Discrete Mathematics Letters
www.dmlett.com

Discrete Math. Lett. **4** (2020) 31–36

On subset sums of pseudo–recursive sequences*

Bence Bakos[1], Norbert Hegyvári[1,†], Máté Pálfy[1], Xiao-Hui Yan[2]

[1]*ELTE TTK, Eötvös University, Institute of Mathematics, H-1117 Pázmány st. 1/c, Budapest, Hungary*
[2]*Nanjing Normal University, Nanjing, Jiangsu, China*

(Received: 1 May 2020. Received in revised form: 15 July 2020. Accepted: 3 August 2020. Published online: 6 August 2020.)

Abstract

Let $a_0 = a \in \mathbb{N}$, $\{M_i\}_{i=1}^{\infty}$ be an infinite set of integers and $\{b_1, b_2, \dots, b_k\}$ be a finite set of integers. We say that $\{a_i\}_{i=0}^{\infty}$ is a *pseudo-recursive sequence* if $a_{n+1} = M_{n+1}a_n + b_{j_{n+1}}$ ($b_{j_{n+1}} \in \{b_1, b_2, \dots b_k\}$) holds. In the first part of the paper, we investigate the subset sum of a generalized version of $A_\alpha := \{a_n = \lfloor 2^n\alpha \rfloor : n = 0, 1, 2, \dots\}$, which is a special pseudo-recursive sequence. In the second part, we use A_α for an encryption algorithm.

Keywords: subset sums; Cantor's representation of integers; encoding a codeword.

2020 Mathematics Subject Classification: 11B30, 11B75, 11L03.

1. Introduction

Let $\alpha \in \mathbb{R}$, $1 \le \alpha < 2$, be any real number and consider the sequence $A_\alpha := \{a_n = \lfloor 2^n\alpha \rfloor : n = 0, 1, 2, \dots\}$. This sequence was advised by Rényi and was used by Erdős to investigate some geometric configuration in Hilbert spaces [3]. If we express α in base 2, $\alpha = 1.\xi_1\xi_2 \dots$, ($\xi_i \in \{0, 1\}$, $\sum_i (1 - \xi_i) = \infty$), then one can see $\{a_n\}$ as a *pseudo-recursive* sequence satisfying the identity

$$a_n = 2a_{n-1} + \xi_n;\ n \ge 1. \tag{1}$$

Generally, let $a_0 = a \in \mathbb{N}$, $\{M_i\}_{i=1}^{\infty}$ be an infinite set of integers and $\{b_1, b_2, \dots b_k\}$ be a finite set of integers. We say that $\{a_i\}_{i=1}^{\infty}$ is a *pseudo-recursive sequence* if the identity

$$a_{n+1} = M_{n+1}a_n + b_{j_{n+1}}$$

holds, where $b_{j_{n+1}} \in \{b_1, b_2, \dots b_k\}$) for $n \ge 0$. One of the aims of this paper is to investigate subset sums of a more general pseudo-recursive sequence which was induced by a sequence of Cantor.

The set of subset sums of A_α is defined for $1 \le \alpha < 2$, by

$$P(A_\alpha) := \Big\{ \sum_{i=0}^{\infty} \varepsilon_i a_i : a_i \in A_\alpha;\ \varepsilon_i \in \{0, 1\} \text{ for all } i; \sum_i \varepsilon_i < \infty \Big\}. \tag{2}$$

This set is related to the binary representation of integers (see related results in [4]).

Cantor advised a representation of all non-negative real numbers in the form

$$x = \lfloor x \rfloor + \sum_{i=1}^{\infty} \frac{\eta_i(x)}{q_1 q_2 \cdots q_i},$$

where $\lfloor x \rfloor$ is the integer part of x, $q_i \ge 2$ are integers ($i = 1, 2, \dots$), $0 \le \eta_i(x) < q_i$ are the 'digits' and there are infinitely many i for which $\eta_i(x) < q_i - 1$ holds (see [2]). The related general radix representation of a non negative integer N is also due to Cantor: let $M_1, M_2, \dots$ be an infinite sequence of integers with $M_i \ge 2$, ($i = 1, 2, \dots$) then

$$N = a_1 + a_2M_1 + a_3M_1M_2 + \cdots + a_{n+1}M_1M_2 \cdots M_n,$$

where $0 \le a_i \le M_i - 1$.

The generalized Rényi type sequence would be the following: let $\{q_i\}_{i=1}^{\infty}$ be an infinite (and fixed) sequence of integers with $q_i \ge 2$ ($i = 1, 2, \dots$) and let $Q_n := \prod_{i=1}^{n} q_i$, $Q_0 := 1$. Represent any α, $1 \le \alpha < 2$ in base $\{q_i\}_{i=1}^{\infty}$ and take

$$A_\alpha = \{a_n = \lfloor Q_n\alpha \rfloor : n \in \mathbb{N}\}.$$

*Dedicated to Professor Kálmán Győry on the occasion of his 80th birthday

†Corresponding author (hegyvari.norbert@renyi.hu)

We can define a set for the subset sums of this generalized A_α in a similar way as we did in (2):

$$R(A_\alpha) := \Big\{ \sum_{i=0}^{\infty} \varepsilon_i a_i : a_i \in A_\alpha;\ \varepsilon_i \in \{0,1,\dots,q_{i+1}-1\} \text{ for all } i; \sum_i \varepsilon_i < \infty \Big\}. \tag{3}$$

In some cases we will use the finite version of this, where the summation goes from 0 to n and we denote it by $R(\{a_0,a_1,\dots,a_n\})$. In Section 3 we will show that the elements of $R(A_\alpha)$ also fulfil some *pseudo-recursive identity*, and in our argument we analyze the structure of the set $R(A_\alpha)$.

In Section 4, we discuss an encryption algorithm, based on the set of subset sums of $A_\alpha := \{a_n = \lfloor 2^n\alpha \rfloor : n = 0,1,\dots\}$. The coding process briefly is the following (see Section 4 for more details).

Let $c_n = \xi_1\xi_2\dots\xi_n$ be the n digit codeword, that Alice wants to send to Bob. Alice chooses α to have the following form in base 2: $\alpha = 1.\xi_1\xi_2\dots\xi_n\dots$ (she can extend arbitrarily).

Alice and Bob previously agree on a secret key γ, $0<\gamma<1$. The encrypted (and public) message will be an integer N which is sent by Alice to Bob. She calculates this N in a way to ensure that γN falls in a certain 'gap' of $P(A_\alpha)$.

We will enable everyone to ask about elements of a set $S \subseteq [1,N]$ of integers , which is defined by α (see Section 4). Take the a query function $f : [1,N] \mapsto \{0,1\}$, $f(x) = 0$, if $x \notin S$ and $f(x) = 1$, if $x \in S$. Everyone can query an $(x_0, x_0+1, \dots x_0+L)$ sequence of integers such that $(f(x_0), f(x_0+1), \dots, f(x_0+L))$=$(0,0,\dots,0,1)$. So, we can query x_0 and if it is not in S we can query x_0+1 and so on until we find an element of S or we reach $N+1$.

In Section 4, we will prove that Bob can find out the message with about $\log_2 N$ queries and that an eavesdropper cannot do better, than a $\frac{cN}{\log_2^2 N}$ long query sequence on average.

2. Notation

For the sets $A, B \subset \mathbb{N}$, the sum (difference) is defined by $A \pm B := \{a \pm b : a \in A;\ b \in B\}$ and the restricted sum of these two sets is defined as $A \dot{+} B := \{a+b : a \in A;\ b \in B;\ a \neq b\}$.

For a finite and non empty set $X = \{x_1 < x_2 < \dots < x_r\} (\subset \mathbb{N})$, the length of the biggest gap is

$$\Delta_X = \max\{t \in \mathbb{N} : \exists y_t \in X;\ x_1 \le y_t < x_r;\ [y_t+1,\dots,y_t+t] \cap X = \emptyset\},$$

(if such t does not exist, then $\Delta_X = 0$). So, essentially we have $\Delta_X = \max_{1\le i<r}(x_{i+1}-x_i)-1$. We say that $[y_{\Delta_X}+1, y_{\Delta_X}+\Delta_X]$ (or if $\Delta_X = 0$, the empty set) is the biggest gap. If we fix X then we write briefly Δ instead of Δ_X.

Throughout the paper, $\log_2 N$ will denote the logarithm in base 2.

3. The structure of $R(A_\alpha)$

In this section, we are going to investigate the structure of the set $R(A_\alpha)$. Here we shall use the notation given just before equation (3), so $1 \le \alpha < 2$ is written in base $\{q_i\}_{i=1}^\infty$, $Q_0 = 1$, $Q_i = \prod_{i=1}^n q_i$ and $A_\alpha = \{a_n = \lfloor Q_n\alpha \rfloor : n \in \mathbb{N}\}$. Let η_n denote the n'th digit of α in the Cantor type representation. Firstly, we show that the elements of A_α ensure a pseudo-recursion.

Theorem 3.1. *For every $n \ge 0$, the pseudo-recursion*

$$a_{n+1} = q_{n+1}a_n + \eta_{n+1}$$

holds.

Proof. Write

$$\begin{aligned} Q_n\alpha &= Q_n + \eta_1\frac{Q_n}{q_1} + \eta_2\frac{Q_n}{q_1q_2} + \dots + \eta_n\frac{Q_n}{q_1q_2\cdots q_n} + \eta_{n+1}\frac{Q_n}{q_1q_2\cdots q_nq_{n+1}} + \dots \\ &= Q_n + \eta_1\frac{Q_n}{q_1} + \eta_2\frac{Q_n}{q_1q_2} + \dots + \eta_n\frac{Q_n}{q_1q_2\cdots q_n} + H_n. \end{aligned}$$

Since $Q_j = \prod_{i=1}^j q_i$, thus the fractions $\frac{Q_n}{q_1}; \frac{Q_n}{q_1q_2}; \dots; \frac{Q_n}{q_1q_2\cdots q_n} = 1$ are integers. Now, we show that $H_n < 1$ and hence

$$a_n = \lfloor Q_n\alpha \rfloor = Q_n + \eta_1\frac{Q_n}{q_1} + \eta_2\frac{Q_n}{q_1q_2} + \dots + \eta_{n-1}\frac{Q_n}{q_1q_2\dots q_{n-1}} + \eta_n. \tag{4}$$

Indeed simplifying in Q_n, we obtain

$$H_n = \eta_{n+1}\frac{1}{q_{n+1}} + \eta_{n+2}\frac{1}{q_{n+1}q_{n+2}} + \dots,$$

$\eta_r \le q_r - 1$ for $r \ge n+1$, and there exists an s for which the inequality is strict, so we obtain

$$\begin{aligned} H_n &\le (q_{n+1}-1)\frac{1}{q_{n+1}} + (q_{n+2}-1)\frac{1}{q_{n+1}q_{n+2}} + \cdots + (q_s-2)\frac{1}{q_{n+1}\cdots q_s} + \dots \\ &\le 1 - \frac{1}{q_{n+1}\cdots q_s} < 1. \end{aligned} \tag{5}$$

Now $Q_{n+1} = Q_n \cdot q_{n+1}$, so multiplying (4) by q_{n+1}, we get

$$q_{n+1}a_n = Q_n \cdot q_{n+1} + \eta_1 \frac{Q_n \cdot q_{n+1}}{q_1} + \eta_2 \frac{Q_n \cdot q_{n+1}}{q_1 q_2} + \cdots + \eta_n \cdot q_{n+1}.$$

Now if we rewrite (4) with $n+1$ instead of n and we subtract the previous expression from it we get that $a_{n+1} - q_{n+1}a_n = \eta_{n+1}$, as we wanted.

□

Proposition 3.1. *Let* $N(n) := \sum_{i=0}^{n}(q_{i+1}-1)a_i$. *The set* $R(A_\alpha) \cap [0, N(n)]$ *is symmetric with respect to the middle point, i.e.*

$$R(A_\alpha) \cap [0, N(n)] = N(n) - (R(A_\alpha) \cap [0, N(n)]).$$

Proof. Pick an element x from $R(A_\alpha) \cap [0, N(n)]$. The element x can be written as $x = \sum_{i=0}^{n} \varepsilon_i a_i$, $a_i \in A_\alpha$, $\varepsilon_i \in \{0, 1, \dots, q_{i+1} - 1\}$. Now,

$$y = N(n) - x = \sum_{i=0}^{n}(q_{i+1}-1)a_i - \sum_{i=0}^{n}\varepsilon_i a_i = \sum_{i=0}^{n}(q_{i+1}-1-\varepsilon_i)a_i = \sum_{i=0}^{n}\varepsilon_i' a_i$$

where $\varepsilon_i' \in \{0, 1, \dots, q_{i+1}-1\}$ which implies that $y \in R(A_\alpha) \cap [0, N(n)]$. When $x \in N(n) - (R(A_\alpha) \cap [0, N(n)])$ the argument is the same.

□

Proposition 3.2.

$$R(A_\alpha) \cap [a_n, a_{n+1}) = \bigcup_{k=1}^{q_{n+1}-1} \{ka_n + (R(A_\alpha) \cap [0, a_n))\}.$$

Moreover,

$$|R(A_\alpha) \cap [0, a_n)| = q_n q_{n-1} \cdots q_1$$

i.e. each member of the set $R(A_\alpha)$ *has a unique representation.*

Proof. Since $a_{n+1} = q_{n+1}a_n + \eta_{n+1}$, it follows $a_{n+1} \ge q_{n+1}a_n = (q_{n+1}-1)a_n + a_n$ and hence, by induction,

$$a_{n+1} \ge \sum_{i=0}^{n}(q_{i+1}-1)a_i + a_0. \tag{6}$$

Now by (6),

$$R(A_\alpha) \cap [a_n, a_{n+1}) = \bigcup_{k=1}^{q_{n+1}-1} \{ka_n, ka_n + a_0, \cdots\cdots, ka_n + (q_1-1)a_0, ka_n + a_1, \cdots, ka_n + (q_1-1)a_0 + (q_2-1)a_1, \cdots$$

$$\cdots, ka_n + (q_1-1)a_0 + \cdots + (q_n-1)a_{n-1}\} = \bigcup_{k=1}^{q_{n+1}-1} \{ka_n + (R(A_\alpha) \cap [0, a_n))\}.$$

So,

$$|R(A_\alpha) \cap [a_n, a_{n+1})| = (q_{n+1}-1)|R(A_\alpha) \cap [0, a_n)|.$$

It follows that

$$\begin{aligned} |R(A_\alpha) \cap [0, a_n)| &= 1 + \sum_{i=1}^{n} |R(A_\alpha) \cap [a_{i-1}, a_i)| = 1 + (q_1-1) + (q_2-1)(1 + (q_1-1)) + \cdots \\ &= q_n q_{n-1} \cdots q_1, \end{aligned}$$

which can be easily seen by induction. □

Corollary 3.1. *Let* $R(A_\alpha) = \{0 = r_0 < r_1 < r_2 < \cdots\}$. *For any index* k, *let*

$$(i)\ k = k_n q_{n-1} \cdots q_1 + k_{n-1} q_{n-2} \cdots q_1 + \cdots + k_2 q_1 + k_1,\ \ 0 \le k_i \le q_i - 1, k_n \neq 0.$$

Then

$$(ii)\ r_k = k_n a_{n-1} + k_{n-1} a_{n-2} + \cdots + k_2 a_1 + k_1 a_0.$$

Conversely, if $a = k_n a_{n-1} + k_{n-1} a_{n-2} + \cdots + k_2 a_1 + k_1 a_0 \in R(A_\alpha)$, *then*

$$a = r_{k_n q_{n-1} \cdots q_1 + k_{n-1} q_{n-2} \cdots q_1 + \cdots + k_2 q_1 + k_1}.$$

Proof. We can easily see that the function f from $\mathbb{N}$ to $R(A_\alpha)$ which carries $k \in \mathbb{N}$ (given in the form (i)) to the element given in the form as in (ii) is strictly increasing (by (6) from the previous proof) and surjective. The converse direction follows from that it is exactly the inverse function of f. □

Proposition 3.3. *Let* Δ_k *be the length of the biggest gap of* $B_k := R(A_\alpha) \cap [1, a_k]$. *Then the sequence* $\{\Delta_k\}_{k=1}^n$, $n \ge 2$ *forms an increasing sequence. Moreover*

$$\Delta_k = \sum_{j=1}^{k} \eta_j \tag{7}$$

and the corresponding sequence to the biggest gap is $\sum_{i=0}^{k-1}(q_{i+1} - 1)a_i + 1, \ldots, a_k - 1$.

Proof. We use induction on k. Firstly, we remark that for any n:

$$R(\{a_0, a_1, \ldots, a_n\}) = R(\{a_0, a_1, \ldots, a_{n-1}\}) + \{j a_n : j = 0, \ldots, q_{n+1} - 1\}. \tag{8}$$

Now, look at the case $k = 1$:

$$B_1 = \{a_0 = 1, \ldots, (q_1 - 1)a_0, a_1\}$$

so the lengths of the gaps are 0s and $(a_1 - (q_1 - 1)a_0) - 1$. Now $a_1 = q_1 a_0 + \eta_1$ (by Theorem 3.1), thus $(a_1 - (q_1 - 1)a_0) - 1 = \eta_1 = \Delta_1$ and the corresponding sequence is $(q_1 - 1)a_0 + 1, \ldots, a_1 - 1$ or $\emptyset$. So for $k = 1$ the statement is true.

Assume now that (7) is true for $k \ge 1$. Now, by (8) and the inductive hypothesis in the intervals $[(j-1)a_k, ja_k]$, $1 \le j \le q_{k+1} - 1$ the biggest gap is Δ_k. Using Theorem 3.1 and the inductive hypothesis again in the last interval $[(q_{k+1} - 1)a_k, a_{k+1}]$ the biggest gap is $\Delta_k + \eta_{k+1}$ and the corresponding sequence is $\sum_{i=0}^{k}(q_{i+1} - 1)a_i + 1, \ldots, a_{k+1} - 1$, which proves the proposition. □

4. Encryption using the set $A_\alpha \dot{+} A_\alpha$

Now, we are ready to analyze our encryption scheme, which was introduced in the end of Section 1. Before the theorems, we shall repeat the process here in a bit more detail.

Let c_n be the binary codeword (the message) with n digits: $c_n = \xi_1 \xi_2 \ldots \xi_n$. Alice chooses α for the message c_n such that $\alpha = 1.\xi_1 \xi_2 \ldots \xi_n \ldots$ in base 2 (she can extend arbitrarily after ξ_n, the only assumption is that the digit 0 appears infinitely many times). We will use the set $A_\alpha := \{a_n = \lfloor 2^n \alpha \rfloor : n = 0, 1, 2, \ldots\}$ and the set of its subset sums $P(A_\alpha)$ defined in (2). Since it is a special case of the generalization we investigated in the previous section (namely $q_i = 2$ for all i) we can use those results here.

Alice and Bob previously agree on the secret key γ, $0 < \gamma < 1$. Alice chooses a random integer $N \in [\frac{\sum_{i=0}^{n-1} a_i + 1}{\gamma}, \frac{a_n}{\gamma})$. The encrypted message (the ciphertext) will be this integer N, which is sent by Alice to Bob.

Let S be the set given by

$$S := (A_\alpha \dot{+} A_\alpha) \cap [1, N]. \tag{9}$$

The set S is available to everyone via a query sequence: Let us define the function $f : [1, N] \to \{0, 1\}$, $f(x) = 0$, if $x \notin S$ and $f(x) = 1$, if $x \in S$. Everyone can query an $(x_0, x_0 + 1, \ldots x_0 + L)$ sequence of integers such that $(f(x_0), f(x_0 + 1), \ldots, f(x_0 + L))$=$(0, 0, \ldots, 0, 1)$. So we can query x_0 and if it is not in S we can query $x_0 + 1$ and so on until we find an element of S. The length of the query sequence is $L_\alpha(N, x_0) := L$.

Firstly, we prove the following:

Theorem 4.1. *If Alice sends the message* N, *for which*

$$\gamma N \in [\sum_{i=0}^{n-1} a_i + 1, a_n) \tag{10}$$

holds, then Bob can get the message by asking a query sequence with length $L_\alpha(N) \le \log_2 N + 2$.

Proof. Write $\alpha = 1 + \sum_i \xi_i 2^{-i}$ which is now hidden and we are interested in $a_n = \lfloor 2^n \alpha \rfloor$. Note that $A_\alpha \dot{+} A_\alpha \subseteq P(A_\alpha)$, i.e. we can use the structure of $P(A_\alpha)$. Let $R = \lfloor \gamma N \rfloor$. Due to the choice of N, R is in the biggest gap of $P(\{a_0, \dots, a_n\})$. Thus the smallest number which is at least R and belongs to this subset sum is just a_n, and hence $a_n + a_0 = a_n + 1$ is the first element of $A_\alpha \dot{+} A_\alpha$, which is at least R.

So the query sequence of Bob should be $(x_0 = R, R+1, \dots, R+L)$. By Proposition 3.3 we get that

$$L \le a_n + 1 - \sum_{k=0}^{n-1} a_k = \Delta_n + 2 = \sum_{k=1}^{n} \xi_k + 2.$$

Since for every k, $\xi_k \le 1$, thus $L \le n + 2 \le \log_2 N + 2$. So the length of Bob's query sequence is at most $\log_2 N$ and the element he finds at the end is $a_n + 1$. From this Bob can easily get c_n, since a_n has the form of $a_n = 1\xi_1\xi_2 \dots \xi_n$. □

Let Eve be an eavesdropper (a passive attacker; i.e. she can catch the encoded ciphertext and also can ask a query sequence). We are interested in how long Eve needs to query on average to find an element of S. The appropriate mathematical phrasing would be the following:

After we fixed the secret key γ and N (the encrypted message), enumerate the elements of S:
$S := \{s_1 < s_2 < \dots < s_k\}$ and let X be the random variable which says that how long Eve needs to query if she picks one element of $[1, N]$ uniformly at random. More precisely, if $n \in [1, N]$, then $X(n) :=$ the length of the query sequence started at n. We have the following theorem:

Theorem 4.2. *The expected length of query sequence of an eavesdropper Eve, who chooses the start of the query sequence uniformly at random in* $[1, N]$ *is*

$$\mathbb{E}(X) \ge \frac{cN}{\log_2^2 N}, \tag{11}$$

($c > 0$ absolute).

Proof. First, we calculate the expected value of a query sequence, assuming that the first number Eve asked is in a fixed gap. We introduce the following events:

$$B_0 := \{\text{the number } n, \text{ that Eve chose is in } [1, s_1)\}$$

$$B_k := \{\text{the number } n, \text{ that Eve chose is in } [s_k, N]\}$$

and for $i = 1, \dots, k-1$:

$$B_i := \{ \text{ the number } n, \text{ that Eve chose is in } [s_i, s_{i+1})\}.$$

Note that the previously defined events form a complete system of events. For $1 \le i < k$, by the decoding scheme, we get that:

$$\mathbb{E}(X \mid B_i) = 0 \cdot \frac{1}{s_{i+1} - s_i} + (s_{i+1} - (s_i + 1)) \cdot \frac{1}{s_{i+1} - s_i} + (s_{i+1} - (s_i + 2)) \cdot \frac{1}{s_{i+1} - s_i} + \dots + 1 \cdot \frac{1}{s_{i+1} - s_i} = \frac{s_{i+1} - s_i - 1}{2}$$

and with the same argument it is easy to see that $\mathbb{E}(X \mid B_0) = \frac{s_1}{2}$ and $\mathbb{E}(X \mid B_k) = \frac{N - s_k}{2}$. With the help of the law of total expectation we get:

$$\mathbb{E}(X) = \sum_{i=0}^{k} \mathbb{P}(B_i)\mathbb{E}(X \mid B_i) = \frac{s_1}{N} \cdot \frac{s_1}{2} + \sum_{i=1}^{k-1} \frac{s_{i+1} - s_i}{N} \cdot \frac{s_{i+1} - s_i - 1}{2} + \frac{N - s_k + 1}{N} \cdot \frac{N - s_k}{2} \ge$$

$$\ge \frac{s_1^2}{k+1} \cdot \frac{k}{2N} + \sum_{i=1}^{k-1} \frac{(s_{i+1} - s_i - 1)^2}{k+1} \cdot \frac{k}{2N} + \frac{(N - s_k)^2}{k+1} \cdot \frac{k}{2N} \ge$$

$$\ge \frac{k}{2N} \left(\frac{s_1 + \sum_{i=1}^{k-1} (s_{i+1} - s_i - 1) + (N - s_k)}{k+1} \right)^2 = \frac{k}{2N} \Big(\frac{N - k + 1}{k+1} \Big)^2 \ge \frac{(N-k+1)^2}{8kN}$$

where in the second inequality we used the Cauchy inequality. It is easy to see that $|A_\alpha \cap [1, N]| \le c' \log N$. Further we have that $S = (A_\alpha \dot{+} A_\alpha) \cap [1, N]$ and the representations of the elements in $A_\alpha \dot{+} A_\alpha$ is unique, since we have seen that in $P(A_\alpha)$ the representation is unique, and $A_\alpha \dot{+} A_\alpha \subseteq P(A_\alpha)$. So we get that $k \le c'' \log_2^2 N$, which gives the desired result. □

5. Concluding remarks

1. Against an eavesdropper we shall restrict the length of the query sequence by N^β (for some parameter β). This way it is possible to ensure that Eve cannot be sure about the right codeword even if she finds an x_L relatively quickly.

If Eve (the eavesdropper) finds an x_L after a query sequence, then she has to find α with the information she got so far. How can she do that? She knows that she found an element of $A_\alpha \dot{+} A_\alpha$ so she tries to decompose x_L as a sum of b_1 and b_2 where one of them is the prefix of the other one in base 2 (because they should both have the form of $\lfloor 2^i\alpha \rfloor$). Seemingly there can be many decompositions, but Eve can eliminate some of these by doing the following: take b_1' and b_2' such that $b_1' = \left\lfloor \frac{b_1}{2^j} \right\rfloor$, $b_2' = \left\lfloor \frac{b_2}{2^t} \right\rfloor$ and $b_1' > b_2'$. If b_i really has the form $\lfloor 2^i\alpha \rfloor$ then $b_1' + b_2'$ is in S. So if Eve has queried $b_1' + b_2'$ (and got 0 as an answer) then b_1 and b_2 is not the correct decomposition. If this is not the case, namely for all b_1' and b_2' the sum $b_1' + b_2'$ was not queried, then we call this b_1, b_2 pair *acceptable* for Eve. With the restriction on the length of the query sequence we have a result about the number of acceptable decompositions for Eve. Namely even if she manages to find an x_L quicker than N^β, she has at least $c \log_2 N$ acceptable b_1, b_2 pairs (where c is a constant).

2. We learnt in section 4 that Bob has to know and keep as a secret for decoding the parameter γ. Since the bound N for the query sequence always varies, hence for eavesdropper has no chance to detect the value of γ, i.e. Bob can use this parameter without restriction.

3. Results in cryptography has a long list. Interestingly (by the knowledge of the authors) papers which relate to subset sums are few (see e.g. [1, 5]). Although the general knapsack problem is known to be NP–complete.

Acknowledgments

The second named author is supported by grant K–129335. The first and third named authors are supported by the European Union, co–financed by the European Social Fund (EFOP–3.6.3–VEKOP–16–2017–00002).

References

[1] E. F. Brickell, A. M. Odlyzko, Cryptanalysis: a survey of recent results, *Proc. IEEE* **76** (1988) 578–593.
[2] G. Cantor, Über die einfachen Zahlensysteme, *Z. Angew. Math. Phys.* **14** (1869) 121–128.
[3] P. Erdős, Geometrical and set–theoretical properties of subsets of Hilbert–space (in Hungarian), *Mat. Lapok* **19** (1968) 255–258; MR40 708.
[4] N. Hegyvári, Some remarks on a problem of Erdős and Graham, *Acta Math. Hungar.* **53** (1989) 149–154.
[5] R. Impagliazzo, M. Naor, Efficient cryptographic schemes provably as secure as subset sum, *Proc. 30th IEEE Symposium on Foundations of Computer Science*, IEEE, 1989, pp. 236–241.