RAIRO-Theor. Inf. Appl. 45 (2011) 225–234
DOI: 10.1051/ita/2011018

Available online at:
www.rairo-ita.org

UNIQUE DECIPHERABILITY IN THE ADDITIVE MONOID OF SETS OF NUMBERS *, **

Aleksi Saarela [1]

Abstract. Sets of integers form a monoid, where the product of two sets A and B is defined as the set containing $a + b$ for all $a \in A$ and $b \in B$. We give a characterization of when a family of finite sets is a code in this monoid, that is when the sets do not satisfy any nontrivial relation. We also extend this result for some infinite sets, including all infinite rational sets.

Mathematics Subject Classification. 68R05, 68Q45.

1. Introduction

The product of two languages A and B is defined as the language containing all words uv, where $u \in A$ and $v \in B$. Then the set of all languages is a monoid. Some problems that are easy for words are very hard in this monoid of languages. For example, if $xy = yx$ for two words x, y, then x and y are powers of a common word, but no similar result holds for languages. In fact, the maximal language commuting with a given finite language is not necessarily even recursively enumerable [7]. As another example, it is undecidable whether $AB^iC = DE^iF$ for all i, where A, B, C, D, E, F are given finite sets [6].

We will study some problems in the case when the languages are unary. The monoid of unary languages is isomorphic to the additive monoid of sets of natural numbers, so we will actually formulate everything in terms of sets of numbers.

Keywords and phrases. Unique decipherability, rational set, sumset.

* *Supported by the Turku University Foundation.*

** *Supported by the Academy of Finland under grant 121419.*

[1] Department of Mathematics and Turku Centre for Computer Science TUCS, University of Turku, 20014 Turku, Finland; amsaar@utu.fi

Article published by EDP Sciences

A set of words A is said to be a code, if no word of A^* has two different representations as products of elements of A. Then it is also said that A has the unique decipherability property. Codes have been studied a lot, see *e.g.* [1], and they are fundamentally important in message transmission.

The notion of unique decipherability can be extended to other monoids, *e.g.* to the monoid of languages. We are interested in this problem for unary languages, that is in the additive monoid of sets of integers, where the product of two sets A and B is defined as the set containing all sums $a+b$, where $a \in A$ and $b \in B$. Now a family of sets can be defined to be a code, if no set has two different representations as a product of these sets. Because the monoid is commutative, representations are not considered different if they differ only by the order of the sets.

The problem of determining whether a given family of finite sets of natural numbers is a code was proved to be decidable in [3]. We will extend this result by giving a complete characterization of these codes, and by generalizing it for some infinite sets. We will also study the power equality problem, that is the problem of determining whether some powers of two sets are equal.

We will begin in Section 2 by giving the required definitions and by proving some results about powers of sets of integers. These results are related to the Frobenius problem, see *e.g.* [9] for a survey [5] for a generalization for words and [4] for related algebraic results. The main result of this section is that if the elements of a set do not have a common divisor, then sufficiently large powers of the set contain almost all integers between their minimums and maximums. This result is very important in the later sections.

In Section 3 we consider the power equality problem. For example, we show that it is sufficient to consider powers that are of linear size with respect to the maximum of the sets. The results in this section form the basis for the solution of the unique decipherability problem, and are also of independent interest.

In Section 4 we give a characterization of codes in the additive monoid of finite sets of integers. In particular, we prove that a family of three sets is never a code, *i.e.* three sets always satisfy a nontrivial relation. We prove a similar result for certain infinite sets, including all infinite rational sets.

2. Additive powers of a set

Let M be a monoid. The subsets of M form a monoid, where the product of two sets $A, B \subseteq M$ is defined to be $AB = \{uv : u \in A, v \in B\}$. We are interested in the case of unary languages, that is the case of $M = \{a\}^*$, where a is a letter. This monoid M is isomorphic with the additive monoid of nonnegative integers $\mathbb{N}_0$, where the isomorphism is $a^k \mapsto k$. Also the monoid of unary languages is isomorphic with the monoid of sets of nonnegative integers, where the isomorphism is $\{a^{k_1}, a^{k_2}, \ldots\} \mapsto \{k_1, k_2, \ldots\}$. Thus we will formulate everything in terms of sets of numbers. Often we can allow the sets to contain also negative integers. We will mostly consider finite sets.

If $m, n \in \mathbb{Z}$, $k \in \mathbb{N}_0$ and $A, B \subseteq \mathbb{Z}$, then we use the following notation:

$$\begin{aligned}
[m,n] &= \{a \in \mathbb{Z} : m \le a \le n\}, \\
[m,\infty) &= \{a \in \mathbb{Z} : a \ge m\}, \\
(-\infty,n] &= \{a \in \mathbb{Z} : a \le n\}, \\
AB &= \{a+b : a \in A, b \in B\}, \\
A^k &= \{a_1 + \cdots + a_k : a_1, \ldots, a_k \in A\}, \\
A^* &= \bigcup_{k=0}^{\infty} A^k, \\
A+n &= \{a+n : a \in A\}, \\
A \cdot n &= \{an : a \in A\}, \\
A/n &= \{a/n : a \in A\}.
\end{aligned}$$

We will often need to assume that the elements of a set do not have a common divisor, or that the minimum of a set is zero. Thus we let

$$S_n = \{A \subseteq [0,n] : 0, n \in A, \gcd A = 1\}.$$

If $A \in S_n$, then let $\widetilde{A} = \{n - a : a \in A\}$ be the "reverse" of A. Now $\widetilde{AB} = \widetilde{A}\widetilde{B}$.

Let $A = \{0, a_1, \ldots, a_r\} \subset \mathbb{N}_0$ and $\gcd A = 1$. It is well known that every sufficiently large integer can be represented in the form

$$a_1 x_1 + \cdots + a_r x_r, \tag{2.1}$$

where $x_1, \ldots, x_r \in \mathbb{N}_0$. The Frobenius problem asks, what is the largest integer that cannot be represented in this way. This integer is called the Frobenius number of A and we denote it by $G(A)$. The numbers (2.1) form the set A^*, so $G(A)$ is the largest integer not in A^*.

We define $F_m(A)$ to be the smallest integer such that

$$A^* \cap [0,m] \subseteq A^{F_m(A)}.$$

We assume that $0 \in A$, so $A \subseteq A^2 \subseteq A^3 \subseteq \ldots$ and $F_m(A)$ exists for every m. The number $F_m(A)$ tells how large the coefficients $x_1, \ldots, x_r$ need to be: if $n \le m$ and n has a representation of the form (2.1), then n has such a representation, where $x_1 + \cdots + x_r \le F_m(A)$.

There are many results concerning the size of the Frobenius number. We use the following result from [2].

Lemma 2.1. *If $A = \{a_0, \ldots, a_r\} \in S_n$, where $0 = a_0 < \cdots < a_r = n$, then $G(A) \le a_1 n - a_1 - n \le n^2 - 2n$.*

We also need an upper bound for $F_m(A)$.

Lemma 2.2. *If $A = \{a_0, \dots, a_r\} \in S_n$, where $0 = a_0 < \cdots < a_r = n$, then $F_m(A) \le n - 1 + m/n$.*

Proof. Let $g = G(A)$ and $a \in A^* \cap [0, m]$. If $a \le g + n$, then $a \in A^k$, where $k = \lfloor (g+n)/a_1 \rfloor$. If $a > g+n$, then $a = g+i+n(a-g-i)/n$, where $i \in \{1, \dots, n\}$ is such that $g + i \equiv a \mod n$. Now $g + i \in A^k$ and $n(a-g-i)/n \in A^l$, where again $k = \lfloor (g+n)/a_1 \rfloor$ and $l = (a-g-i)/n$, and thus $a \in A^{k+l}$. So with the help of Lemma 2.1 we get the result

$$F_m(A) \le k + l \le \frac{g+n}{a_1} + \frac{a-g-1}{n} \le n - 1 + \frac{m}{n}. \qquad \square$$

Next we examine the structure of A^k for large k. If $A \in S_n$, then $A^k \subseteq [0, kn]$. Because A^* contains almost every natural number, A^k contains almost every number from the interval $[0, kn]$. Only some numbers from the beginning and from the end are missing. These missing numbers will be essentially the same for all large values of k (of course the large missing numbers will be getting larger and larger as k grows). This is formalized by the following theorem.

Theorem 2.3. *Let $A \in S_n$, $C = A^* \cap [0, n^2 - 2n]$ and $\widetilde{D} = (\widetilde{A})^* \cap [0, n^2 - 2n]$. Now*

$$A^k = C \cup \left[n^2 - 2n + 1, kn - n^2 + 2n - 1\right] \cup (D + kn - n^2 + 2n)$$

for all $k \ge 2n - 2$.

Proof. Let $k \ge 2n - 2$. By Lemma 2.2,

$$F_{\lfloor kn/2 \rfloor}(A), F_{\lfloor kn/2 \rfloor}\left(\widetilde{A}\right) \le n - 1 + k/2 \le k.$$

Now we get

$$\begin{aligned} A^k \cap \left[0, \lfloor kn/2 \rfloor\right] &= A^* \cap \left[0, \lfloor kn/2 \rfloor\right] \\ &= \left(A^* \cap \left[0, n^2 - 2n\right]\right) \cup \left(A^* \cap \left[n^2 - 2n + 1, \lfloor kn/2 \rfloor\right]\right) \\ &= C \cup \left[n^2 - 2n + 1, \lfloor kn/2 \rfloor\right]. \end{aligned}$$

Here the first equality holds, because $k \ge F_{\lfloor kn/2 \rfloor}(A)$, and the last equality follows from Lemma 2.1. Similarly we get $\widetilde{A}^k \cap [0, \lfloor kn/2 \rfloor] = \widetilde{D} \cup [n^2 - 2n + 1, \lfloor kn/2 \rfloor]$. The claim follows. $\square$

3. Power equality problem

We will study the power equality problem, that is the problem of determining whether some powers of two finite sets $A, B \subset \mathbb{Z}$ are equal. The following lemma tells how this problem can be reduced to the case where $\min A = \min B = 0$.

Lemma 3.1. *Let* $\min A_i = m_i < \max A_i = n_i$, $A_i = B_i + m_i$, *and* $k, l > 0$. *Now* $A_1^k = A_2^l$ *if and only if* $B_1^k = B_2^l$ *and* $m_1 n_2 = m_2 n_1$.

Proof. The sets A_1^k and A_2^l are equal if and only if

$$B_1^k + km_1 = B_2^l + lm_2. \tag{3.1}$$

If the sets A_1^k and A_2^l are equal, then their minimums and maximums are equal, that is $km_1 = lm_2$ and $kn_1 = ln_2$. From this and (3.1) it follows that $B_1^k = B_2^l$ and $m_1 n_2 = m_2 n_1$.

On the other hand, if $B_1^k = B_2^l$, then $k(n_1 - m_1) = \max B_1^k = \max B_2^l = l(n_2 - m_2)$. Multiplying this by $m_1 m_2$ gives $km_1(m_2 n_1 - m_1 m_2) = lm_2(m_1 n_2 - m_1 m_2)$. If $m_1 n_2 = m_2 n_1$, then $km_1 = lm_2$ and (3.1) holds. □

It is clear that if $\min A = \min B = 0$, then some powers of A and B can be equal only if $\gcd A = \gcd B = d$, and if this is the case, then $A^k = B^l$ if and only if $(A/d)^k = (B/d)^l$. Thus we can assume that $d = 1$.

Example 3.2. Let $A^2 = B^2$. If the two smallest elements of A are 0 and a, then the two smallest elements of A^2 are also 0 and a. Thus 0 and a must also be the two smallest elements of B. Similarly the two largest elements of A and B must be the same.

The example of $A \neq B$, $A^2 = B^2$, where the largest element of A is as small as possible, is $A = \{0, 1, 3, 4\}, B = \{0, 1, 2, 3, 4\}$ (or *vice versa*). In this case $A^2 = B^2 = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.

The sets A, B can also be selected to be maximal in the sense that they are not proper subsets of any set D such that $A^2 = D^2$. For example, if

$$A = \{0, 1, 3, 7, 8, 9\},\ B = \{0, 1, 3, 6, 8, 9\},\ C = \{0, 1, 2, 6, 8, 9\},$$

then $A^2 = B^2 = C^2 \neq D^2$ for all D such that $A \subset D$, $B \subset D$ or $C \subset D$.

Theorem 3.3. *Let* $m \leq n$, $A \in S_m$ *and* $B \in S_n$. *There are* $i, j > 0$ *such that* $A^i = B^j$ *if and only if*

$$\begin{aligned} &A^k \cap [0, n^2 - 2n] = B^k \cap [0, n^2 - 2n] \quad \textit{and} \\ &\widetilde{A}^k \cap [0, n^2 - 2n] = \widetilde{B}^k \cap [0, n^2 - 2n], \end{aligned} \tag{3.2}$$

where $k = 2n - 2$.

Proof. If $A^i = B^j$, then $im = jn$ and $A^{ki} = B^{kj}$ for all k. Thus there are such i, j if and only if there are such $i, j \geq 2n - 2$. If $A^i = B^j$, where $i, j \geq 2n - 2$, then (3.2) holds by Lemma 2.2.

Next, assume that (3.2) holds. Consider two arbitrary integers $i, j \geq 2n - 2$ satisfying $im = jn$. Then $A^i = B^j$ by Theorem 2.3. □

Theorem 3.3 gives a condition for the existence of the required numbers i and j, and this leads to an algorithm for solving the power equality problem. The next

theorem gives a similar condition, which is perhaps not as useful algorithmically, but may be easier in some other ways.

Theorem 3.4. *Let $A \in S_m$ and $B \in S_n$. There are $i, j > 0$ such that $A^i = B^j$ if and only if $A^* = B^*$ and $(\widetilde{A})^* = (\widetilde{B})^*$.*

Proof. If $m \le n$, $A^i = B^j$ and $C \in \{A, B, \widetilde{A}, \widetilde{B}\}$, then

$$C^* = \left(C^{2n-2} \cap \left[0, n^2 - 2n\right]\right) \cup \left[n^2 - 2n + 1, \infty\right)$$

by Lemmas 2.1 and 2.2. Now $A^* = B^*$ and $(\widetilde{A})^* = (\widetilde{B})^*$ by Theorem 3.3. On the other hand, if $A^* = B^*$ and $(\widetilde{A})^* = (\widetilde{B})^*$, then (3.2) holds by Lemma 2.2. □

We can use Theorem 3.3 to prove that if $A^k = B^k$ holds for some k, then it holds for all sufficiently large k. We are not aware whether $A^k = B^k$ implies $A^{k+1} = B^{k+1}$.

Theorem 3.5. *If $A, B \in S_n$ and $A^k = B^k$ for some $k > 0$, then $A^k = B^k$ for all $k \ge 2n - 2$.*

Proof. If $A^k = B^k$ for some k, then by Theorem 3.3 equation (3.2) holds for $k = 2n - 2$, and by Lemma 2.2 it holds for all larger k as well. The claim now follows from Theorem 2.3. □

Theorem 3.5 raises the following question: if n is fixed, then what is the smallest number m such that if $A, B \in S_n$ and $A^k = B^k$ for some $k > 0$, then $A^k = B^k$ for all $k \ge m$? Theorem 3.5 proves that $m \le 2n - 2$, and the following example proves that $m \ge n - 2$.

Example 3.6. Let $A = \{0, 1, n-1, n\}$. Now $A^{n-3} \ne [0, n]^{n-3}$, but $A^{n-2} = [0, n]^{n-2}$. The inequality holds, because $n-2 \notin A^{n-3}$. The equality holds, because every element of $[0, n]^{n-2} = [0, (n-2)n]$ is of the form $an + b$, where $a \in [0, n-3]$ and $b \in [0, n]$, and if $a + b \le n - 2$, then $an + b \in \{0, 1, n\}^{n-2}$, and if $a + b > n - 2$, then $an + b = (a + b - n + 1)n + (n - b)(n - 1) \in \{0, n-1, n\}^{n-2}$.

4. Unique decipherability problem

In this section we will study the unique decipherability problem in the monoid of sets of integers. The motivation for the terms "code" and "decipherability" comes from the theory of languages. There, a language A is called a code, if the following holds: if $u_1, \ldots, u_m, v_1, \ldots, v_n \in A$ and $u_1 \ldots u_m = v_1 \ldots v_n$, then $m = n$ and $u_i = v_i$ for all i. A good reference on the theory of codes is [1].

In the commutative monoid of sets of integers the definition of a code can be written as follows. A family of sets $\{A_1, \ldots, A_s\}$ is a *code*, or has the *unique decipherability property*, if no set has two essentially different representations as a product of these sets, or more formally if there are no numbers $k_1, \ldots, k_s, l_1, \ldots, l_s$ such that $A_1^{k_1} \ldots A_s^{k_s} = A_1^{l_1} \ldots A_s^{l_s}$ and $k_i \ne l_i$ for some i.

Theorem 4.1. *Let $A_1, \dots, A_s$ be finite sets of integers. Let $\min A_i = m_i$ and $\max A_i = n_i$. The sets A_i form a code if and only if $s = 1$ and $A_1 \neq \{0\}$ or $s = 2$ and $m_1 n_2 \neq m_2 n_1$.*

Proof. Let $A_1^{k_1} \dots A_s^{k_s} = A_1^{l_1} \dots A_s^{l_s}$. The minimums and maximums of these sets must be the same, that is

$$\begin{aligned} m_1(k_1 - l_1) + \cdots + m_s(k_s - l_s) &= 0 \quad \text{and} \\ n_1(k_1 - l_1) + \cdots + n_s(k_s - l_s) &= 0. \end{aligned}$$

This can be viewed as a pair of equations with s unknowns $k_i - l_i$ and coefficients m_i, n_i. This pair of equations has nontrivial solutions if and only if the rank of the matrix

$$\begin{pmatrix} m_1 & \dots & m_s \\ n_1 & \dots & n_s \end{pmatrix}$$

is smaller than s. The rank is s if and only if $s = 1$ and $A_1 \neq \{0\}$ or $s = 2$ and $m_1 n_2 \neq m_2 n_1$.

If the rank is s, then necessarily $k_i = l_i$ for all i. This means that the sets A_i form a code.

If the rank is smaller than s, then we can select the numbers k_i and l_i to be positive integers so that $k_j \neq l_j$ for some j. Let $A_i = B_i + m_i$. Let $d = \gcd(B_1 \cup \cdots \cup B_s)$ and $C_i = B_i/d$. If $D = C_1^{k_1} \dots C_s^{k_s}$ and $E = C_1^{l_1} \dots C_s^{l_s}$, then $D^* = (C_1 \cup \cdots \cup C_s)^* = E^*$ and $(\widetilde{D})^* = (\widetilde{C}_1 \cup \cdots \cup \widetilde{C}_s)^* = (\widetilde{E})^*$. Now from Theorem 3.4 it follows that $D^k = E^l$ for some k, l. Because

$$\begin{aligned} d \max D &= k_1(n_1 - m_1) + \cdots + k_s(n_s - m_s) \\ &= l_1(n_1 - m_1) + \cdots + l_s(n_s - m_s) = d \max E, \end{aligned}$$

it must be $k = l$. This means that

$$\begin{aligned} \left(A_1^{k_1} \dots A_s^{k_s}\right)^k &= \left(B_1^{k_1} \dots B_s^{k_s}\right)^k + k(k_1 m_1 + \cdots + k_s m_s) \\ &= D^k \cdot d + k(k_1 m_1 + \cdots + k_s m_s) \\ &= E^k \cdot d + k(l_1 m_1 + \cdots + l_s m_s) \\ &= \left(B_1^{l_1} \dots B_s^{l_s}\right)^k + k(l_1 m_1 + \cdots + l_s m_s) = \left(A_1^{l_1} \dots A_s^{l_s}\right)^k \end{aligned}$$

and the sets A_i do not form a code. □

A subset of a monoid is *rational*, if it is obtained from finite sets by repeatedly using the operations of union, product and star. In other words, all finite sets are rational, and if A and B are rational, so are $A \cup B$, AB and A^*. In the case of the additive monoid $\mathbb{N}_0$, a set $A \subseteq \mathbb{N}_0$ is rational if and only if it is ultimately periodic, that is if there are finite sets B, C and a number n such that $A = B \cup C\{n\}^*$.

We have given a characterization of all codes in the additive monoid of finite sets of integers. Next it would be natural to study the unique decipherability problem

for rational sets. We can indeed generalize Theorem 4.1, and the condition we need is actually weaker than rationality: some power of some set must contain an infinite rational set. The next lemma gives some equivalent conditions.

Lemma 4.2. *Let $A \subset \mathbb{Z}$ be infinite and $\min A > -\infty$. The following are equivalent:*

(i) *A^k contains an infinite rational set for some k;*
(ii) *A^k contains an infinite arithmetic progression for some k.*

If these conditions hold and $0 \in A$, then $\{\gcd A\}^* \setminus A^k$ is finite for some k. If also $\min A = 0$, then $A^* = A^l$ for some l (this is called the finite power property*).*

Proof. Every arithmetic progression is a rational set, and every infinite rational set contains an infinite arithmetic progression, so (i) and (ii) are equivalent.

Let $0 \in A$ and let $a, b \in \mathbb{Z}$ be such that $a + bn \in A^k$ for every $n \geq 0$. Because A^* contains all sufficiently large multiples of $\gcd A$, there are numbers c, l such that every multiple of $\gcd A$ that is in the interval $[c+1, c+b]$ is also in A^l. Now A^{l+k} contains every number that is greater than $a + c$ and divisible by $\gcd A$.

Let $\min A = 0$ and let $\{\gcd A\}^* \setminus A^k$ be finite. Now $A^k \subseteq A^{k+1} \subseteq A^{k+2} \subseteq \cdots \subseteq \{\gcd A\}^*$ and only finitely many of these inclusions can be proper, so $A^l = A^{l+1} = A^{l+2} = \cdots = A^*$ for some l. □

It is not necessary for any of the conditions in Lemma 4.2 to hold for $k = 1$. For example, if A is the set of all squares, then $A^4 = \mathbb{N}_0 = A^*$ by Lagrange's four-square theorem.

Theorem 4.3. *Let $A_1, \ldots, A_s$ be sets of integers. Let $\min A_i = m_i > -\infty$ for all i. Let A_1^k contain an infinite arithmetic progression for some k. The sets A_i form a code if and only if $s = 1$ and $m_1 \neq 0$.*

Proof. If $s = 1$ and $m_1 = 0$, then $A_1^l = A_1^* = A_1^{l+1}$ for some l by Lemma 4.2. If $s = 1$ and $m_1 \neq 0$, then $A_1^k \neq A_1^l$ for all $k \neq l$, because $\min A_1^k = km \neq lm = \min A_1^l$ for all $k \neq l$.

Let $s \geq 2$. There are $k_1, k_2, l_1, l_2 > 0$ such that $k_1 m_1 + k_2 m_2 = l_1 m_1 + l_2 m_2$, but $k_1 \neq l_1$ or $k_2 \neq l_2$. Let $A_i = B_i + m_i$. Now B_1^k contains an infinite arithmetic progression, and the same is true for the sets $(B_1^{k_1} B_2^{k_2})^k$ and $(B_1^{l_1} B_2^{l_2})^k$. By Lemma 4.2, there is a number l such that $(B_1^{k_1} B_2^{k_2})^* = (B_1^{k_1} B_2^{k_2})^l$ and $(B_1^{l_1} B_2^{l_2})^* = (B_1^{l_1} B_2^{l_2})^l$. Also $(B_1^{k_1} B_2^{k_2})^* = (B_1 \cup B_2)^* = (B_1^{l_1} B_2^{l_2})^*$. Now

$$\begin{aligned}
\left(A_1^{k_1} A_2^{k_2}\right)^l &= \left(B_1^{k_1} B_2^{k_2}\right)^l + l(k_1 m_1 + k_2 m_2) \\
&= \left(B_1^{k_1} B_2^{k_2}\right)^* + l(k_1 m_1 + k_2 m_2) \\
&= \left(B_1^{l_1} B_2^{l_2}\right)^* + l(l_1 m_1 + l_2 m_2) \\
&= \left(B_1^{l_1} B_2^{l_2}\right)^l + l(l_1 m_1 + l_2 m_2) = \left(A_1^{l_1} A_2^{l_2}\right)^l
\end{aligned}$$

and the sets A_i do not form a code. □

In Theorem 4.3 we assumed that every infinite set is one-way infinite, *i.e.* has a finite minimum. The case where every set has a finite maximum is of course symmetric. We can also consider the two-way infinite case, *i.e.* the case when at least one set has arbitrarily large elements, and at least one (possibly the same) has arbitrarily small elements. This is done in the following theorem.

Theorem 4.4. *Let A and B be (not necessarily distinct) infinite sets of integers. Let the sets $(A \cap [0,\infty))^k$ and $(B \cap (-\infty,0])^k$ contain infinite arithmetic progressions for some k. The sets A, B do not form a code.*

Proof. Now $(AB)^k$ contains increasing and decreasing infinite arithmetic progressions. Let $m \in (AB)^k$ and $C + m = (AB)^k$. Let a, b be such that

$$\gcd C = \gcd(C \cap [a,\infty)) = \gcd(C \cap (-\infty, b]).$$

By Lemma 4.2, there is a number l_1 such that $(C \cap [a,\infty))^{l_1}$ contains all but finitely many of the positive numbers divisible by $\gcd C$. Similarly, there is a number l_2 such that $(C \cap (-\infty, b])^{l_2}$ contains all but finitely many of the negative numbers divisible by $\gcd C$. If $l > l_1, l_2$, then $C^l = \{\pm \gcd C\}^*$. Now

$$\begin{aligned}((AB)^k)^{l+\gcd C} &= C^{l+\gcd C} + ml + m \gcd C = \{\pm \gcd C\}^* + ml + m \gcd C \\ &= \{\pm \gcd C\}^* + ml = C^l + ml = ((AB)^k)^l\end{aligned}$$

and the sets A, B do not form a code. □

We have shown that three finite sets of integers do not form a code, and we have generalized this for certain infinite sets. However, no similar result holds for all infinite sets. The next example shows that there are arbitrarily large codes in the additive monoid of sets of integers.

Example 4.5. Let $A_i = \{1\} \cup \{(i + js)! : j \in \mathbb{N}_0\}$ for $i = 1, \ldots, s$. Let $B = A_1^{k_1} \ldots A_s^{k_s}$. We prove that the sets A_i form a code by showing that the set B uniquely determines the exponents k_i.

Let j be such that $js > \min B = k_1 + \cdots + k_s$. Now $k(i + js)! + \min B - k \in B$ for $k \leq k_i$, but not for $k = k_i + 1$. Thus every k_i is determined by B.

It remains an interesting question what can be said about the unique decipherability problem (or the power equality problem) in the case of non-unary languages. This question probably requires an entirely different approach. In [3] it was proved that the unique decipherability problem is decidable for sets of finite languages if some letter appears exactly once in every word of every language. It is also known that the set of finite prefix sets is a free monoid, *i.e.* generated by a code [8]. Perhaps similar results could be proved for some other classes of languages.

References

[1] J. Berstel and D. Perrin, *Theory of Codes*. Academic Press (1985).
[2] A. Brauer, On a problem of partitions. *Amer. J. Math.* **64** (1942) 299–312.
[3] Ch. Choffrut and J. Karhumäki, Unique decipherability in the monoid of languages: an application of rational relations, in *Proceedings of the Fourth International Computer Science Symposium in Russia* (2009) 71–79.
[4] R. Gilmer, *Commutative Semigroup Rings*. University of Chicago Press (1984).
[5] J.-Y. Kao, J. Shallit and Z. Xu, The frobenius problem in a free monoid, in *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science* (2008) 421–432.
[6] J. Karhumäki and L.P. Lisovik, The equivalence problem of finite substitutions on ab^*c, with applications. *Int. J. Found. Comput. Sci.* **14** (2003) 699–710.
[7] M. Kunc, The power of commuting with finite sets of words. *Theor. Comput. Syst.* **40** (2007) 521–551.
[8] D. Perrin, Codes conjugués. *Inform. Control.* **20** (1972) 222–231.
[9] J.L. Ramírez Alfonsín, *The Diophantine Frobenius Problem*. Oxford University Press (2005).

Communicated by Ch. Choffrut.
Received December 22, 2009. Accepted February 2, 2011.