A SOLUTION OF THE EQUATION $(x_1^2 \dots x_n^2)^3 = (x_1^3 \dots x_n^3)^2$

ŠTĚPÁN HOLUB

1. Introduction

There exist n-tuples of words $u_1, \dots, u_n$ such that raising to the k-power behaves like an endomorphism, i.e.

$$u_1^k \dots u_n^k = (u_1 \dots u_n)^k. \tag{1.1}$$

This equality is trivial if $u_1, \dots, u_n$ are powers of a common word w, but there are as well non-trivial cases. We shall say that the n-tuple $u_1, \dots, u_n$ is *k-invariant* if it is not trivial and (1.1) holds. One might ask whether there exists a n-tuple that is k-invariant for two different integers $k_1, k_2 > 1$. In [4] it is shown that there exists no k-invariant n-tuple for three different integers greater than one, but for two integers the problem is open. The most simple case is that with $k_1 = 2$, $k_2 = 3$. Note that if (1.1) holds for $k = 2, 3$ than the equality

$$(u_1^2 \dots u_n^2)^3 = (u_1^3 \dots u_n^3)^2 \tag{1.2}$$

holds as well.

It is not difficult, using the computer, to find a lot of 2-invariant or 3-invariant n-tuples, but it turns out that it is not so easy to find a n-tuple realizing the equality (1.2). In this paper we describe a method that allows to get such a n-tuple in rather short time.

2. Classical method

In this section we shall describe the classical method of searching a solution of an equation (cf. e.g. [1]). First let us introduce basic notions.

Let Σ be a finite *alphabet*. Elements of Σ are called *letters* and sequences of letters are called *words*. The sequence of length zero is called the *empty word*. The set of all words (all non-empty words) is denoted by Σ^* (Σ^+, resp.). It is a monoid (semigroup, resp.) under the operation of concatenation.

Let T be a finite set of unknowns. Every

$$(e, e') \in T^+ \times T^+$$

we shall call an *equation* in unknowns from T.

We shall say that a morphism $\varphi : T^+ \to \Sigma^+$ is a *solution* of an equation $(e, e') \in T^+ \times T^+$ in the semigroup Σ^+ if and only if the equality $\varphi(e) = \varphi(e')$ holds.

We shall say that a solution $\varphi : T^+ \to \Sigma^+$ is *cyclic* if and only if there exists a word $v \in \Sigma^+$ such that $\varphi(x)$ is a power of v for every $x \in T$

Now let us have the equation

$$(x_1^2 \dots x_n^2)^3 = (x_1^3 \dots x_n^3)^2 \tag{2.1}$$

and $\varphi : X^+ \to \Sigma^+$, $X = \{x_1, \ldots, x_n\}$, be a solution of it in Σ^+. Denote $\varphi(x_i) = u_i$, $d_i = |u_i|$, $1 \le i \le n$, $d = \sum_{i=1}^n d_i$ and

$$w = \varphi((x_1^2 \ldots x_n^2)^3) = \varphi((x_1^3 \ldots x_n^3)^2). \tag{2.2}$$

The word w is a sequence $y_1 \ldots y_t$, with $t = 6d$, where y_i, $1 \le i \le t$, is a variable for a letter from the alphabet Σ. The equality (2.2) imposes certain identifications on y_i.

Definition 2.1. Let $\approx$ be the smallest equivalence relation on the set $\{y_1, \ldots, y_t\}$ such that

1. $y_i \approx y_j$ if $i \equiv j \mod 2d$
2. $y_i \approx y_j$ if $i \equiv j \mod 3d$
3. $y_i \approx y_j$ if
$$2\sum_{k=i}^{i-1} d_k < i, j \le 2\sum_{k=i}^{i} d_k$$
for some $1 \le k \le n$, and $i \equiv j \mod d_k$
4. $y_i \approx y_j$ if
$$3\sum_{k=i}^{i-1} d_k < i, j \le 3\sum_{k=i}^{i} d_k$$
for some $1 \le k \le n$, and $i \equiv j \mod d_k$

Items 1. and 3. express the fact that y_i and y_j are the same letter of a word u_i thanks to the left side of the equation (2.1), while items 2. and 4. express the same fact deduced from the right side.

We say that the n-tuple $(d_1, \ldots, d_n)$ is the *type* of the solution φ. Clearly the equivalence $\approx$ is defined by the type of the solution. Certainly for any solution of the type $(d_1, \ldots, d_n)$ the equality $y_i = y_j$ must hold as soon as $y_i \approx y_j$.

Definition 2.2. We shall say that φ is the *canonical solution* of the type $(d_1, \ldots, d_n)$ if and only if

$$y_i = y_j \Leftrightarrow y_i \approx y_j.$$

The number of classes of equivalence $\approx$ we shall call the *rank* of the canonical solution.

We say that the $(n, \max(d_1, \ldots, d_n))$ is the size of the type $(d_1, \ldots, d_n)$. We say that the type with size (n, m) is smaller than the type with the size (n', m') if and only if $n \le n'$ and $m \le m'$. It is a partial ordering of canonical solutions.

Observation 2.3. It is sufficient to restrict ourselves to the types $(d_1, \ldots, d_n)$ with $\gcd(d_1, \ldots, d_n) = 1$ (greatest common divisor). Indeed let us suppose that

$$\gcd(d_1, \ldots d_n) = g > 1.$$

It is easy to see that the canonical solution of the type $(d_1, \ldots, d_n)$ results from the canonical solution of the type $(d_1/g, \ldots, d_n/g)$ substituting each letter by a word of length g. Furthermore two words substituted for two different letters have no common letter.

Observation 2.4. Every canonical solution that is minimal non cyclic solution in respect to the above ordering, is of rank two. Indeed, it must be of the rank at least two to be non cyclic and if it was of the rank more than three we would get a smaller non-trivial solution by a morphism mapping one of the letters to the empty word.

Definition 2.5. We say that a canonical solution φ is *antihomogeneus* if it is of rank two and every word u_i contains both letters.

Using the facts described in this section we can search for a non-trivial solution of (2.1) taking different types and constructing their canonical solutions.

3. The weak equivalence

Clearly the time necessary to check all types smaller than a given one grows exponentially with both parameters of the size. Using the following method it is possible to check in the same time all solutions of double length with restriction to antihomogeneus ones [1] .

Let φ be a non-cyclic antihomogeneus solution of (2.1) of rank two in $\{A, B\}^+$. (We can understand the letters A, B as names of the two equivalence classes of $\approx$.) Denote by a_i the number of occurrences of the letter A in the word u_i. The equivalence $\approx$ for the type $(a_1, \dots, a_n)$ has obviously just one class. Now we shall define a bit "weaker" equivalence $\sim$ on $\{y'_1, \dots, y'_s\}$, with $s = \sum_{i=1}^n a_i$. The equivalence $\approx$ is generated by the relation "$y_i \approx y_j$, if they are the same letter in a word u_i". The equivalence $\sim$ is going to be generated by the relation "$y_i \sim y_j$, if $y_i y_{i+1}$ is the same word (of length two) as $y_j y_{j+1}$" [2] .

That relation does not unify the last letter of a word u_i followed by another copy of u_i (we shall call such a letter *a final letter of the first type*) with the last letter of the same word followed by the word u_{i+1} (*a final letter of the second type*) [3] . The precise definition of $\sim$ is as follows.

Definition 3.1. Let $\sim$ be the smallest equivalence relation on the set $\{y'_1, \dots, y'_s\}$ such that

1. $y_i \approx y_j$ if $i \equiv j \mod 2d$
2. $y_i \approx y_j$ if $i \equiv j \mod 3d$
3. $y_i \approx y_j$ if
$$2\sum_{k=i}^{i-1} d_k < i, j < 2\sum_{k=i}^{i} d_k$$
for some $1 \le k \le n$ and $i \equiv j \mod d_k$
4. $y_i \approx y_j$ if
$$3\sum_{k=i}^{i-1} d_k < i, j < 3\sum_{k=i}^{i} d_k$$
for some $1 \le k \le n$ and $i \equiv j \mod d_k$

If $y'_i \sim y'_j$ then corresponding letters A are in w followed by the same power of letter B (we define B^0 to be the empty word). If the word w contained only one maximal power of the letter B, it would be cyclic. It implies that if φ is a non-cyclic canonical solution then the equivalence $\sim$ has at least two classes.

[1] Such a restriction is not fatal as it can be proved (see [3]) that if there exists a non-trivial 2,3-invariant n-tuple then the minimal one is antihomogeneus.

[2] Further we consider the word w in the cyclic way.

[3] We identify u_{n+1} with u_1.

Now the question is whether having a type $(a_1, \ldots, a_n)$ with the equivalence $\sim$ of at least two classes, we can construct a non-cyclic solution of the equation (2.1). Let us suppose that we have such a solution, with a_i being the number of occurrences of the letter A in the word u_i.

Denote by $\eta_{i,i}$, $1 \le i \le n$, the number of letters B that are placed between the last and the first letter A of the word u_i in the word $u_i u_i$, and by $\eta_{i,i+1}$, $1 \le i \le n$, the same for the edge between words u_i and u_{i+1}.[3] The equivalence $\sim$ generates an equivalence between numbers $\eta_{i,j}$. Denote that equivalence $\bowtie$.

Denote by α_i (β_i), $1 \le i \le n$, the number of letters B on the beginning (on the end resp.) of the word u_i. We get the following system of $2n$ linear equations

$$\begin{aligned} \alpha_i + \beta_i &= \eta_{i,i}, & 1 \le i \le n \\ \alpha_i + \beta_{i+1} &= \eta_{i,i+1}, & 1 \le i < n \\ \alpha_1 + \beta_n &= \eta_{n,n+1} & \end{aligned} \tag{3.1}$$

The matrix of the system (3.1) is not regular and it is easy to see that a sufficient condition for the existence of a solution is that in each class of the equivalence $\bowtie$ the number of variables of the type $e_{i,i}$ is the same as the number of variables of the type $e_{i,i+1}$. In such a case there exists a solution such that $e_{i,j}$ are non-negative integers, $e_{i,j} = e_{i',j'}$ holds if and only if $e_{i,j} \bowtie e_{i',j'}$, and all α_i, β_i, $1 \le i \le n$ are non-negative integers. It means that it is enough to prove that in each class of the equivalence $\sim$ the number of representatives among the final letters of the first type is the same as the number of representatives among the final letters of the second type.

To prove this, fix C, an equivalence class of $\sim$. Denote by a the number of non-final letters contained in C, by f_1 the number of the final letters of the first type and by f_2 the number of the final letters of the second type contained in C. Now we shall count $|w|_C$, the total number of occurrences of elements from C in the word w. Looking at the left side of the equation (2.1) we get

$$|w|_C = 6a + 3f_1 + 3f_2,$$

while looking at the right side of the equation we get

$$|w|_C = 6a + 4f_1 + 2f_2.$$

From that we deduce $f_1 = f_2$, q.e.d.

4. Conclusion

We can conclude that each type $(a_1, \ldots, a_n)$ having at least two classes of the weak equivalence $\sim$, generates a non cyclic solution of the equation (2.1). The method of the weak equivalence can be of course used even for other similar equations.

Using the method we have found following solution of the equation (2.1):

$n = 7$	$u_4 = ABA$
$u_1 = ABABAABABAABAABABA$	$u_5 = BAABA$
$u_2 = ABABA$	$u_6 = ABABA$
$u_3 = ABAAB$	$u_7 = ABABAABAABABAABABA$

It is the canonical solution of the type (18,5,5,3,5,5,18) and it was constructed using the weak equivalence of the type (7,2,2,1,2,2,7).

References

1. K. I. Appel, F. M. Djorup, *On the equation $z_1^n z_2^n \dots z_k^n = y^n$ in a free semigroup*, Trans. Am. Math. Soc. **134** (1968), pp. 461–470.
2. M. Lothaire, *Combinatorics on words*, Cambridge University Press, 1983.
3. Š. Holub, *O rovnicích $(x_1^s \dots x_n^s)^r = (x_1^r \dots x_n^r)^s$ ve volných pologrupách*, M.D. thesis, Charles University, Prague, 1998.
4. Š. Holub, *Local and global periodicity*, submitted, 1998.

Department of mathematics, Charles University, Sokolovská 83, Prague 18000
E-mail address: holub@karlin.mff.cuni.cz