ACTA MATHEMATICA SPALATENSIA
Vol.2 (2022) 1–28
https://doi.org/10.32817/ams.2.1

Received: 09 Feb 2021
Accepted: 29 Mar 2021

Šare's algebraic systems

Mario Essert, Darko Žubrinić

Abstract

We study algebraic systems M_Γ of free semigroup structure, where Γ is a well ordered finite alphabet, discovered in 1970s within the Theory of Electric Circuits by Miro Šare, and finding recent applications in Multivalued Logic, as well as in Computational Linguistics. We provide three simple axioms (reversion axiom (1) and two compression axioms (2) and (3)), which generate the corresponding equivalence relation between words. We also introduce a class of incompressible words, as well as the quotient Šare system $\widetilde{M_\Gamma}$. The main result is contained in Theorem 1, announced by Šare without proof, which characterizes the equivalence of two words by means of Šare sums. The proof is constructive. We describe an algorithm for compression of words, study homomorphisms between quotient Šare systems for various alphabets Γ (Theorem 4), and introduce two natural Šare categories $\check{\mathbf{S}}\mathbf{a}(M)$ and $\check{\mathbf{S}}\mathbf{a}(\widetilde{M})$. Quotient Šare systems are regular semigroups, but not inverse semigroups.

Keywords: Šare algebraic systems or M-systems, jorbs, free semigroups over alphabets, Šare's sum, compression of jorbs, regular semigroups, Šare's categories
2010 Math. Subj. Class.: 20M05, 08A50

1 Introduction

In this paper we describe a class of algebraic systems M_Γ (depending on a well ordered alphabet Γ), introduced in 1970s by Miro Šare

(1918–2005) in [6, 7]. Here are a few examples of well ordered alphabets stemming from Arithmetic, Linguistics, and Electrical Engineering: (a) $\Gamma = \{0, 1, \dots, 9\}$, ($b$) $\Gamma =$ the set of letters of the Latin alphabet, (c) $\Gamma = \{a, b, c\}$, with $a < b < c$, and $C = aa$, $R = bb$ and $L = cc$, where C, R, and L stand for capacitance, resistance and inductance, respectively. For motivations originating from Electrical Engineering, and in particular, for compression axioms (2) and (3) provided below (relevant for Šare's algebraic system M_Γ introduced in Definition 1), see the Introduction in [1]. This enables to *qualitatively* represent electrical networks using words with an even number of letters from $\Gamma = \{a, b, c\}$. Two letters are needed for the representation of C, R and L since each of the corresponding electrical elements (capacitor, resistor, and inductor) has two nodes. Applications in Computational Linguistics and in Multivalued Logic can be seen in the rest of the indicated paper.

The main result of this paper is stated in Theorem 1. It characterizes the equivalence $abcd \sim ad$ in terms of a suitable alternating sum introduced by Šare, which is easy to compute. Theorem 1 was announced by Miro Šare in [8, Theorem 1, case 1.9], but without proof. The proof provided here is constructive, in the sense that if the Šare sufficient condition in Theorem 1 is true, then we not only have equivalence, but we know which admissible substitutions (based on compression/decompression axioms (2) and (3) below) one has to apply in order to achieve it.

Theorem 1 enables us to characterize incompressible words in M_Γ; see Theorem 3. Incompresssibility of a word is defined in terms of minimizing its length among all the words equivalent to it; see Definition 5. In Proposition 4 we show that the Šare quotient semigroups $\widetilde{M_\Gamma}$ do not belong to the class of inverse semigroups (a classical treatise on inverse semigroups is, e.g., [5]). We also provide a compression algorithm.

The paper is organized as follows. In Section 2 we introduce basic definitions and establish some auxilliary results. In Section 3 we define the Šare sums and prove our main result stated in Theorem 1. Section 4 discusses the problem of maximal compression of jorbs. In Section 5 we introduce four Šare categories. In the Appendix we complete the proof of Theorem 1.

2 Basic definitions and auxilliary results

By a *well ordered alphabet* (also called *alphabet*, for short) we mean any well ordered nonempty set, denoted by $(\Gamma, \leq)$. The elements of Γ are called *letters*. For simplicity, we assume that the alphabet Γ is finite.

Definition 1. *By M_Γ, we denote the set of all words $w = \alpha_1 \dots \alpha_{2n}$, consisting of an* even *number of letters $\alpha_k \in \Gamma$ (in arbitrary order) from the alphabet Γ, where $k = 1, \dots, 2n$.*

Note that, according to this definition, the alphabet Γ is clearly not contained in M_Γ, and moreover, these two sets are disjoint.

We introduce the usual binary operation of *concatenation* of words in M_Γ, which is obviously associative. We denote it by $x \cdot y$ or just by xy, for any two words x and y in M_Γ. Thus, M_Γ becomes a free semigroup of rank n, where $n = |\Gamma|$. This semigroup (satisfying the accompanying axioms (1), (2) and (3)) was introduced in this generality by Šare in [6] (in the context of the two-letter alphabet) already in 1970 in [7].

For any given letter $\alpha \in \Gamma$ and for any positive integer k, we define α^k as concatenation of k copies of α. The expression α^k can be contained in a larger word $w \in M_\Gamma$, so that k does not have to be even. We can analogously define w^k for any word $w \in M_\Gamma$. The *length* of any word $w = x_1 \dots x_n \in M_\Gamma$ (where $x_1, \dots, x_n \in \Gamma$), denoted by $\ell(w)$, is $\ell(w) = n$.

We shall occasionally need the set words consisting of k letters,

$$\Gamma_k = \{\alpha_1 \cdots \alpha_k : \alpha_1, \dots, \alpha_k \in \Gamma\},$$

where k is a positive integer, not necessarily even. Clearly,

$$M_\Gamma = \bigcup_{k=1}^{\infty} \Gamma_{2k}.$$

It will be also useful to consider the semigroup of all words (not necessarily of even length): $W_\Gamma = \bigcup_{k=1}^{\infty} \Gamma_k$. (The set W_Γ is often denoted by Γ^+ in the literature.) Here, $M_\Gamma \subset W_\Gamma$, and moreover, M_Γ is a subsemigroup of W_Γ with respect to concatenation. Also, M_Γ is a *language* over the prescribed alphabet Γ; see e.g. [2, 3, 4] for a more detailed information.

Any k-word $w_k \in \Gamma_k$ is said to be a *subword* of a given word $w \in M_\Gamma$, if either $w_k = w$, or w can be obtained from w_k by concatenating on the left of w_k, or on the right, or both. By saying that x is a *k-word*, we mean that its length is any positive integer k, not necessarily even.

The elements of the set Γ_2 consisting of the product of any two letters are called the *atoms* (or *generators*) of the Šare system M_Γ. In other words, the set of atoms in M_Γ is

$$\Gamma_2 = \{\alpha_i\alpha_j : \alpha_i, \alpha_j \in \Gamma\}.$$

Each word $w \in M_\Gamma$ can be obtained as the product of atoms. However, the atoms in the decomposition are not uniquely determined by w. For example, in Figure 4, we have that $aa\,cb = cb\,bc$ in $M_{\{a,b,c\}}$.

We can concatenate $w_1 \in \Gamma_k$ and $w_2 \in \Gamma_l$, to obtain $w_1w_2 \in \Gamma_{k+l}$. We can perform concatenation of consecutive subwords of a given word $w \in M_\Gamma$, and this operation is also associative. Words $w \in \Gamma_k$ for *odd* k are called *quarks*. Letters from the alphabet Γ are also quarks.

2.1 Reversion and (de)compression of words

It will be convenient to introduce the *reversal* of k-words, defined as the function $E : \Gamma_k \to \Gamma_k$ given by

$$E(\gamma_1\gamma_2\dots\gamma_k) = \gamma_k\dots\gamma_2\gamma_1,$$

for any word $w = \gamma_1\gamma_2\dots\gamma_k \in \Gamma_k$. We also write $\overline{w} := E(w)$, for short. This operation can be in a natural way extended to reversion $E : M_\Gamma \to M_\Gamma$.

In the following definition, we introduce the notion of Šare's system M_Γ (also called *m-system*, according to the original terminology by M. Šare (ʃá:re). It first appeared in 1973 in his doctoral dissertation [6]. See also [8]. The crucial role is played by the notion of equivalence $\sim$ among words of even length, which Šare denoted as a mere equality, i.e. by $=$. We prefer to denote it here by $\sim$, for several obvious reasons. One of them is that the compression axiom $\alpha\beta\beta\gamma \sim \alpha\gamma$ appearing in Eq. (3) below, if written as $\alpha\beta\beta\gamma = \alpha\gamma$, would inevitebly imply that $\ell(\alpha\beta\beta\gamma) = \ell(\alpha\gamma)$, i.e. the contradiction $4 = 2$. The equivalence between words generates a relation of equivalence on the set M_Γ; see Definition 4 below.

Definition 2. *Let Γ be a well ordered alphabet. We say that M_Γ is a* Šare system *generated by Γ if it fulfills the following three properties (Šare's axioms)* :
(*a*) *For any letter $\gamma \in \Gamma$ and for any k-word $w \in \Gamma_k$, where k is an even number, we have the following* reversion property:

$$\gamma w\gamma \sim \gamma\overline{w}\gamma. \tag{1}$$

(*b*) *For any two letters* $\alpha, \beta \in \Gamma$*,* the first compression *law is true*:

$$\alpha\beta\alpha \sim \alpha, \tag{2}$$

where we understand that the combination $\alpha\beta\alpha$ *is a subword of a given word* $w \in M_\Gamma$*.* (*See also Remark 2 below.*)

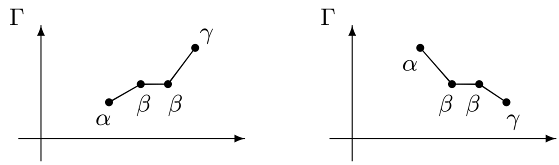

Figure 1: The product $\alpha\beta^2\gamma$, for β between α and γ, is equivalent to $\alpha\gamma$. The figure is selfexplanatory, corresponding to cases of increasing and decreasing sequences $(\alpha, \beta, \beta, \gamma)$ in the product. See Eq. (3).

(*c*) *For any ordered triple* (α, β, γ) *of letters in* Γ (*that is, such that either* $\alpha \le \beta \le \gamma$ *or* $\alpha \ge \beta \ge \gamma$*; see Figure 1*)*,* the second compression *law is true*:

$$\alpha\beta^2\gamma \sim \alpha\gamma, \tag{3}$$

where we understand that the combination $\alpha\beta^2\gamma$ *is a subword of* (*or equal to*) *a given word* $w \in M_\Gamma$*.*

The following simple result shows that the relation $\sim$ behaves well with respect to the product of words.

Lemma 1. *For any four words* $x_1, x_2, y_1, y_2 \in W_\Gamma$ *such that* $xx_1, yy_1 \in M_\Gamma$*, we have that*

$$x_1 \sim y_1 \text{ and } x_2 \sim y_2 \quad \Longrightarrow \quad x_1y_1 \sim x_2y_2.$$

Remark 1. *In a concrete situation, compression property* (2) *can be used either in the sense of compression* (*i.e.,* $\alpha\beta\alpha \mapsto \alpha$)*, or in the sense of* decompression (*i.e.,* $\alpha \mapsto \alpha\beta\alpha$)*. This should be clear from the context, and in both cases we most often say to have used compression property* (2)*, for simplicity. For example, in* $\alpha\beta\alpha\gamma \sim \alpha\gamma \sim \alpha\gamma\alpha\gamma$*, we have compressed* $\alpha\beta\alpha$ *to* α *in the first equivalence, and then decompressed* γ *to* $\gamma\alpha\gamma$ *in the second, where* $\alpha, \beta, \gamma \in \Gamma$*. Similarly for compression property* (3)*. Also, remark that compression property* (2) *involves quarks.*

Remark 2. *Note that if* $\alpha, \beta \in \Gamma$*, then* $\alpha\beta\alpha \notin M_\Gamma$*, so that equivalence in* (2) *is not equivalence of words in* M_Γ*. For any letter* $\varepsilon \in \Gamma$*, the first compression property in* (2) *implies that* $\varepsilon\alpha\beta\alpha \sim \varepsilon\alpha$*, which* is *equivalence* *of words in M_Γ. Properties (a), (b), and (c) are used within larger words in M_Γ. For example, for any three words $w_1, w, w_2 \in M_\Gamma$, property (a) applies as follows: $w_1\gamma w\gamma w_2 \sim w_1\gamma\overline{w}\gamma w_2$, and this is equivalence of words in M_Γ; etc. It is worth noticing that the expressions $\alpha\beta\alpha$ and $\alpha\beta^2\gamma$ appearing in (de)compression axioms (2) and (3) are invariant with respect to reversion.*

The following useful result shows that if two words in M_Γ are equivalent, then their initial letters must coincide, as well as their terminal letters.

Proposition 1. *For any four letters $\alpha_j, \beta_j \in \Gamma$, where $j = 1, 2$, and for any $x, y \in M_\Gamma$, the following* boundary property *is true*:

$$\alpha_1 x\beta_1 \sim \alpha_2 y\beta_2 \quad \Rightarrow \quad \alpha_1 = \alpha_2 \text{ and } \beta_1 = \beta_2. \tag{4}$$

Proof. Note that the property stated in Eq. (4) is true for axioms (1), (2), and (3), appearing in Definition 2. Since each equivalence of words is obtained by consecutive use of these three axioms, the claim follows. □

Definition 3. (*Jorbs and jorbology*) *According to terminology introduced by M. Šare in* [8], *the elements (words) in M_Γ are called* jorbs (*to be pronounced as yorbs*). *The set M_Γ is a semigroup with respect to concatenation.*

The product of any two quarks is clearly a jorb, while the product of any jorb and a quark (and vice versa) is a quark.

Remark 3. *The notion of* jorb, *introduced in the above Definition 3, is obtained by reversion as* $E(\text{broj}) = \text{jorb}$. *Here,* broj – *number (in Croatian), with Γ taken as the usual Latin alphabet. For example,* $E(\text{number}) = \text{rebmun}$. *In this sense, we can speak about Šare's* theory of jorbs, *or of the* jorbology (*the analog of which in English would be 'rebmunology'*).

The following lemma is useful in immediate compression of the words. The last equivalence in (5) below, for $n = 2$ shows that the product of any two elements of the alphabet (i.e., of any atom) is idempotent.

Lemma 2. *For any two positive integers k and l and any two letters $\alpha, \beta \in \Gamma$, we have that*

$$\alpha^k\beta\alpha^l \sim \alpha^{k+l-1}.$$

In particular, $\alpha^2\beta\alpha^2 \sim \alpha^3$.
For for any positive integer n, we have that

$$\alpha^{2n} \sim \alpha^2, \quad \alpha^{2n+1} \sim \alpha, \quad (\alpha\beta)^n \sim \alpha\beta. \tag{5}$$

Proof. The first claim follows from $\alpha^k\beta\alpha^l \sim \alpha^{k-1}(\alpha\beta\alpha)\alpha^{l-1} \sim \alpha^{k+l-1}$. (In case when $k = 1$ or $l = 1$, we agree that the corresponding term α^0 is in fact absent.)

The second equivalence in (5) follows by letting $k = l = n$. The first equivalence follows by multiplying the second one by α.

The last equivalence in (5) for $n = 2$ follows by multiplying Eq. (2) by β from the right. The case when $n \geq 2$ is obtained by mathematical induction. □

The following lemma is a generalization of (de)compression property (3).

Lemma 3. *Let a finite ordered sequence of letters $a \leq \gamma_1 \leq \cdots \leq \gamma_k \leq b$ be given in Γ. Then,*

$$a\gamma_1^2 \dots \gamma_k^2 b \sim ab. \tag{6}$$

The same conclusion holds if we reverse the order of the sequence: $a \geq \gamma_1 \geq \cdots \geq \gamma_k \geq b$.

Proof. The proof follows by mathematical induction. For $k = 1$, the claim is equivalent to axiom (3). Assume that the claim in the lemma is true for some positive integer k. Then for γ_{k+1} such that $\gamma_k \leq \gamma_{k+1} \leq b$ we have that

$$a\gamma_1^2 \dots \gamma_k^2\gamma_{k+1}^2 b = (a\gamma_1^2 \dots \gamma_k^2\gamma_{k+1})\gamma_{k+1}b \sim (a\gamma_{k+1})\gamma_{k+1}b = a\gamma_{k+1}^2 b \sim ab,$$

where we have used the inductive hypotheses in the first equivalence, and compression axiom (3) in the second. Analogous proof can be performed in the case of reverse order of letters. □

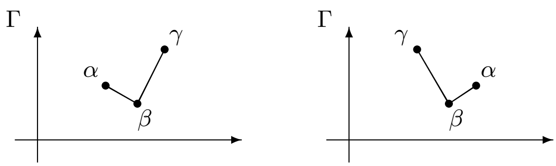

Figure 2: The quark $\alpha\beta\gamma$ appearing on the left, described in Lemma 4, is equivalent to $\alpha^2\gamma$. The quark $\gamma\beta\alpha$ on the right is equivalent to $\gamma\alpha^2$.

In the following lemma, we say that a letter α is *between* letters β and γ in Γ if either $\beta \leq \alpha \leq \gamma$ or $\gamma \leq \alpha \leq \beta$. See Figure 2.

Lemma 4. *Let three letters α, β, γ be given in a well ordered alphabet Γ satisfying (de)compression axioms* (2) *and* (3). *If α is between β and γ, then*

$$\alpha\beta\gamma \sim \alpha^2\gamma, \quad \gamma\beta\alpha \sim \gamma\alpha^2.$$

Proof. Decompressing by (3), and then compressing by (2), we have that

$$\alpha\beta\gamma \sim \alpha(\beta\alpha^2\gamma) = (\alpha\beta\alpha)\alpha\gamma \sim \alpha^2\gamma.$$

The second equivalence is proved in a similar fashion. □

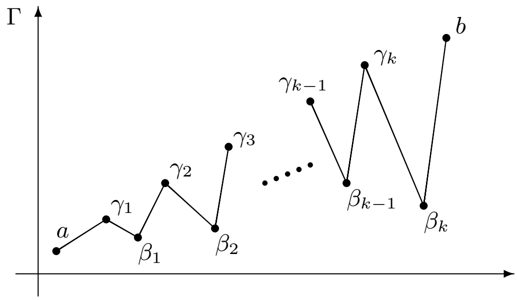

Figure 3: Zizagging product $a\Big(\prod_{j=1}^k \gamma_j\beta_j\Big)b$, described in Proposition 2, is equivalent to ab. Horizontal axis indicates the order of letters in the product.

In Proposition 2 below, we compress the zizagging product in (7). See Figure 3.

Proposition 2. (Zigzagging product) Let a finite ordered sequence of letters $a \le \gamma_1 \le \cdots \le \gamma_k \le b$ be given in Γ. Assume that a set of letters $\{\beta_1, \dots, \beta_k\}$ is given such that $\beta_j \le \gamma_j$ for each $j = 1, \dots, k$. Then (see Figure 3),

$$a\Big(\prod_{j=1}^k \gamma_j\beta_j\Big)b \sim ab. \tag{7}$$

An analogous claim holds if we reverse all inequalities in the assumptions.

Proof. First remark that, since γ_j is between β_j and γ_{j+1}, then by Lemma 4, $\gamma_j\beta_j\gamma_{j+1} \sim \gamma_j^2\gamma_{j+1}$. We proceed by inducton with respect to k. From this we see by mathematical induction with respect to k that $\prod_{j=1}^k \gamma_j\beta_j \sim \prod_{j=1}^k \gamma_j^2$. The claim follows by using Lemma 3:

$$a\Big(\prod_{j=1}^k \gamma_j\beta_j\Big)b \sim a\Big(\prod_{j=1}^k \gamma_j^2\Big)b \sim ab.$$

□

$\cdot$	aa	ab	ba	bb
aa	aa	ab	aa	$aabb$
ab	aa	ab	$abba$	ab
ba	ba	$baab$	ba	bb
bb	$bbaa$	bb	ba	bb

Figure 4: Multiplication table of the atoms (modulo equivalence $\sim$) of the Šare system $M_{\{a,b\}}$. The set of atoms Γ_2 is not a grupoid for $\Gamma = \{a, b\}$. Among 16 products, four them are incompressible.

$\cdot$	aa	ab	ac	ba	bb	bc	ca	cb	cc
aa	aa	ab	ac	aa	$aabb$	$aabc$	aa	$aacb$	$aacc$
ab	aa	ab	ac	$abba$	ab	ac	$abca$	ab	$abcc$
ac	aa	ab	ac	$acba$	ab	ac	$acca$	$accb$	ac
ba	ba	$baab$	$baac$	ba	bb	bc	ba	bb	$bacc$
bb	$bbaa$	bb	bc	ba	bb	bc	ba	bb	$bbcc$
bc	$bcaa$	bb	bc	ba	bb	bc	$bcca$	$bccb$	bc
ca	ca	$caab$	$caac$	ca	cb	$cabc$	ca	cb	cc
cb	$cbaa$	cb	cc	ca	cb	$cbbc$	ca	cb	cc
cc	$ccaa$	$ccab$	cc	$ccba$	$ccbb$	cc	ca	cb	cc

Figure 5: Multiplication table of the atoms (modulo $\sim$) of the Šare system $M_{\{a,b,c\}}$. Note, for example, that $aa\,cb = a(acb) \sim aabb$, $ca\,bc = (cab)c \sim cb\,bc$. Among 81 products, 27 of them are incompressible.

Since the multiplication of jorbs in Šare system M_Γ, generated by a given well ordered alphabet Γ reduces to multiplication of its atoms, it is meaningful to have in mind the corresponding multiplication table of the atoms. Taking into account the axioms of reversion (1) and compression (2) and (3), we see that that the product of two atoms in M_Γ does not have to be an atom already in the case of the well ordered alphabet consisting of two letters, $\Gamma = \{a, b\}$; see Figure 4. Indeed, we have that the products are maximally compressed (here, as well as in the following table, reversion axiom (1) is not needed). In other words, the set of atoms is not even a groupoid under the multiplication.

In the case of the three element well ordered alphabet $\Gamma = \{a, b, c\}$, we have the multiplication table of 9 corresponding atoms shown in Figure 5 Here, we also needed just the last two of the Šare axioms (i.e., compression axioms (2) and (3)) for the associated system M_Γ are used.

For example, *acba* $\sim$ *abba*, using Lemma 4. Similaraly, *abca* $\sim$ *abba*, etc. It is clear that the multiplication of atoms (and more generally, of jorbs) is associative, since the multiplication of letters in the alphabet is associative as well.

Remark 4. *Note that the multiplication table of the atoms in the Šare system* $M_{\{a,b\}}$ *is embedded into the one corresponding to* $M_{\{a,b,c\}}$*; see Figures 4 and 5. Also remark that* $M_{\{a,b\}}$ *can be identified with* $M_{\{a,c\}}$*, since these two subsemigroups of* $M_{\{a,b,c\}}$ *are clearly isomorphic. More generally, for any two alphabets* Γ_1 *and* Γ_2*, such that* $|\Gamma_1| < |\Gamma_2|$*, it is clear that the Šare semigroup* M_{Γ_1} *is isomorphic to a subsemigroup of* M_{Γ_2}*, i.e.,* M_{Γ_1} *can be embedded into* M_{Γ_2}*.*

Remark 5. (*Complete noncommutativity; open problem*) *From the multiplication tables shown in Figures 4 and 5, we notice that corresponding Šare systems* M_Γ *are* completely noncomutative *in the following sense: for any two atoms* $x, y \in M_\Gamma$*, such that* $x \neq y$*, we have that* $xy \neq yx$*. It would be interesting to know if analogous property holds for any well ordered finite alphabet* Γ*.*

2.2 Valuation and distance of letters of an alphabet

By a *valuation* of letters in an alphabet Γ we mean any integer-valued function $v : \Gamma \to \mathbb{Z}$ which is increasing, and the absolute value of the difference of valuations of any two consecutive letters in the alphabet is equal to 1. It is clear that a valuation v of a given alphabet Γ is determined uniquely up to an additive integer constant.

For example, the letters α, β, γ,... of the Greek alphabet Γ can be valuated with 1, 2, 3,... (but also with 0, 1, 2,...). The alphabet $\Gamma = \{0, 1, \dots, 9\}$ can be valuated by the values of the alphabet itself (but also with values translated by 1, i.e., with $1, 2, \dots, 10$).

For a given alphabet Γ, we define the *distance* $\delta(a, b)$ between any two letters $a, b \in \Gamma$, by

$$\delta(a, b) = |v(a) - v(b)|.$$

It is clear that the distance does not depend on the choice of the valuation of the alphabet Γ. Also, $0 \leq \delta(a, b) \leq |\Gamma| - 1$, and the bounds for $\delta(a, b)$ are optimal. Furthermore, $\delta(a, b)$ can achieve all of the values in the set $[0, |\Gamma| - 1] \cap \mathbb{Z}$.

The distance δ is clearly a *metric* on the alphabet Γ. It is obvious that in the triangle inequality, we have equality (that is, $\delta(a, c) = \delta(a, b) + \delta(b, c)$)

if and only if the value of $v(b)$ is *between* the values of $v(a)$ and $v(c)$ (by this we mean that $v(b) \in [v(a), v(c)]$ if $v(a) \leq v(b)$, or $v(b) \in [v(c), v(a)]$ if $v(c) \leq v(a)$).

Remark 6. *It is sometimes convenient to identify the letters of an alphabet* Γ *with the set of its valuations, so that* $\Gamma \subset \mathbb{Z}$*. For example, n is the cardinality of* Γ*, then we can write* $\Gamma = \{1, 2, \ldots, n\}$*. This will be done in the proof of Theorem 1 below.*

3 Šare's sum of jorbs

To any four-letter subjorb $abcd$ of a given jorb $w \in M_\Gamma$ we associate the following important numerical expression $\lambda(abcd)$, that we call *Šare's sum* of $abcd$:

$$\lambda(abcd) := \delta(a, b) - \delta(b, c) + \delta(c, d). \tag{8}$$

The next theorem shows (somewhat surprisingly) that it is possible to characterize equivalence $abcd \sim ab$ by means of the indicated alternating sum of distances appearing in (8). This result was first published in [8, Theorem 1, case 1.9, p. 19], but without the proof.

Theorem 1. (Šare, see [8, Theorem 1, case 1.9]) *Assume that compression properties* (2) *and* (3) *hold in* M_Γ*. Then, for any four letters* $a, b, c, d \in \Gamma$ *we have that*

$$\delta(a, b) - \delta(b, c) + \delta(c, d) = \delta(a, d) \quad \Longleftrightarrow \quad abcd \sim ad. \tag{9}$$

We postpone the proof of this theorem until after Lemma 6 below. It will be convenient to extend the definition of Šare's sum λ to any jorb $a_1\alpha_2 \ldots, a_k \in M_\Gamma$, where $a_j \in \Gamma$ for all $j = 1, 2, \ldots, k$ (recall that k is even):

$$\lambda(a_1\alpha_2 \ldots a_k) := \delta(a_1, a_2) - \delta(a_2, a_3) + \cdots + \delta(a_{k-2}, a_{k-1}) - \delta(a_{k-1}, a_k). \tag{10}$$

The alternating sum on the right-hand side of (10) has $k - 1$ (i.e., odd number of) terms. Note also that $\lambda(a_1a_2) = \delta(a_1, a_2)$. Thus, Eq. (10) defines the function

$$\lambda : M_\Gamma \to \mathbb{Z}, \quad \lambda(a_1\alpha_2 \ldots a_k) := \sum_{j=1}^{k-1} (-1)^{j-1} \delta(a_j, a_{j+1}).$$

Šare's sum λ for any jorb in M_Γ was introduced in [8, Definition 11 on p. 23], but with the opposite signs on the right-hand side of (10) then we have here. Of course, this slight change is inessential.

The following lemma shows that applying axioms (3), (2), and (1) to a jorb, does not change the Šare sum of the jorb.

Lemma 5. *Let an arbitrary* 4*-word* $abcd \in M_\Gamma$ *be given, where* $a, b, c, d \in \Gamma$*. Then we have the following:*

(i) *if* $b = c$ *and* $a \le b \le d$*, then* $\lambda(abcd) = \delta(a, d)$*;*

(ii) *if* $a = c$*, then* $\lambda(abcd) = \delta(a, d)$*; the same for* $b = d$*;*

(iii) *if* $w \in M_\Gamma$ *is any word, then* $\lambda(awa) = \lambda(a\overline{w}a)$*. (Furthermore,* $\lambda(w) = \lambda(\overline{w})$*.)*

Proof. The proof follows easily by direct inspection.
(i) If $b = c$ and $a \le b \le d$, then

$$\lambda(abcd) = \lambda(abbd) = \delta(a, b) - \delta(b, b) + \delta(b, d) = \delta(a, b) + \delta(b, d) = \delta(a, d).$$

(ii) If $a = c$, then $abcd = abad$, so that

$$\lambda(abcd) = \lambda(abad) = \delta(a, b) - \delta(b, a) + \delta(a, d) = \delta(a, d);$$

the same for $b = d$.
(iii) If $w = \alpha_1\alpha_2 \dots \alpha_{k-1}\alpha_k \in M_\Gamma$ is any word (with k even), then $\lambda(awa) = \delta(a, \alpha_1) - \delta(\alpha_1, \alpha_2) + \cdots - \delta(a_{k-1}, \alpha_k) + \delta(\alpha_k, a)$. This expression is clearly equal to

$$\begin{aligned}\lambda(a\overline{w}a) &= \lambda(a\alpha_k\alpha_{k-1} \dots \alpha_2\alpha_1 a)\\ &= \delta(a, \alpha_k) - \delta(\alpha_k, \alpha_{k-1}) + \cdots - \delta(a_2, \alpha_1) + \delta(\alpha_1, a).\end{aligned}$$

□

The following lemma follows easily from the above Lemma 5.

Lemma 6. *For any two jorbs* $x, y \in M_\Gamma$*, if* $x \sim y$ *then* $\lambda(x) = \lambda(y)$*.*

Proof. It suffices to note that y is obtained from x using axiom of reversion (1), and axioms of (de)compression (2) and (3) finitely many times. The claim follows from Lemma 2. □

Our proof of the sufficiency part of Theorem 1, provided below in part (a), is constructive. More precisely, the construction of the desired sequence of equivalence $ab \sim \cdots \sim abcd$ (or vice versa), using axioms of (de)compression (2) and (3) only, is in fact an algorithm consisting of 24 cases.

Proof of Theorem 1. Part (a). We indicate the proof of the sufficiency part, i.e., of implication $\Rightarrow$. Assume without loss of generality (see Remark 6) that $\Gamma = \{1, 2, \dots, n\}$, where $n = |\Gamma|$.

For a given ordered quadruple (a, b, c, d) of integers in Γ, we have $4! = 24$ possible orders indicated by the associated permutations, so that we have to consider the corresponding 24 cases.

Case 1. Assume that $a \le b \le c \le d$. Then, the condition $\delta(a, b) - \delta(b, c) + \delta(c, d) = \delta(a, d)$ reads as $(b - a) - (c - b) + (d - c) = d - a$, so that $b = c$. Therefore, $abcd \sim ab^2d \sim ad$ due to the compression property (3). Recall that, here, $b^2 = bb$ is taken in the sense of concatenation.

Case 2. If $b \le a \le c \le d$, then condition $\delta(a, b) - \delta(b, c) + \delta(c, d) = \delta(a, d)$ reads as $(a - b) - (c - b) + (d - c) = d - a$, i.e., $a = c$. Hence, from compression property (2) we conclude that $abcd \sim abad \sim ad$.

Case 3. Assume that $a \le c \le b \le d$. Then, the condition $\delta(a, b) - \delta(b, c) + \delta(c, d) = \delta(a, d)$ reads as $(b - a) - (b - c) + (d - c) = d - a$, which is always fulfilled. Thus, using the compression properties (2) and (3), we have that

$$ad \sim abad \sim abaccd \sim abccaccd \sim abcccd \sim abcd.$$

Case 4. If $d \le c \le b \le a$, then condition $\delta(a, b) - \delta(b, c) + \delta(c, d) = \delta(a, d)$ reads as $(a - b) - (b - c) + (c - d) = a - d$, i.e., $b = c$, and hence, we have an analogous situation to Case 1.

The remaining 20 cases are treated analogously. See the Appendix near the end of this paper.

Part (b). In order to prove the converse implication (i.e., of $\Leftarrow$), assume that $abcd \sim ab$. This means that, if we start with ab, then the expression $abcd$ is obtained from ab by applying axioms of reversion (1) and compressions (2) and (3) finitely many times. The inital value of $\lambda(ad) = \delta(a, d)$ is not changed after any such step, in light of Lemma 6. This completes the proof of part (b), as well as the proof of the theorem. □

Now we describe a suitable algorithm for compressing any given jorb, by means of its pseudocode. Let $\omega \in M_\Gamma$, $\ell(\omega)$ - length of a jorb, $w[i]$ - letter at position i in ω, $\omega[i:j]$ - slice, subjorb of ω from position i to j, $v()$ - valuation function, $reverse(\omega)$ - reversing string ω, $\omega.insert(i, \omega')$ - function for insertion of subjorb ω' at i-position of ω, $append(\omega')$ - appending subjorb into the list.

Corollary 1. *For any four letters* $a, b, c, d \in \Gamma$ *we have that*

$$abcd \sim ad \quad \Leftrightarrow \quad dcba \sim da.$$

Proof. If $abcd \sim ad$, then by Theorem 1,

$$\delta(a,b) - \delta(b,c) + \delta(c,d) = \delta(a,d).$$

Writing this equality as

$$\delta(d,c) - \delta(c,b) + \delta(b,a) = \delta(d,a),$$

then again by Theorem 1 it follows that $dcba \sim da$. The converse implication follows along the same lines. □

Remark 7. *Note that, if* $abca \sim a^2$ *for some three letters* $a, b, c \in \Gamma$, *then by Corollary 1 we have that* $abca \sim acba$. *But from reversion axiom* (1) *we know that* $abca \sim acba$ *without any additional condition on the expression* $abca$.

Thanks to Theorem 1, we can derive the following consequences.

Corollary 2. *Assume that a finite sequence of ordered quadruples of letters*

$$(a_j, b_j, c_j, d_j) \in \Gamma^4, \quad j = 1, \dots, k$$

is given, such that $\lambda(a_j b_j c_j d_j) = \delta(a_j, d_j)$ *for each* j. *Then,*

$$\prod_{j=1}^{k} a_j b_j c_j d_j \sim \prod_{j=1}^{k} a_j d_j. \tag{11}$$

Furthermore, if the following zigzagging condition is satisfied, $a_1 \le d_1 \le d_2 \le \cdots \le d_{k-1} \le d_k$ *and* $a_{j+1} \le d_j$ *for all* $j = 1, \dots, k-1$, *then*

$$\prod_{j=1}^{k} a_j b_j c_j d_j \sim a_1 d_k. \tag{12}$$

Analogous claim holds also if we reverse all inequalities in the assumptions preceding Eq. (12).

Proof. The first claim follows immediately from Theorem 1. The second claim follows from the first one, by rewriting the product appearing on the right-hand side of (11) and then making use of Proposition 2:

$$\prod_{j=1}^{k} a_j d_j = a_1 \Big(\prod_{j=1}^{k-1} d_j a_{j+1} \Big) d_k \sim a_1 d_k.$$

□

In the following result we find the minimum and maximum of the set $\lambda(\Gamma_k) := \{\lambda(w) : w \in \Gamma_k\}$, where k is even.

Proposition 3. *Let k be any even positive integer. Then*

$$\min \lambda(\Gamma_k) = -(n-1)\Big(\frac{k}{2} - 1\Big), \quad \max \lambda(\Gamma_k) = (n-1)\frac{k}{2}.$$

Furthermore, the minimum and maximum are both achieved on precisely two different jorbs from Γ_k, that are provided in the proof.

Proof. Let $\alpha := \min \Gamma$ and $\omega := \max \Gamma$, i.e., α and ω are the first and last letters of alphabet Γ, respectively. Let us define the folowing four atoms:

$$A = \alpha\alpha, \quad B = \omega\omega, \quad C = \alpha\omega, \quad D = \omega\alpha.$$

Then let us define an element of Γ_k of the form $w'_{min} := ABAB\dots$, where A and B appear alternately, and the last atom in w'_1 is A or B, depending on whether $k/2$ is odd or even, respectively. Since $\delta(\alpha, \omega) = n - 1$, and this is the maximal possible distance of any two letters in Γ, we see that the value

$$\lambda(w'_{min}) = \lambda(\alpha\alpha\omega\omega\alpha\alpha\omega\omega\dots) = -\Big(\frac{k}{2} - 1\Big)\,\delta(\alpha, \omega) = -\Big(\frac{k}{2} - 1\Big)(n-1)$$

is the smallest possible. The same minimal value is obtained for $w''_{min} := BABA\dots$ It is easy to see that w'_{min} and w''_{min} are the only minima of $\lambda|_{\Gamma_k}$.

Analogously, $w'_{max} := CDCD\dots$ and $w''_{max} := DCDC\dots$ are the two unique points of maxima of $\lambda|_{\Gamma_k}$. □

Remark 8. *If the product of two atoms $A = ab$ and $B = cd$ has nontrivial compression (i.e., it is equivalent to an atom in this case), then we must have that $AB \sim ad$, in light of Proposition 1. So, due to Theorem 1, the number of* (*nontrivially*) *compressible words in Γ_4 is equal to the number of solutions of equation $\lambda(abcd) = \delta(ad)$.*

4 Incompressible jorbs

In the following two definitions, we introduce the notion of equivalence among jorbs in M_Γ, and the notion of incompressible jorbs.

Definition 4. *Two jorbs* $w_1, w_2 \in M_\Gamma$ *are said to be* equivalent*, denoted by* $w_1 \sim w_2$*, if* w_2 *can be obtained from* w_1 *by consecutive use of the axiom of reversion* (1) *and of the two axioms of compression* (2) *and* (3)*. This relation is clearly a relation of equivalence on the Šare system* M_Γ*. For each jorb* $w \in M_\Gamma$*, we denote by*

$$[w] = \{w' \in M_\Gamma : w' \sim w\}$$

the corresponding equivalence class.

It is natural to define the *quotient Šare system* $\widetilde{M_\Gamma}$ as

$$\widetilde{M_\Gamma} = M_\Gamma/_\sim = \{[w] : w \in M_\Gamma\}.$$

It is also a semigroup, with respect to multiplication of classes defined via concatenation of their representatives: $[x][y] = [xy]$, for any $[x], [y] \in \widetilde{M_\Gamma}$. This multiplication is well defined, since it does not depend on the choice of representatives. Indeed, if $x \sim x_1$ and $y \sim y_1$, then clearly $xy \sim x_1y_1$. In other words, $[x] = [x_1]$ and $[y] = [y_1]$ imply that $[xy] = [x_1y_1]$. Note also that if x is a quark, then $[x]$ is not well defined.

In the following proposition we show that each element of the form $[ab] \in \widetilde{M_\Gamma}$, generated by an atom $ab \in \Gamma_2$, is *regular* (that is, for each $[ab] \in \widetilde{M_\Gamma}$ there exists $[x] \in \widetilde{M_\Gamma}$ such that $[ab] = [ab][x][ab]$). The definition of regular elements and regular semigroups can be seen in [2, p. 44] or in [5, I.7.1 Definition on p. 33]. The concept of regularity was introduced by John von Neumann (1936) in ring theory. The proposition also shows that Šare quotient semigroups are neither regular nor inverse semigroups, when $|\Gamma| \geq 2$. The definition of general inverse semigroup can be seen e.g. in [5, II.1.1 Definition on p. 71].

Proposition 4. *Šare quotient systems* $\widetilde{M_\Gamma}$ *are generated by idempotents. All elements* $[ab] \in \widetilde{M_\Gamma}$ *generated by the atoms* $ab \in \Gamma_2$ *are regular. The elements of the form* $[aabb] \in \widetilde{M_\Gamma}$ *are neither idempotent nor regular for* $a \neq b$ *in* Γ*. Furthermore, Šare systems are neither regular nor inverse semigroups for* $|\Gamma| \geq 2$*.*

Proof. To prove that elements of the form $[ab]$ are idempotents, note that the last equivalence in Eq. (5) implies that $[ab]^n = [ab]$ for all $n \in \mathbb{N}$. In particular, $[ab]^3 = [ab]$, so that $[ab]$ is regular with $[x] = [ab]$. The idempotency $[aabb]$ for $a \neq b$ follows from Theorem 2 below.

If $a, b, c, d \in \Gamma$ are such that $\lambda(abcd) = \delta(ab)$, then according to Theorem 1 we have that $abcd \sim ad$. In particular, $[ab][cd] = [abcd] = [ad]$, and hence,

$$[ab][cd][ab] = [ad][ab] = [adab] = [ab], \tag{13}$$

where in the last equality we have used compression axiom (2). On the other hand,

$$[cd][ab][cd] = [cd][abcd] = [cd][ad] = [cdad] = [cd]. \tag{14}$$

As we can see, if the alphabet Γ contains at least two elements, then for any given element $[ab] \in \widetilde{M}_\Gamma$, we have multiple solutions $[cd] \in \widetilde{M}_\Gamma$ satisfying simultaneously equations $[ab][cd][ab] = [ab]$ and $[cd][ab][cd] = [cd]$ appearing in (13) and (14). This follows immediately from the proof of Theorem 1; see also the Appendix. Consequently, the Šare quotient semigroup is not an inverse semigroup.

Alternative (and more direct) proof. Let a and b be any two different letters from the alphabet Γ, and let us define $x = [aa] \in \widetilde{M}_\Gamma$. Then the following system with $y \in \widetilde{M}_\Gamma$ as an unknown,

$$xyx = x \quad \text{and} \quad yxy = y$$

appearing in the definition of general inverse semigroup, possesses two obvious different solutions $y = [aa]$ and $y = [ab]$. Indeed, for $y = [aa]$ we have that $xyx = yxy = [aa]^3 = [(aa)^3] = [aa] = x = y$, where we have used the last equivalence in Eq. (5) of Lemma 2. If $y = [ab]$, then the compression axiom (2) implies that $xyx = [aa\,ab\,aa] = [aa(aba)a] = [a^3a] = [aa] = x$, and $yxy = [ab\,aa\,ab] = [(aba)a^2b] = [a^3b] = [ab] = y$. Consequently, $\widetilde{M}_\Gamma$ is not an inverse semigroup. □

It is of obvious interest to find the 'best' representative of any given class $[w] \in \widetilde{M}_\Gamma$, in the sense of minimizing the length representatives, i.e., of $\ell(w)$. For this reason, we now pass to the definition of incompressible jorbs. (Šare uses the term 'canonical jorb' instead; see [7, Definition 019 on p. 6].)

Definition 5. *We say that a jorb $w \in M_\Gamma$ is* incompressible, *if its length $\ell(w)$ is minimal in $[w]$, that is,*

$$\ell(w) = \min\{\ell(w') : w' \sim w\}.$$

It is interesting to note that in the case of the alphabet $\Gamma = \{a, b\}$ consisting of two letters only, the *incompressible* jorb w_{zip} equivalent to a given jorb $w \in M_\Gamma$ *is uniquely* determined by w; see [7, Theorem 05 on p. 6]. According to [7, Theorem 04 on p. 6], the set of all incompressible words w of the Šare system $M_{\{a,b\}}$ can be characterized as follows:

(i) either w is an atom (that is, equal to aa, ab, ba, or bb), or (*ii*) any two consecutive atoms appearing in w are mutually *dual* (the *dual* of aa is bb, and vice versa, while the dual of ab is ba, and vice versa). (See also Figure 4.)

An immediate consequence of this result is the following theorem.

Theorem 2. *Let $\Gamma = \{a, b\}$. For any even positive integer k, we have precisely four incompressible yorbs in Γ_k:*

$$w_{1k} = aa\,bb\,aa\,bb \cdots \in \Gamma_k, \quad w_{2k} = ab\,ba\,ab\,ba \cdots \in \Gamma_k,$$
$$w_{3k} = ba\,ab\,ba\,ab \cdots \in \Gamma_k, \quad w_{4k} = bb\,aa\,bb\,aa \cdots \in \Gamma_k.$$

Equivalently, the set of incompressible jorbs in the Šare system $M_{\{a,b\}}$ is equal $\{w_{ik} : k \in \mathbb{N},\ i = 1, 2, 3, 4\}$. All these jorbs are mutually nonequivalent, that is, $[w_{ik}] \cap [w_{jl}] = \emptyset$, whenever $(i, k) \neq (j, l)$. In particular, the corresponding quotient Šare system is equal to $\widetilde{M}_{\{a,b\}} = \{[w_{ik}] : k \in \mathbb{N},\ i = 1, 2, 3, 4\}$.

Remark 9. (Open problem) *In the notation of the above Theorem 2, we see that $[w_{ik}][w_{jl}] = [w_{mn}]$, for some m and n depending on i, k, j, l. It is clear that*

$$[w_{ik}]^2 = \begin{cases} [w_{i,2k}] & \text{if } k/2 \text{ is even,} \\ [w_{i,2k-2}] & \text{if } k/2 \text{ is odd.} \end{cases}$$

When $i \neq j$, we have massive cancellations. For example, $[w_{14}][w_{24}] = [w_{12}]$. It would be of interest to find explicit expressions of the functions $m = m(i, k, j, l)$ and $m = m(i, k, j, l)$.

Remark 10. (Open problem) *Assume that $|\Gamma| = n$. In the multiplication table of n^2 atoms in Γ_2, there are n^4 products. What is the number $a(n)$ of irreducable products? According to Figures 4 and 5, we have that $a(2) = 4$ and $a(3) = 27$.*

In the sequel, it will be convenient to introduce the following notation.

Definition 6. *For any jorb $w \in M_\Gamma$, we let $\partial_-(w)$ and $\partial_+(w)$ be the first and the last letter of w, respectively. The resulting functions $\partial_\pm : M_\Gamma \to \Gamma$ are called* left and right boundary functions *or* Šare's boundary functions.

We provide a pseudocode of an algorithm for compressing a given jorb $\omega \in M_\Gamma$ of arbitrary length. If n is the length of ω, that is, $\ell(\omega) = n$, then the complexity of this algorithm is of order $O(n)$ when $n \to \infty$. This means that it is tractable.

Algorithm 1 RVS(ω) - reversing, axiom (1); ZIP3(ω) - the first compression low, axiom (2); ZIP4(ω) - the second compression low, axiom (3); EXT(ω) - the jorb's expansions, Eq. (6)

Input for all functions *jorb* ω
Output *list of jorbs for rvs() function, transformed jorb* ω *for all other functions*
RVS(ω):
rev_list $\leftarrow$ [ ]
for $i = 1$ to $\ell(\omega) - 1$ **do**
    for $j = \ell(\omega) - 1$ to $i + 1$ **step** -1 **do**
        if ($j - i$ is even) **and** ($\omega[i] = \omega[j]$) **and** ($\omega[i+1:j] \neq \omega[j:i]$ **then**)
            rev_list $\leftarrow$ *append* $(\omega[0:i] + reverse(\omega[i:j]) + \omega[j:\ell(\omega)])$
return *rev_list*
ZIP3(ω):
$i \leftarrow 0$
while $i < \ell(\omega) - 2$ **do**
    if $\omega[i] = \omega[i+2]$ **then**
        $\omega = \omega[0:i] + \omega[i+2:\ell(\omega)]$
    else
        $i \leftarrow i + 1$
return ω
ZIP4(ω):
$i \leftarrow 0$
while $i < \ell(\omega) - 3$ **do**
    if $\delta(\omega[i], \omega[i+1]) - \delta(\omega[i+1], \omega[i+2]) + \delta(\omega[i+2], \omega[i+3]) = \delta(\omega[i], \omega[i+3])$ **then**
        $\omega = \omega[0:i+1] + \omega[i+3:\ell(\omega)]$
        $i \leftarrow i - 1$
    $i \leftarrow i + 1$
return ω
EXT(ω):
for $i = 1$ to $\mathrm{l}(\omega - 1)$ **do**
    $a \leftarrow v(\omega[i]); \quad b \leftarrow v(\omega[i+1])$
    if $a + 1 < b$ **then**
        $k \leftarrow 1$
        for $j = a + 1$ to b **do**
            $\omega.insert(i + k, \Gamma[j] + \Gamma[j])$
            $k \leftarrow \mathrm{k} + 1$
            $i \leftarrow \mathrm{i} + 1$
    if $a > b + 1$ **then**
        $k \leftarrow 1$
        for $j = b - 1$ to a **step** -1 **do**
            $\omega.insert(i + k, \Gamma[j] + \Gamma[j])$
            $k \leftarrow \mathrm{k} + 1$
            $i \leftarrow \mathrm{i} + 1$
return ω

Here is the main progam.

Algorithm 2 COMPRESS(ω) MAIN PROGRAM

```
while True do
    ω′ ← zip3(zip4(ext(ω)))
    ω″_list ← rvs(ω′)
    smaller ← False
    for ω″ in ω″_list do
        ω‴ ← zip3(zip4(ω″))
        if ℓ(ω‴) = ℓ(ω′) then
            continue
        else
            ω ← ω‴
            smaller ← True
            break
    if smaller = False then
        break
return ω
```

In the following theorem, we characterize all incompressible jorbs in M_Γ. We also introduce the set of all subjorbs z of a given jorb w for which the boundary letters of z coincide, that is, $\partial_-(z) = \partial_+(z)$ (so that $z \sim \overline{z}$, by axiom (1), where $\overline{z}$ denotes reversion of z). By $R(w)$ we denote the set of jorbs w_r that can be obtained from w by reversion of such subjorbs z of w. (Note that the length of w_r is left unchanged, i.e., $\ell(w_r) = \ell(w)$ for all $w_r \in R(w)$.) In other words, the set $R(w)$ is the set of all jorbs that can be obtained from w by applying reversion axiom (1). The set $R(w)$ may be empty.

Theorem 3. (Incompressible jorbs in M_Γ) *The set of incompressible jorbs in M_Γ is equal to $\Gamma_2 \cup G$ where G is the set of jorbs $w \in M_\Gamma \setminus \Gamma_2$ such that for all possible subjorbs x of jorbs in $\{w\} \cup R(w)$, of length equal to 4, we have that*

$$\lambda(x) \neq \delta\big(\partial_-(x), \partial_+(x)\big). \tag{15}$$

Proof. It is clear that each of the atoms of M_Γ (i.e., each element of Γ_2) is incompressible. Next, it suffices to note that by Theorem 1, condition (15) is equivalent to incompressibility of x, since otherwise (i.e., if we had equality in (15)), x would be compressible to $\partial_-(x)\partial_+(x) \in \Gamma_2$. Also note that compression axioms (2) and (3) deal only with with subjorbs of w of length at most 4. □

Remark 11. *For each $w \in M_\Gamma$, there exists an* incompressible *jorb w_{zip} equivalent to it. Is w_{zip} uniquely determined by w? In general, the* *answer is no. For example, if* $\Gamma = \{0, 1, 2, \dots, 9\}$*, then for* $w = 123221$*, by using* (*de*)*compression axioms* (2) *and* (3)*, we obtain two different zipped* (*i.e., maximally compressed*) *jorbs:*

$$w = 1(232)21 \sim w'_{zip} = 1221, \quad w = 12(3221) \sim w''_{zip} = 1231.$$

Of course, by transitivity we have that $w'_{zip} \sim w''_{zip}$*.*

In the following lemma, we say that a jorb w_{zip} is a *zipped jorb* (or *maximally compressed* jorb) with respect to a given jorb $w \in M_\Gamma$, if $w \sim w_{zip}$ and w_{zip} is incompressible.

Lemma 7. (*a*) *If* w'_{zip} *and* w''_{zip} *are any two zipped jorbs of a given* $w \in M_\Gamma$*, then* $w'_{zip} \sim w''_{zip}$*.*

(*b*) *Furthermore,* $w_1 \sim w_2$ *if and only if* $w_{1,zip} \sim w_{2,zip}$*, where* $w_{1,zip}$ *is any zipped jorb of* w_1*, and* $w_{2,zip}$ *is any zipped jorb of* w_2

Proof. (*a*) Since $w \sim w'_{zip}$ and $w \sim w''_{zip}$, then by symmetry and transitivity of relation '$\sim$', we have that $w'_{zip} \sim w''_{zip}$. Claim (*b*) follows immediately from (*a*). □

Proposition 5. *For any word* $w \in M_\Gamma$*, the corresponding equivalence class* $[w] \in \widetilde{M}_\Gamma$ *is infinite.*

Proof. Each letter α appearing in w can be decompressed to α^{2n+1} with arbitrary positive integer n, using Eq. (5) in Lemma 2. Hence, if for example $w = \alpha w'$, then we have that $\{\alpha^{2n+1} w' : n \in \mathbb{N}\} \subset [w]$. □

According to terminology of Šare introduced in [8, p. 19], the set of all words in M_Γ of minimal length (that is, of length 2), is called the *zero-base* of the semigroup M_Γ. It coincides with the set of *atoms* of M_Γ, that is, with Γ_2. Any given word $w \in M_\Gamma$ can be obtained as a product of atoms. Moreover, the order of the atoms in the product is uniquely determined by w. Each atom is incompressible and idempotent (see the last equivalence in Eq. (5) of Lemma 5).

Remark 12. *The set of atoms* Γ_2 *can be considered as a 'derived' alphabet for the Šare system* M_Γ*. Note, however, that reversion and compression rules described by Eqs.* (1)*,* (2)*, and* (3) *are formulated in terms of elements of the primary alphabet* Γ*, and not of* Γ_2*. This is the reason of introducing the set of atoms of the semigroup* M_Γ *of jorbs. See also Remark 3. Furthermore, the set* Γ_2 *is not closed under multiplication, provided* Γ *consists of at least two letters; see Figure 4.*

Remark 13. (*Open problem*) *For any even number k, find the number $I(n,k)$ of mutually nonequivalent incompressible jorbs in Γ_k, in dependence with prescribed values of an even positive integer k and $n = |\Gamma|$. For example, if $k = 2$, then this number is equal to $I(n,2) = n^2$, since all the atoms in Γ_2 are incompressible and mutually nonequivalent. For $k = 4$, see Remark 8. If $n = 2$, then according to Theorem 2, we have that $I(2,k) = 4$ for all k.*

Remark 14. *The Šare sum behaves nicely with respect to the product of jorbs x and $y \in M_\Gamma$* (see [8, p. 24]) :

$$\lambda(xy) = \lambda(x) + \lambda(y) - \delta\big(\partial_+(x), \partial_-(y)\big).$$

Here, $\partial_\pm$ denote positive and negative boundary functions introduced in Definition 6. Consequently, denoting by $M_\Gamma(c)$ the set of all $w \in M_\Gamma$ for which $\partial_-(w) = \partial_+(w) = c$, where c is a fixed letter from the alphabet Γ, we have that the restriction $\lambda_c = \lambda|_{M_\Gamma(c)}$ is homomorphism of semigroups $(M_\Gamma(c), \cdot)$ and $(\mathbb{Z}, +)$, since $\delta(c,c) = 0$:

$$\lambda_c(xy) = \lambda_c(x) + \lambda_c(y), \quad \textit{for all } x, y \in M_\Gamma(c).$$

Assuming that $[x] = [y]$ (that is, $x \sim y$), by Lemma 6 we know that $\lambda(x) = \lambda(y)$. Hence, homomorphism λ_c induces a homomorphism $\tilde{\lambda}_c : \widetilde{M_\Gamma}(c) \to \mathbb{Z}$, defined by $\tilde{\lambda}_c([x]) = \lambda(x)$, where $\widetilde{M_\Gamma}(c) = \{[x] : x \in M_\Gamma(c)\}$. In other words, $\tilde{\lambda}_c([x][y])] = \tilde{\lambda}_c([x]) + \tilde{\lambda}_c([y])$ for all $[x], [y] \in \widetilde{M_\Gamma}(c)$.

5 Homomorphisms between Šare systems and Šare's categories

In this section we define homomorphisms between Šare semigroups, as well as between the corresponding quotient semigroups. We also introduce the associated cannonical Šare semigroups. They induce the corresponding categories, that we briefly describe.

5.1 Homomorphisms, embeddings and isomorphisms between Šare systems

Assume that two Šare systems M_{Γ_1} and M_{Γ_2} are generated by two finite and well ordered alphabets Γ_1 and Γ_2. A function $f : M_{\Gamma_1} \to M_{\Gamma_2}$ is said to be *homomrphism* of Šare systems if $f(xy) = f(x)f(y)$ for all $x, y \in M_{\Gamma_1}$, and if $x_1 \sim x_2$ in M_{Γ_1} then $f(x_1) \sim f(x_2)$ in M_{Γ_2}, for any pair (x_1, x_2) (or, equivalently, $[x_1] = [x_2]$ implies that $[f(x_1)] = [f(x_2)]$).

It generates in a natural way a homomorphism $\tilde{f} : \widetilde{M}_{\Gamma_1} \to \widetilde{M}_{\Gamma_2}$ of corresponding quotient semigroups, defined by $\tilde{f}([x]) = [f(x)]$.

If $\tilde{f}$ is a monomorphism of quotient semigroups, then we say that $\widetilde{M}_{\Gamma_1}$ is *embedded* into $\widetilde{M}_{\Gamma_2}$ via $\tilde{f}$, whereby we identify $\widetilde{M}_{\Gamma_1}$ with $\tilde{f}(\widetilde{M}_{\Gamma_1})$, which is a subsemigroup of $\widetilde{M}_{\Gamma_2}$. In this case we can write $\widetilde{M}_{\Gamma_1} \leq \widetilde{M}_{\Gamma_2}$.

The following diagram is commutative, in which $\pi_j : M_{\Gamma_j} \to \widetilde{M}_{\Gamma_j}$ are the canonical projections, defined by $\pi_j(x_j) = [x_j]$, for $x_j \in M_{\Gamma_j}$ and $j = 1, 2$:

$$\begin{array}{ccc} M_{\Gamma_1} & \xrightarrow{f_1} & M_{\Gamma_2} \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi_2} \\ \widetilde{M}_{\Gamma_1} & \xrightarrow{\tilde{f}_1} & \widetilde{M}_{\Gamma_2} \end{array}$$

Analogously for epimorphisms and isomorphisms. It is clear that the two quotient Šare systems are isomorphic if and only if the corresponding alphabets are equipotent. If the quotient semigroups $\widetilde{M}_{\Gamma_1}$ and $\widetilde{M}_{\Gamma_2}$ are isomorphic, we write $\widetilde{M}_{\Gamma_1} \simeq \widetilde{M}_{\Gamma_2}$.

Any monotone function $f_0 : \Gamma_1 \to \Gamma_2$ (monotone with respect to well orderings in the alphabets) induces in a natural way a homomorphism $f : M_{\Gamma_1} \to M_{\Gamma_2}$ defined by $f(a_1 \dots a_k) = f_0(a_1) \dots f_0(a_k)$ for all jorbs $a_1 \dots a_k \in \Gamma_k$ and for all even positive integers k. Monotonicity is needed because of (de)compression axiom (3). It is clear that

$$\tilde{f}(\widetilde{M}_{\Gamma_1}) = \widetilde{M}_{f_0(\Gamma_1)},$$

where $f_0(\Gamma_1)$ is the corresponding subalphabet of Γ_2. And vice versa: any homomorphism $\tilde{f} : \widetilde{M}_{\Gamma_1} \to \widetilde{M}_{\Gamma_2}$ is generated by uniquely determined monotone function $f_0 : \Gamma_1 \to \Gamma_2$. (Recall that the alphabets are assumed to be well ordered.) If the function f_0 is strictly monotone, then the corresponding function $\tilde{f}$ is monomorphic, and we have that $\widetilde{M}_{f_0(\Gamma_1)} \leq \widetilde{M}_{\Gamma_2}$.

Assume that $|\Gamma_1| = |\Gamma_2| = n$. Then, there are precisely two different isomorphisms $\tilde{f} : \widetilde{M}_{\Gamma_1} \to \widetilde{M}_{\Gamma_2}$, generated by one increasing and one decreasing bijection $f_0 : \Gamma_1 \to \Gamma_2$. More generally, we have the following result.

Theorem 4. *Assume that $|\Gamma_1| = n_1$, $|\Gamma_2| = n_2$. If $n_1 \leq n_2$, then there are precisely $2\binom{n_2}{n_1}$ monomrphisms $\tilde{f} : \widetilde{M}_{\Gamma_1} \to \widetilde{M}_{\Gamma_2}$. For any two positive* *integers n_1 and n_2, the number of all homomorphisms $\tilde{f} : \widetilde{M}_{\Gamma_1} \to \widetilde{M}_{\Gamma_2}$ is equal to $2\binom{n_1+n_2-1}{n_1} - n_2$.*

Proof. To each subset of n_1 elements of Γ_2, we can assign uniquely determined (strictly) increasing function $f_0 : \Gamma_1 \to \Gamma_2$ with an image equal to this subset. Therefore the number of increasing functions f_0 is equal to $\binom{n_2}{n_1}$. Analogously for the number of decreasing functions f_0. Since each such function f_0 generates a monomorphism (and all of the monomorphisms are of this form), the first claim follows.

To prove the second claim, we first find the number of nondecreasing functions $f_0 : \Gamma_1 \to \Gamma_2$. Given such f_0, let m_k be the cardinality of the f_0-preimage of the k-th element in Γ_2, where $k = 1, \dots, n_2$. Then $m_1 + \cdots + m_{n_2} = n_1$. The number of nonegative integer solutions $(m_1, \dots, m_{n_2})$ of this equation is equal to the number of combinations with repetition, $\binom{n_1+n_2-1}{n_1}$, and this is the number of homomorphisms $\tilde{f} : \widetilde{M}_{\Gamma_1} \to \widetilde{M}_{\Gamma_2}$ generated by all nondecreasing functions f_0. The number of functions $\tilde{f}$ generated by nonincreasing functions f_0 is the same, and the second claim follows. (Remark that constant functions f_0 were counted twice.) □

We also introduce *cannonical Šare system M_n* defined as the Šare system generated by the usual numerical alphabet $\{1, 2, \dots, n\}$, that is, $M_n := M_{\{1,\dots,n\}}$. It is clear that $M_\Gamma \simeq M_n$ if and only if $|\Gamma| = n$. By $\widetilde{M}_n$ we denote the corresponding cannonical quotient Šare systme, for any $n \in \mathbb{N}$.

We have the natural embedding $i_n : M_n \to M_{n+1}$, generated by the increasing injective map $i_n : \{1, \dots, n\} \to \{1, \dots, n, n+1\}$, where $i_n(j) = j$ for all $j = 1, \dots, n$. In this way, we obtain an infinite sequence of naturally embedded Šare systems,

$$M_1 \le M_2 \le \cdots \le M_n \le M_{n+1} \le \dots$$

Analogously, each induced map $\tilde{i}_n : \widetilde{M}_n \to \widetilde{M}_{n+1}$ is a natural embedding of $\widetilde{M}_n$ into $\widetilde{M}_{n+1}$, so that we have an infinite sequence of subsemigroups:

$$\widetilde{M}_1 \le \widetilde{M}_2 \le \cdots \le \widetilde{M}_n \le \widetilde{M}_{n+1} \le \dots$$

5.2 Šare's categories

We introduce the *Šare category*, denoted by $\mathbf{Ša}(M)$, the objects of which are Šare systems M_Γ for all possible well ordered finite alphabets Γ, while the morphisms between the objects are homomorphisms between Šare systems. The corresponding *quotient Šare category*, denoted by $\check{\mathbf{S}}\mathbf{a}(\widetilde{M})$, consists of all possible quotient semigroups of the form $\widetilde{M}_\Gamma$ as objects (for all possible well ordered finite alphabets Γ), and of all possible homomorphisms between them as morphisms. The associated two *cannonical Šare subcategories* of $\check{\mathbf{S}}\mathbf{a}(M)$ and $\check{\mathbf{S}}\mathbf{a}(\widetilde{M})$ are denoted by $\check{\mathbf{S}}\mathbf{a}(M,\mathbb{N})$ and $\check{\mathbf{S}}\mathbf{a}(\widetilde{M},\mathbb{N})$, respectively. All of them are subcategories of the category $\mathbf{S}$ of all semigroups:

$$\check{\mathbf{S}}\mathbf{a}(M,\mathbb{N}) \subset \check{\mathbf{S}}\mathbf{a}(M) \subset \mathbf{S}, \quad \check{\mathbf{S}}\mathbf{a}(\widetilde{M},\mathbb{N}) \subset \check{\mathbf{S}}\mathbf{a}(\widetilde{M}) \subset \mathbf{S}.$$

We can easily build covariant functors between Šare categories. For example, a natural covariant functor $F : \check{\mathbf{S}}\mathbf{a}(M,\mathbb{N}) \to \check{\mathbf{S}}\mathbf{a}(\widetilde{M},\mathbb{N})$ that we name *Šare's functor*, consisting of the sequence $(\pi_n)_{n\geq 1}$ of projections between the corresponding objects, as well as of the sequence $(f_n, \tilde{f}_n)_{n\geq 1}$ of ordered pairs of the corresponding morphisms (i.e., homomorphisms between Šare systems), is indicated in the following commutative diagram:

$$\begin{array}{ccccccccccc}
M_1 & \xrightarrow{f_1} & M_2 & \xrightarrow{f_2} & \cdots & \xrightarrow{f_{n-1}} & M_n & \xrightarrow{f_n} & M_{n+1} & \xrightarrow{f_{n+1}} & \cdots \\
\downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi_2} & & & & \downarrow{\scriptstyle \pi_n} & & \downarrow{\scriptstyle \pi_{n+1}} & & \\
\widetilde{M}_1 & \xrightarrow{\tilde{f}_1} & \widetilde{M}_2 & \xrightarrow{\tilde{f}_2} & \cdots & \xrightarrow{\tilde{f}_{n-1}} & \widetilde{M}_n & \xrightarrow{\tilde{f}_n} & \widetilde{M}_{n+1} & \xrightarrow{\tilde{f}_{n+1}} & \cdots
\end{array}$$

5.3 Final remarks

In this paper, we have dealt with the set $\mathcal{A}_0$ of three axioms consisting of reversion (1) and two (de)compressions (2) and (3). It generated the Šare system $M_\Gamma = M_\Gamma(\mathcal{A}_0)$ and the quotient system $\widetilde{M}_\Gamma = \widetilde{M}_\Gamma(\mathcal{A}_0)$. In future applications, for some particular classes of problems, it is possible to envisige a different finite set $\mathcal{A}$ of axioms on M_Γ, yielding a new Šare system $M_\Gamma(\mathcal{A})$ and the corresponding quotient system $\widetilde{M}_\Gamma(\mathcal{A})$, as well as the associated Šare categories $\check{\mathbf{S}}\mathbf{a}(M,\mathcal{A})$ and $\check{\mathbf{S}}\mathbf{a}(\widetilde{M},\mathcal{A})$.

6 Appendix

The 24 cases indicated in the proof of Theorem 1 (corresponding to its sufficiency part (a)) are discussed here. We note that their number can be further reduced by half, since the (de)compression axioms (2) and (3) are symmetric in the sense that if, for example, we reverse the order of letters appearing in $\alpha\beta\alpha$, nothing is changed. Hence, Case 1 corresponding to $a \leq b \leq c \leq d$ is equivalent to the case with reverse order, i.e., $d \le c \le b \le a$, etc. Therefore, it suffices consider the 12 cases corresponding to $a \le \dots$ and $b \le \dots$ only, since the cases of the form $b \le \dots$ and $d \le \dots$ have their symmetric counterparts among the preceding two cases.

Case 1. Assume that $a \le b \le c \le d$. See Case 1 in the proof of Theorem 1.

Case 2. Assume that $a \le b \le d \le c$. Condition in (9) reduces to $b = c$, hence, $abcd \sim abbd \sim ad$, since b is between a and d (or, from $b \le d \le c$ it follows that $b = d = c$, so that $abcd = ad^3 \sim ad$, by Lemma 2).

Case 3. Assume that $a \le c \le b \le d$. See Case 3 in the proof of Theorem 1.

Case 4. Assume that $a \le c \le d \le b$. Condition in (9) reduces to $0 = 0$. From $d \in [c, b]$ (in the first equivalence) and $d \in [a, b]$ (in the third equivalence), by using decompression formula (3), we obtain that

$$abcd \sim abd^2cd \sim abd^2 \sim ad^2bd^2 = ad(dbd)d \sim ad^3 \sim ad.$$

Case 5. Assume that $a \le d \le b \le c$. Condition in (9) reduces to $b = d$. Hence, $abcd \sim adcd \sim ad$.

Case 6. Assume that $a \le d \le c \le b$. Condition in (9) reduces to $c = d$, so that (since $d \in [a, b]$, using decompression property (3) in the second equality below, and then compression property (2) in the last equivalence),

$$abcd = abdd \sim addbdd = ad(dbd)d \sim ad^3 \sim ad.$$

Case 7. Assume that $b \le a \le c \le d$. See Case 2 in the proof of Theorem 1.

Case 8. Assume that $b \le a \le d \le c$. Condition in (9) reduces to $a = d$. Since $a \in [b, c]$, we have that (see decompression property (2), used in the equivlanece below)

$$abcd = abca \sim abaaca = (aba)(aca) \sim aa = ad.$$

Case 9. Assume that $b \le c \le a \le d$. Condition in (9) reduces to $a = c$, so that $abcd = abad \sim ad$.

Case 10. Assume that $b \le c \le d \le a$. Condition in (9) reduces to $d = c$, so that

$$abcd = abd^2 \sim ad^2bd^2 = ad(dbd)d \sim ad^3 \sim ad,$$

where we have used that $d \in [b, a]$, along with decompression property (3) and Lemma 2.

Case 11. Assume that $b \leq d \leq a \leq c$. Condition in (9) reduces to $0 = 0$. Since $d \in [b, c]$ (used in the first equality, along with decompression property (3)) and $d \in [b, a]$ (used in the second equality), we have that

$$abcd = abd^2cd \sim abdd \sim ad^2bd^2 = ad(dbd)d \sim ad^3 \sim ad.$$

Case 12. Assume that $b \leq d \leq c \leq a$. Condition in (9) reduces to $0 = 0$. Since $d \in [b, c]$, using decompression property (3), and then using $d \in [b, a]$, we obtain that

$$abcd = abd^2cd = abd^2 \sim ad^2bd^2 = ad(dbd)d \sim ad^3 \sim ad.$$

Acknowledgements

We express our deep gratitude to an anonymous referee for careful reading of the paper and for his useful remarks. We are also grateful to our colleague Damir Kirasić (of the University of Zagreb) for his help during the preparation of this article.

References

[1] Essert, M., I. Kuzmanović, I. Vazler, and T. Žilić, *Theory of M-system*, Log. J. IGPL, Vol. **25** (2017) 5, 836-858.

[2] Howie, J. M., *An Introduction to Semigroup Theory*, Academic Press, London, 1976.

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

[4] Lothaire, M., *Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, Addison-Wesley Publishing Company, London, 1983.

[5] Petrich, M., *Inverse semigroups*, John Wiley and Sons, 1984.

[6] Šare, M., *Algebra, topologija i logika električnih mreža* (Algebra, Topology, and Logic of Electrical Circuits), University of Zagreb (Faculty of Electrical Engineering), Zagreb 1972 (PhD dissertation, in Croatian; depending strongly on [7])

[7] Šare, M., *m-Brojevi* (m-Numbers), University of Zagreb (Faculty of Electrical Engineering), Zagreb 1970 (309 pp, in Croatian)

[8] Šare, M., *Jorbologija* (Jorbology), Element, Zagreb 2000 (in Croatian)

Mario Essert
University of Zagreb, Faculty of Electrical Engineering and Computing,
Department of Applied Mathematics, Unska 3, 10000 Zagreb, Croatia
E-mail address: messert@inet.hr

Darko Žubrinić
University of Zagreb, Faculty of Electrical Engineering and Computing,
Department of Applied Mathematics, Unska 3, 10000 Zagreb, Croatia
E-mail address: darko.zubrinic@gmail.com