Ricci and G. Weiss (Lecture Notes in Mathematics, Vol. 908), V+325 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — 34,50.

L. K. Hua, Introduction to number theory, XVIII+572 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 96,—.

B. Huppert—N. Blackburn, Finite groups II—III (Grundlehren der mathematischen Wissenschaften, Band 242—243), XIII+531 and IX+454 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 138,— and DM 128,—.

Integer Programming and Related Areas. A classified bibliography 1978—1981. Edited by R. von Randow (Lecture Notes in Economics and Mathematical Systems, Vol. 197), XIV+338 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 52,—.

K. Ireland—M. Rosen, A classical introduction to modern number theory (Graduate Texts in Mathematics, Vol. 84), XIII+341 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 72,—.

O. G. Jorsboe—L. Mejlbro, The Carleson-Hunt theorem on Fourier series (Lecture Notes in Mathematics, Vol. 911), IV+123 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 18,—.

S. M. Khaleelulla, Counterexamples in topological vector spaces (Lecture Notes in Mathematics, Vol. 936), XXI+179 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 24,—.

M. Klemm, Symmetrien von Ornamenten und Kristallen (Hochschultext), VII+214 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 36,—.

A. I. Kostrikin, Introduction to algebra (Universitext), XIII+575 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 69,—.

A. Legrand, Homotopie des espaces de sections (Lecture Notes in Mathematics, Vol. 941), VII+132 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 19,80.

J. H. van Lint, Introduction to coding theory (Graduate Texts in Mathematics, Vol. 86), IX+171 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 58,—.

G. E. Martin, The foundations of geometry and the non-Euclidean plane (Undergraduate Texts in Mathematics), XVI+509 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 58,—.

————, **Transformation geometry.** An introduction to symmetry (Undergraduate Texts in Mathematics), XII+237 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982, — DM 67,—.

J. Martinet, Singularities of smooth functions and maps (London Mathematical Society Lecture Note Series, 58), XIV+256 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1982. — £ 12.50.

Martingale Theory in Harmonic Analysis and Banach Spaces. Proceedings of the NSF—CBMS Conference held at the Cleveland State University, Cleveland, Ohio, July 13—17, 1981. Edited by J.-A. Chao and W. A. Woyczynski (Lecture Notes in Mathematics, Vol. 939), VIII+21 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 28,—.

Mathematical Models as a Tool for the Social Sciences. Edited by B. J. West, VII+120 pages, Gordon and Breach, New York—London—Paris, 1980. — $ 29.25.

E. E. Moise, Introductory problem courses in analysis and topology (Universitext), VII+94 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 29,—.

M. A. Naimark—A. I. Stern, Theory of group representations (Grundlehren der mathematischen Wissenschaften, Band 246), IX+568 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 140,—.

M. Nermuth, Information structures in economics studies in the theory of markets with imperfect information (Lecture Notes in Economics and Mathematical Systems, Vol. 196), VIII+236 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 34,—.

Number Theory. Proceedings of the Third Matscience Conference held at Mysore, India, June 3—6, 1981. Edited by K. Alladi (Lecture Notes in Mathematics, Vol. 938), IX+177 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 24,—.

Numerical Analysis. Proceedings of the Third IIMAS Workshop held at Cocoyoc, Mexico, January 1981. Edited by J. P. Hennart (Lecture Notes in Mathematics, Vol. 909), VII+247 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 29,—.

Numerical Analysis. Proceedings of the 9th Biennial Conference held at Dundee, Scotland, June 23—26, 1981. Edited by G. A. Watson (Lecture Notes in Mathematics, Vol. 912), XIV+245 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 29,—.

Probability Measures on Groups. Proceedings of the Sixth Conference held at Oberwolfach, June 28—July 4, 1981. Edited by H. Heyer (Lecture Notes in Mathematics, Vol. 928), X+477 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 55,—.

A. G. Ramm, Iterative methods for calculating static fields and wave scattering by small bodies, XII+122 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 48,—.

R. B. Reisel, Elementary theory of metric spaces. A course in constructing mathematical proofs (Universitext), XI+120 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 24,—.

The Riemann Problem, Complete Integrability and Artihmetic Applications. Proceedings of a Seminar held at the Institut des Hautes Etudes Scientifiques, Bures-sur Yvette, France, and at Columbia University, New York, USA 1979—80. Edited by D. Chudnovsky and G. Chudnovsky (Lecture Notes in Mathematics, Vol. 925), VI+373 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 39,—.

G. F. Roach, Green's functions, VII+325 pages, Cambridge University Press, Cambridge—London—New York—Melbourne, 1982. — £7.95.

D. J. S. Robinson, A course in the theory of groups (Graduate Texts in Mathematics, Vol. 80), XVII+481 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 98, —.

M. Sakai, Quadrature domains (Lecture Notes in Mathematics, Vol. 934), IV+133 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 19,80.

P. Schenzel, Dualisierende Komplexe in der lokalen Algebra und Buchsbaum-Ringe (Lecture Notes in Mathematics, Vol. 907), VII+161 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 21,50.

Z. Semadeni, Schauder bases in Banach spaces of continuous functions (Lecture Notes in Mathematics, Vol. 918), V+136 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 18,—.

Séminaire d'Algèbre Paul Dubreil et Marie-Paule Malliavin, Proceedings, Paris 1981 (34ème Année). Edité par M.-P. Malliavin (Lecture Notes in Mathematics, Vol. 924), VI+461 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 48,50.

Séminaire de Probabilités XVI 1980/81. Edité par J. Azéma et M. York (Lecture Notes in Mathematics, Vol. 920), V+622 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 70,—.

Séminaire de Probabilité XVI 1980/81. Supplement: Géometrie Differentielle Stochastique. Edité par J. Azéma et M. Yor (Lecture Notes in Mathematics, Vol. 921), III+285 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 29,—.

Séminaire de Théorie du Potentiel Paris, No. 6. Directeurs: M. Brelot, G. Choquet et J. Deny. Rédacteurs: F. Hirsch et G. Mokobodzki (Lecture Notes in Mathematics, Vol. 906), IV+328 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 34,50.

Séminaire Pierre Lelong—Henri Skoda (Analyse) Années 1980/81 et Colloque de Wimereux, Mai 1981. "Les fonctions plurisousharmoniques en dimension finie ou infinie", organisé an l'honneur de Pierre Lelong. Edité par P. Lelong et H. Skoda (Lecture Notes in Mathematics, Vol. 919), VII+386 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982, — DM 39,—.

J. G. Simmonds, A brief on tensor analysis (Undergraduate Texts in Mathematics), XI+92 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 45,—.

R. Sot, Simple morphisms in algebraic geometry (Lecture Notes in Mathematics, Vol. 935), IV+146 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 19.80.

O. Tammi, Extremum problems for bounded univalent functions II (Lecture Notes in Mathematics, Vol. 913), VI+168 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 21,50.

M. L. Warschauer, The Witt group of degree k maps and asymmetric inner product spaces (Lecture Notes in Mathematics, Vol. 914), IV+269 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 29,—.

L. C. Washington, Introduction to cyclotomic fields (Graduate Texts in Mathematics, Vol. 83), XI+389 pages, Springer-Verlag, Berlin—Heidelberg—New York, 1982. — DM 96,—.

M. J. Wygodski, Höhere Mathematik griffbereit, 832 Seiten, Akademie-Verlag, Berlin, 1982. — 24,80 M.

TOMUS XLIV — 1982 — 44. KÖTET

Bibliographie

INDEX

ACTA SCIENTIARUM MATHEMATICARUM

SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1

On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (1061 Budapest, I., Fő utca 32)

ISSN 0324-6523 Acta Univ. Szeged
ISSN 0001-6969 Acta Sci. Math.

INDEX: 26024

82-2672 — Szegedi Nyomda — F. v.: Dobó József igazgató

Felelős szerkesztő és kiadó: Leindler László
A kézirat a nyomdába érkezett: 1982. június 2
Megjelenés: 1983. január 28

Példányszám: 1000. Terjedelem: 18 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint