

# SELF-SIMILAR SUBSETS OF THE CANTOR SET

DE-JUN FENG, HUI RAO, AND YANG WANG

ABSTRACT. In this paper, we study the following question raised by Mattila in 1998: what are the self-similar subsets of the middle-third Cantor set  $\mathcal{C}$ ? We give criteria for a complete classification of all such subsets. We show that for any self-similar subset  $\mathbf{F}$  of  $\mathcal{C}$  containing more than one point every linear generating IFS of  $\mathbf{F}$  must consist of similitudes with contraction ratios  $\pm 3^{-n}$ ,  $n \in \mathbb{N}$ . In particular, a simple criterion is formulated to characterize self-similar subsets of  $\mathcal{C}$  with equal contraction ratio in modulus.

## 1. INTRODUCTION

Let  $\mathcal{C}$  denote the standard middle-third Cantor set. The main goal of this paper is to answer the following open question raised by Mattila [2] in 1998: what are the self-similar subsets of  $\mathcal{C}$ ?

Recall that a non-empty compact set  $\mathbf{F} \subset \mathbb{R}$  is said to be *self-similar* if there exists a finite family  $\Phi = \{\phi_i\}_{i=1}^k$  of contracting similarity maps on  $\mathbb{R}$  such that

$$(1.1) \quad \mathbf{F} = \bigcup_{i=1}^k \phi_i(\mathbf{F}).$$

Such  $\Phi$  is called a linear *iterated function system* (IFS) on  $\mathbb{R}$ . As proved by Hutchinson [5], for a given IFS  $\Phi$ , there is a unique non-empty compact set  $\mathbf{F}$  satisfying (1.1). To specify the relation between  $\Phi$  and  $\mathbf{F}$ , we call  $\Phi$  a *linear generating IFS* of  $\mathbf{F}$ , and  $\mathbf{F}$  the *attractor* of  $\Phi$ . Throughout this paper, we use  $\mathbf{F}_\Phi$  to denote the attractor of a given linear IFS  $\Phi$ . A self-similar set  $\mathbf{F} = \mathbf{F}_\Phi$  is said to be *non-trivial* if it is not a singleton.

The middle-third Cantor set  $\mathcal{C}$  is one of the most well known examples of self-similar sets. It has a generating IFS  $\{x/3, (x+2)/3\}$ .

---

*Key words and phrases.* Cantor set, Self-similar sets, Ternary expansions, Set of uniqueness.  
*2000 Mathematics Subject Classification:* Primary 28A78, Secondary 28A80, 11K16.

The first result of this paper is the following theorem, which is our starting point for further investigations.

**Theorem 1.1.** *Assume that  $\mathbf{F} \subseteq \mathbf{C}$  is a non-trivial self-similar set, generated by a linear IFS  $\Phi = \{\phi_i\}_{i=1}^k$  on  $\mathbb{R}$ . Then for each  $1 \leq i \leq k$ ,  $\phi_i$  has contraction ratio  $\pm 3^{-m_i}$ , where  $m_i \in \mathbb{N}$ .*

The proof of the above theorem is based on a short geometric argument and a fundamental result of Salem and Zygmund on the sets of uniqueness in harmonic analysis.

It is easy to see that if a self-similar set  $\mathbf{F}$  has a generating IFS  $\Phi = \{\phi_i\}_{i=1}^k$  that is *derived* from the IFS  $\Psi := \{x/3, (x+2)/3\}$ , i.e. each map in  $\Phi$  is a finite composition of maps in  $\Psi$ , then  $\mathbf{F} \subseteq \mathbf{C}$ . In light of Theorem 1.1, one may guess that each nontrivial self-similar subset of  $\mathbf{C}$  has a linear generating IFS derived from  $\Psi$ . However, this is not true. The following counter example was constructed in [4].

**Example 1.1.** *Let  $\Phi = \{\frac{1}{9}x, \frac{1}{9}(x+2)\}$ . Choose a sequence  $(\epsilon_n)_{n=1}^\infty$  with  $\epsilon_n \in \{0, 2\}$  so that  $w = \sum_{n=1}^\infty \epsilon_n 3^{-2n+1}$  is an irrational number. Then by looking at the ternary expansion of the elements in  $\mathbf{F}_\Phi + w := \{x + w : x \in \mathbf{F}_\Phi\}$ , it is easy to see that  $\mathbf{F}_\Phi + w \subset \mathbf{C}$ . Observe that  $\mathbf{F}_\Phi + w$  is a self-similar subset of  $\mathbf{C}$  since it is the attractor of the IFS  $\Phi' = \{\frac{1}{9}(x+8w), \frac{1}{9}(x+2+8w)\}$ . However any generating IFS of  $\mathbf{F}_{\Phi'}$  can not be derived from the original IFS  $\{\psi_0 = x/3, \psi_1 = (x+2)/3\}$ , since  $w = \min \mathbf{F}_{\Phi'}$  can not be the fixed point of any map  $\psi_{i_1 i_2 \dots i_n}$  composed from  $\psi_0, \psi_1$  due to the irrationality of  $w$ .*

The above construction actually shows that  $\mathbf{C}$  has uncountably many non-trivial self-similar subsets, and indicates the non-triviality of Mattila's question.

To further describe self-similar subsets of  $\mathbf{C}$ , we consider a special class of linear IFSs. Let  $\mathcal{F}$  denote the collection of linear IFSs  $\Phi = \{\phi_i\}_{i=1}^k$  on  $\mathbb{R}$  that satisfy the following conditions (i)-(iii):

- (i)  $k \geq 2$  and for each  $1 \leq i \leq k$ ,  $\phi_i$  is of the form
- $$(1.2) \quad \phi_i(x) = s_i 3^{-m_i} x + d_i, \quad s_i = \pm 1, \quad m_i \in \mathbb{N}, \quad 0 \leq d_i \leq 1.$$
- (ii)  $s_1 = 1$  and  $d_1 = 0$ .
  - (iii)  $d_i > 0$  for at least one  $1 \leq i \leq k$ . Moreover  $d_i > 0$  if  $s_i = -1$ .

It is easy to check that for each  $\Phi \in \mathcal{F}$ ,  $\mathbf{F}_\Phi \subseteq [0, 2]$  and  $\min \mathbf{F}_\Phi = 0$ . By the symmetry of  $\mathcal{C}$  (i.e.  $\mathcal{C} = 1 - \mathcal{C}$ ), if  $\mathbf{F}$  is a non-trivial self-similar subset of  $\mathcal{C}$ , then  $1 - \mathbf{F}$  is also a self-similar subset of  $\mathcal{C}$ . The following result is a simple consequence of Theorem 1.1 (see Appendix A for a proof).

**Lemma 1.1.** *Let  $\mathbf{F} \subseteq \mathcal{C}$  be a non-trivial self-similar set. Then either  $\mathbf{F}$  or  $1 - \mathbf{F}$  can be written as  $a + \mathbf{E}$ , where  $a \in \mathcal{C}$  and  $\mathbf{E} = \mathbf{F}_\Phi$  for some  $\Phi \in \mathcal{F}$ .*

Hence, to characterize the self-similar subsets of  $\mathcal{C}$ , it is equivalent to characterize the pairs  $(a, \Phi) \in \mathcal{C} \times \mathcal{F}$  so that  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ . This is what we will carry out in this paper.

We start by introducing some notation and definitions. Notice that any real number  $x \in [0, 1]$  admits at least one ternary expansion

$$x = \sum_{n=1}^{\infty} u_n 3^{-n}$$

with  $u_n \in \{0, 2, -2\}$ , and certain rational numbers may admit two different such expansions. For our purpose, we construct a subset  $\Theta \subset \{0, 2, \bar{2}\}^{\mathbb{N}}$ , with convention  $\bar{2} := -2$ , by

$$(1.3) \quad \Theta = \left\{ \mathbf{u} \in \{0, 2, \bar{2}\}^{\mathbb{N}} : \iota(\mathbf{u}) \neq -2, \mathbf{u} \text{ does not end with } 2\bar{2}^{\infty} \text{ and } \bar{2}2^{\infty} \right\},$$

where

$$(1.4) \quad \iota(\mathbf{u}) = \begin{cases} 0, & \text{if } \mathbf{u} = 0^{\infty}, \\ \text{the first non-zero letter appearing in } \mathbf{u}, & \text{if } \mathbf{u} \neq 0^{\infty}. \end{cases}$$

Here we say that  $\mathbf{u}$  ends with  $\mathbf{c} \in \{0, 2, \bar{2}\}^{\mathbb{N}}$  if  $\mathbf{u} = \omega \mathbf{c}$  for some finite or empty word  $\omega$  over the alphabet  $\{0, 2, \bar{2}\}$ . Then it is easily checked that every  $x \in [0, 1]$  admits one and only one ternary expansion  $\sum_{n=1}^{\infty} u_n 3^{-n}$  with  $\mathbf{u} = (u_n)_{n=1}^{\infty} \in \Theta$ .

**Definition 1.1.** *For  $x \in [0, 1]$ , the unique ternary expansion  $x = \sum_{n=1}^{\infty} u_n 3^{-n}$  with  $(u_n)_{n=1}^{\infty} \in \Theta$ , is called the intrinsic ternary expansion of  $x$ .*

Clearly any  $x \in \mathcal{C}$  admits an intrinsic ternary expansion  $\sum_{n=1}^{\infty} b_n 3^{-n}$  with  $b_n \in \{0, 2\}$ . Now we define the notion of intrinsic translation matrix.

**Definition 1.2.** *Let  $\Phi = \{\phi_i(x) = s_i 3^{-m_i} x + d_i\}_{i=1}^k \in \mathcal{F}$ . Let  $\sum_{n=1}^{\infty} u_{i,n} 3^{-n}$  be the intrinsic ternary expansion of  $d_i$ ,  $i = 1, \dots, k$ . Denote by  $U_\Phi$  the following  $k \times \infty$*

matrix with entries in  $\{0, 2, \bar{2}\}$ :

$$(1.5) \quad U_\Phi = (u_{i,n})_{1 \leq i \leq k, n \geq 1}.$$

We call  $U_\Phi$  the intrinsic translation matrix of  $\Phi$ .

For  $n \in \mathbb{N}$ , let  $U_\Phi(n)$  denote the  $n$ -th column of  $U_\Phi$ , i.e.

$$U_\Phi(n) = (u_{1,n}, \dots, u_{k,n})^T,$$

where the superscript  $T$  denotes the transpose. We say that the vector  $U_\Phi(n)$  is *positive* if  $u_{i,n} \in \{0, 2\}$  for all  $1 \leq i \leq k$  and  $u_{j,n} = 2$  for at least one  $j \in \{1, \dots, k\}$ ; correspondingly, we say that  $U_\Phi(n)$  is *negative* if  $u_{i,n} \in \{0, \bar{2}\}$  for all  $1 \leq i \leq k$  and  $u_{j,n} = \bar{2}$  for at least one  $j \in \{1, \dots, k\}$ .

For  $m \in \mathbb{N}$ , let  $\mathcal{F}_m^+$  denote the collection of  $\Phi \in \mathcal{F}$  with uniform contraction ratio  $3^{-m}$ , and  $\mathcal{F}_m$  the collection of  $\Phi \in \mathcal{F}$  with contraction ratios  $\pm 3^{-m}$ .

Now we are ready to present one of our main results.

**Theorem 1.2.** *Let  $m \in \mathbb{N}$  and  $(a, \Phi) \in \mathbf{C} \times \mathcal{F}_m^+$ . Define  $U_\Phi$  as in (1.5). Write  $a = \sum_{n=1}^{\infty} a_n 3^{-n}$  with  $a_n \in \{0, 2\}$ . Then  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  if and only if there exists a finite non-empty set  $\mathcal{M} \subset \mathbb{N}$  such that the following properties (i)-(iv) hold:*

- (i) *Any two numbers in  $\mathcal{M}$  are incongruent modulo  $m$ ;*
- (ii) *For each  $n \in \mathcal{M}$ ,  $U_\Phi(n)$  is either positive or negative;*
- (iii) *For each  $n \in \mathbb{N} \setminus \mathcal{M}$ ,  $U_\Phi(n) = (0, 0, \dots, 0)^T$ ;*
- (iv) *For any  $n \in \mathcal{M}$ , and any integer  $t \geq 0$ ,*

$$a_{n+mt} = \begin{cases} 0, & \text{if } U_\Phi(n) \text{ is positive,} \\ 2, & \text{if } U_\Phi(n) \text{ is negative.} \end{cases}$$

According to the above theorems 1.2, every pair  $(a, \Phi) \in \mathbf{C} \times \mathcal{F}_m^+$  satisfying  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  can be obtained as follows. Choose a finite set  $\mathcal{M} \subset \mathbb{N}$  such that any two numbers in  $\mathcal{M}$  are incongruent modulo  $m$ . Let  $k \geq 2$ . Construct an  $k \times \infty$  matrix  $U$  with entries in  $\{0, 2, \bar{2}\}$  such that the first row of  $U$  consists of the entries 0 only, all the other rows are the sequences in  $\Theta$  (see (1.3) for the definition of  $\Theta$ ) and moreover,  $U$  (instead of  $U_\Phi$ ) fulfils the conditions (ii)-(iii) in Theorem 1.2. There are at most  $2^{km}$  different such matrices  $U$  once  $\mathcal{M}$  has been fixed. For  $i = 1, \dots, k$ , let  $d_i$  be the number in  $[0, 1]$  with intrinsic ternary expansion given by the  $i$ -th row of  $U$ . Choose  $a = \sum_{n=1}^{\infty} a_n 3^{-n} \in \mathbf{C}$  such that  $(a_n)$  satisfies (iv) of Theorem 1.2. If

the cardinality of  $\mathcal{M}$  equals  $m$ , there are finitely many choices for such  $a$ ; otherwise if the cardinality is less than  $m$ , there are uncountable many choices for  $a$ . Let  $\Phi = \{\phi_i(x) = 3^{-m}x + d_i\}_{i=1}^k$ . Then  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Below we give a simple example.

**Example 1.2.** Let  $m = 4$  and  $k = 4$ . Set  $\mathcal{M} = \{2, 3, 5\}$ . Construct a  $4 \times \infty$  matrix  $U$  as below:

$$U = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & 0 & 0 & \cdots \\ 0 & 2 & 2 & 0 & \bar{2} & 0 & \cdots \\ 0 & 2 & 0 & 0 & \bar{2} & 0 & \cdots \end{pmatrix},$$

where the  $n$ -th column  $U(n)$  of  $U$  is null for  $n \geq 6$ . One can check that each row of  $U$  belongs to  $\Theta$ ,  $U(n)$  is either positive or negative for  $n \in \mathcal{M}$ , and  $U(n)$  is null for  $n \in \mathbb{N} \setminus \mathcal{M}$ . Thus the conditions (i)-(iii) in Theorem 1.2 are fulfilled for  $U$ . Notice that  $U = U_\Phi$  for

$$\Phi = \{3^{-4}x, 3^{-4}x + .002, 3^{-4}x + .0220\bar{2}, 3^{-4}x + .0200\bar{2}\},$$

where the translation parts in the above maps are written in its intrinsic ternary expansions. Let  $a = .(2002)^\infty$ . Then the condition (iv) in Theorem 1.2 is also fulfilled. Hence by Theorem 1.2,  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Moreover, by Theorem 1.2, for  $b = \sum_{n=1}^\infty b_n 3^{-n} \in \mathbf{C}$  with  $(b_n)_{n=1}^\infty \in \{0, 2\}^\mathbb{N}$ ,

$$b + \mathbf{F}_\Phi \subseteq \mathbf{C} \iff b_{4n+2} = 0, b_{4n+3} = 0, b_{4n+5} = 2 \text{ for each integer } n \geq 0.$$

Next we consider more general pairs  $(a, \Phi) \in \mathbf{C} \times (\mathcal{F}_m \setminus \mathcal{F}_m^+)$ . Let

$$\Phi = \{\phi_i(x) = s_i 3^{-m}x + d_i\}_{i=1}^k \in \mathcal{F}_m$$

so that there exists  $1 \leq k' < k$  such that  $s_i = 1$  for  $1 \leq i \leq k'$  and  $s_i = -1$  for  $i > k'$ . Write

$$(1.6) \quad \begin{aligned} \mathbf{e}_0 &:= \underbrace{(0, \dots, 0)}_k^T, \\ \mathbf{e}_1 &:= \underbrace{(0, \dots, 0)}_{k'} \underbrace{(2, \dots, 2)}_{k-k'}^T, \\ \mathbf{e}_2 &:= \underbrace{(0, \dots, 0)}_{k'} \underbrace{(\bar{2}, \dots, \bar{2})}_{k-k'}^T. \end{aligned}$$

**Theorem 1.3.** Let  $m \in \mathbb{N}$  and  $(a, \Phi) \in \mathbf{C} \times (\mathcal{F}_m \setminus \mathcal{F}_m^+)$ . Write  $a = \sum_{n=1}^\infty a_n 3^{-n}$  with  $a_n \in \{0, 2\}$ . Define  $U_\Phi$  as in (1.5). Then  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  if and only if there exists a finite non-empty set  $\mathcal{M} \subset \mathbb{N}$  so that the following properties (i)-(iv) hold:

- (i) Any two numbers in  $\mathcal{M}$  are incongruent modulo  $m$ ;
- (ii) For each  $n \in \mathcal{M}$ , the  $n$ -th column  $U_\Phi(n)$  of  $U_\Phi$  is either positive or negative, and for any  $t \geq 0$ ,

$$U_\Phi(n + m(t + 1)) = \begin{cases} \mathbf{e}_1, & \text{if } U_\Phi(n) \text{ is positive,} \\ \mathbf{e}_2, & \text{if } U_\Phi(n) \text{ is negative;} \end{cases}$$

- (iii) For each  $n \in \mathbb{N} \setminus (\mathcal{M} + (m\mathbb{N} \cup \{0\}))$ ,  $U_\Phi(n) = \mathbf{e}_0$ .

- (iv) For any  $n \in \mathcal{M}$ , and any integer  $t \geq 0$ ,

$$a_{n+mt} = \begin{cases} 0, & \text{if } U_\Phi(n) \text{ is positive,} \\ 2, & \text{if } U_\Phi(n) \text{ is negative.} \end{cases}$$

Similar to the remark after Theorem 1.2, we see that Theorem 1.3 provides us a simple algorithm to generate all the possible pairs  $(a, \Phi) \in \mathbf{C} \times (\mathcal{F}_m \setminus \mathcal{F}_m^+)$  so that  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Below we give an example.

**Example 1.3.** Let  $m = 3$ ,  $k' = 2$  and  $k = 4$ . Set  $\mathcal{M} = \{1, 2, 3\}$ . Construct a  $4 \times \infty$  matrix  $U$  as below:

$$U = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 2 & \bar{2} & 2 & 2 & \bar{2} & 2 & \cdots \\ 2 & \bar{2} & 0 & 2 & \bar{2} & 2 & 2 & \bar{2} & 2 & \cdots \end{pmatrix},$$

where the sequence  $(U(n))_{n \geq 3}$  of the columns in  $U$  is periodic with period 3. Let  $a = .(020)^\infty$ . Then  $\mathcal{M}$ ,  $U$  and  $a$  satisfy the conditions (i)-(iv) in Theorem 1.3. Notice that  $U = U_\Phi$  for

$$\Phi = \{3^{-3}x, 3^{-3}x + .202, -3^{-3}x + .002(2\bar{2}2)^\infty, -3^{-3}x + .2\bar{2}0(2\bar{2}2)^\infty\}.$$

By Theorem 1.3,  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ .

So far, we have given a simple and complete characterization of those pairs  $(a, \Phi) \in \mathbf{C} \times \mathcal{F}_m$  satisfying  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  (Theorems 1.2 and 1.3). In Section 6, we will extend this result to the pairs  $(a, \Phi) \in \mathbf{C} \times \mathcal{F}$ . Indeed, when  $\mathcal{F}_m$  is replaced by  $\mathcal{F}$ , we will show that if  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ , then  $(U_\Phi(n))_{n \geq 1}$  is eventually periodic; furthermore, we can provide a finite algorithm to find all the pairs  $(a, \Phi) \in \mathbf{C} \times \mathcal{F}$  satisfying  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ , once the ratios of the maps in  $\Phi$  are pre-given (Theorems 6.1 and 6.2).

The paper is organized as follows. In Section 2, we prove Theorem 1.1. In Section 3, we provide the so-called addition and subtraction principles to judge whether  $a + b \in \mathbf{C}$  (or  $a - b \in \mathbf{C}$ ) for given  $a \in \mathbf{C}$  and  $b \in [0, 1]$ , by using the intrinsic

ternary expansions of  $a, b$ . In Section 4, we use them to formulate a criterion to judge whether  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  for given  $(a, \Phi) \in \mathbf{C} \times \mathcal{F}$ , by investigating certain compatible properties of the intrinsic matrix  $U_\Phi$  and the intrinsic ternary expansion of  $a$ . The proofs of our main results are given in Sections 5 and 6. In Section 7 we give some generalizations and remarks. A needed known result in number theory and some elementary facts about the Cantor set  $\mathbf{C}$  are given in Appendix A.

## 2. THE PROOF OF THEOREM 1.1

Recall that the Cantor set  $\mathbf{C}$  is obtained in the following way. Start with the unit interval  $K_0 = [0, 1]$ . Divide it into three equal sections and remove the open middle third. Thus we have  $K_1 = [0, 1/3] \cup [2/3, 1]$ . Then we continue inductively to obtain a sequence of closed sets  $(K_n)_{n=1}^\infty$ , so that  $K_n$  consists of  $2^n$  closed intervals of length  $3^{-n}$  obtained by removing the middle (open) one-third of each intervals in  $K_{n-1}$ . Finally the Cantor set is given by  $\mathbf{C} = \bigcap_{n=0}^\infty K_n$ . For  $n \geq 0$ , each interval in  $K_n$  is called an  $n$ -th basic interval.

Before proving Theorem 1.1, we shall first prove the following lemma.

**Lemma 2.1.** *Assume that  $\mathbf{F} \subseteq \mathbf{C}$  is a non-trivial self-similar set, generated by a linear IFS  $\Phi = \{\phi_i\}_{i=1}^k$  on  $\mathbb{R}$ . Then we have  $\log |\rho_i| / \log 3 \in \mathbb{Q}$  for each  $1 \leq i \leq k$ , where  $\rho_i$  denotes the contraction ratio of  $\phi_i$ .*

*Proof.* Fix  $\phi$  in  $\Phi$ . Let  $\rho$  be the contraction ratio of  $\phi$ . We prove that  $\log |\rho| / \log 3 \in \mathbb{Q}$ .

Let  $a$  be the fixed point of  $\phi$ . Then  $\phi(x) = a + \rho(x - a)$ . Let  $\mathbf{F}_\Phi$  denote the attractor of  $\Phi$ . Clearly,  $a \in \mathbf{F}_\Phi$ . Take  $b \in \mathbf{F}_\Phi$  such that  $b \neq a$ . Observing that  $\phi^n(b) \in \mathbf{F}_\Phi \subseteq \mathbf{C}$  and  $\phi^n(b) = a + \rho^n(b - a)$  for any  $n \in \mathbb{N}$ , we have

$$a + \rho^{2n}(b - a) \in \mathbf{C} \text{ and } 1 - a - \rho^{2n}(b - a) \in \mathbf{C}, \quad \forall n \in \mathbb{N}.$$

(Observe that  $\mathbf{C}$  is symmetric in the sense  $1 - \mathbf{C} = \mathbf{C}$ ). Hence there exist  $c \in \mathbf{C}$  and  $d > 0$  such that

$$c + \rho^{2n}d \in \mathbf{C}, \quad \forall n \in \mathbb{N}.$$

Suppose on the contrary that  $\log |\rho| / \log 3 \notin \mathbb{Q}$ . We will derive a contradiction. Pick  $0 < \epsilon < \rho^2$ . Since  $\log |\rho| / \log 3 \notin \mathbb{Q}$ , the set  $\{2n \log |\rho| + m \log 3 : n, m \in \mathbb{N}\}$  is

dense in  $\mathbb{R}$ . Therefore, we can find  $n, m \in \mathbb{N}$  such that

$$2n \log |\rho| + m \log 3 + \log d \in (0, \log(1 + \epsilon)),$$

that is,  $\rho^{2n} 3^m d \in (1, 1 + \epsilon)$ . We rewrite it as follows.

$$(2.1) \quad 3^{-m} < \rho^{2n} d < (1 + \epsilon) 3^{-m}.$$

Since  $0 < \epsilon < \rho^2$ , we have  $1/\epsilon > \rho^{-2}$ . This, combining with the first inequality in (2.1), implies that there exists  $p \in \mathbb{N}$  such that

$$(2.2) \quad 3^{-m} \epsilon \leq \rho^{2(n+p)} d < 3^{-m}.$$

Let  $I$  denote the  $m$ -th basic interval containing the point  $c$ , and  $J$  the  $m$ -th basic interval containing the point  $c + \rho^{2n} d$ . Due to (2.1),  $I$  and  $J$  are two different basic intervals with a gap of length  $3^{-m}$ ; then the second inequality in (2.1) implies that, the distance between  $c$  and the right endpoint of  $I$  is less than  $3^{-m} \epsilon$ . Using this information and (2.2), we see that the point  $c + \rho^{2(n+p)} d$  must be located in the gap between  $I$  and  $J$ , contradicting the fact that  $c + \rho^{2(n+p)} d \in \mathcal{C}$ .  $\square$

To prove Theorem 1.1, we need to use a result of Salem and Zygmund [7, 6] in the theory of trigonometric series. Let us consider a trigonometric series

$$\sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx),$$

where the variable  $x$  is real. In one of his pioneer works, Cantor (cf. [8]) showed that if this series converges everywhere to zero, it should vanish identically, i.e.,  $a_n = b_n = 0$  for all  $n$ . This work leads to the following.

**Definition 2.1.** *A subset  $\mathbf{E}$  of the circle  $[0, 2\pi)$  is called a set of uniqueness, if any trigonometric expansion*

$$\sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx),$$

*which converges to zero for  $x \in [0, 2\pi) \setminus \mathbf{E}$ , is identically zero; that is,  $a_n = b_n = 0$  for all  $n$ .*

It is an unsolved fundamental problem to classify all sets of uniqueness. The following significant result is due to Salem and Zygmund.

**Theorem 2.1** (cf. Chapter VI of [6]). *Let  $\mathbf{E}$  be a self-similar subset of  $(0, 2\pi)$  generated by an IFS  $\{\rho x + a, \rho x + b\}$  with  $0 < |\rho| < 1/2$ ,  $a \neq b$ . Then a necessary and sufficient condition for  $\mathbf{E}$  to be a set of uniqueness is that  $1/|\rho|$  is a Pisot number, i.e. an algebraic integer whose algebraic conjugates are all inside the unit disk.*

According to the above result, the Cantor set  $\mathbf{C}$  is a set of uniqueness. Of course, by definition, every subset of  $\mathbf{C}$  is a set of uniqueness.

*Proof of Theorem 1.1.* By Lemma 2.1, for each  $1 \leq i \leq k$ , the contraction ratio of  $\phi_i$  is of the form  $\pm 3^{-p_i}$ , where  $p_i \in \mathbb{Q}$ . Here we need to prove that  $p_i \in \mathbb{N}$ . Without loss of generality, we prove that  $p_1 \in \mathbb{N}$ . Since the attractor of  $\Phi$  is not a singleton, there exists  $i > 1$  such that  $\phi_i$  and  $\phi_1$  have different fixed points. Then the two maps  $\phi_1 \circ \phi_i$  and  $\phi_i \circ \phi_1$  are different and they have the same contraction ratio  $\rho$ . Since the attractor generated by  $\{\phi_1 \circ \phi_i, \phi_i \circ \phi_1\}$  is a subset of  $\mathbf{C}$ , it is a set of uniqueness. By Theorem 2.1,  $1/|\rho|$  is a Pisot number. Since  $1/|\rho| = 3^{p_1+p_i}$  and  $p_1 + p_i \in \mathbb{Q}$ , to guarantee that  $3^{p_1+p_i}$  is a Pisot number, we must have  $p_1 + p_i \in \mathbb{N}$ . Similarly, considering the maps  $\phi_1 \circ (\phi_i^2)$  and  $(\phi_i^2) \circ \phi_1$ , we also have  $p_1 + 2p_i \in \mathbb{N}$ . This forces  $p_1 \in \mathbb{N}$ .  $\square$

### 3. ADDITION AND SUBTRACTION PRINCIPLES

In this section we consider the following basic question: given  $a \in \mathbf{C}$  and  $b \in [0, 1]$ , how can we judge whether  $a + b \in \mathbf{C}$  or  $a - b \in \mathbf{C}$ ? As an answer, using the intrinsic ternary expansions we establish the so called *addition and subtraction principles* illustrated by Lemma 3.1 and Proposition 3.1.

**Definition 3.1.** For  $\mathbf{a} = (a_n)_{n=1}^\infty \in \{0, 2\}^\mathbb{N}$  and  $\mathbf{u} = (u_n)_{n=1}^\infty \in \{0, 2, \bar{2}\}^\mathbb{N}$ , say that  $(\mathbf{a}, \mathbf{u})$  is plus-admissible if  $a_n + u_n \in \{0, 2\}$  for each  $n$ ; in this case we define

$$\mathbf{a} \oplus \mathbf{u} = (a_n + u_n)_{n=1}^\infty.$$

Similarly, say that  $(\mathbf{a}, \mathbf{u})$  is minus-admissible if  $(\mathbf{a}, \bar{\mathbf{u}})$  is plus-admissible, where

$$(3.1) \quad \bar{\mathbf{u}} = (\bar{u}_n)_{n=1}^\infty$$

with  $\bar{i} = -i$  for  $i \in \{0, 2, \bar{2}\}$ .

**Remark 3.1.** It is direct to check that if  $(\mathbf{a}, \mathbf{u}) \in \{0, 2\}^{\mathbb{N}} \times \{0, 2, \bar{2}\}^{\mathbb{N}}$  is plus-admissible, then  $(\mathbf{a} \oplus \mathbf{u}, \bar{\mathbf{u}})$  is plus-admissible, i.e.,  $(\mathbf{a} \oplus \mathbf{u}, \mathbf{u})$  is minus-admissible, and  $\mathbf{a} = (\mathbf{a} \oplus \mathbf{u}) \oplus \bar{\mathbf{u}}$ .

Define  $\pi : \{0, 2, \bar{2}\}^{\mathbb{N}} \rightarrow [-1, 1]$  by

$$(3.2) \quad \pi(\mathbf{u}) = \sum_{n=1}^{\infty} u_n 3^{-n}, \quad \forall \mathbf{u} = (u_n)_{n=1}^{\infty}.$$

The following result directly follows from the above definitions.

**Lemma 3.1.** *Let  $\mathbf{a} \in \{0, 2\}^{\mathbb{N}}$  and  $\mathbf{u} \in \{0, 2, \bar{2}\}^{\mathbb{N}}$ . Define  $\bar{\mathbf{u}}$  as in (3.1). Then the following properties hold:*

(i) *If  $(\mathbf{a}, \mathbf{u})$  is plus-admissible, then*

$$\pi(\mathbf{a}) + \pi(\mathbf{u}) = \pi(\mathbf{a} \oplus \mathbf{u}) \in \mathcal{C}.$$

(ii) *If  $(\mathbf{a}, \mathbf{u})$  is minus-admissible, then*

$$\pi(\mathbf{a}) - \pi(\mathbf{u}) = \pi(\mathbf{a} \oplus \bar{\mathbf{u}}) \in \mathcal{C}.$$

We remark that the converse of the above lemma is not totally true. For example,  $\pi(02^{\infty}) + \pi(02^{\infty}) = 2/3 \in \mathcal{C}$ , though  $(02^{\infty}, 02^{\infty})$  is not plus-admissible; similarly  $\pi(20^{\infty}) + \pi(0\bar{2}^{\infty}) = 1/3 \in \mathcal{C}$ , though  $(20^{\infty}, 0\bar{2}^{\infty})$  is not plus-admissible. Nevertheless, we are going to show that except some special cases, the converse of Lemma 3.1 is true.

Let  $\sigma$  denote the left shift map on  $\{0, 2, \bar{2}\}^{\mathbb{N}}$ , i.e.,  $\sigma((u_n)_{n=1}^{\infty}) = (u_{n+1})_{n=1}^{\infty}$  for  $(u_n)_{n=1}^{\infty} \in \{0, 2, \bar{2}\}^{\mathbb{N}}$ . Recall that  $\Theta$  is defined as in (1.3). Define  $\Gamma, \Gamma' \subset \{0, 2\}^{\mathbb{N}} \times \Theta$  by

$$(3.3) \quad \Gamma = \{(\mathbf{a}, \mathbf{u}) : \exists k \geq 0 \text{ such that } (\sigma^k \mathbf{a}, \sigma^k \mathbf{u}) = (02^{\infty}, 02^{\infty}) \text{ or } (20^{\infty}, 0\bar{2}^{\infty})\},$$

$$(3.4) \quad \Gamma' = \{(\mathbf{a}, \mathbf{u}) : \exists k \geq 0 \text{ such that } (\sigma^k \mathbf{a}, \sigma^k \mathbf{u}) = (20^{\infty}, 02^{\infty}) \text{ or } (02^{\infty}, 0\bar{2}^{\infty})\}.$$

The main result of this section is the following.

**Proposition 3.1.** *Let  $\mathbf{a} = (a_n)_{n=1}^{\infty} \in \{0, 2\}^{\mathbb{N}}$  and  $\mathbf{u} = (u_n)_{n=1}^{\infty} \in \Theta$ . Then we have the following statements:*

- (i) Assume that  $(\mathbf{a}, \mathbf{u}) \notin \Gamma$  and  $\pi(\mathbf{a}) + \pi(\mathbf{u}) \in \mathbf{C}$ . Then the pair  $(\mathbf{a}, \mathbf{u})$  is plus-admissible.
- (ii) Assume that  $(\mathbf{a}, \mathbf{u}) \notin \Gamma'$  and  $\pi(\mathbf{a}) - \pi(\mathbf{u}) \in \mathbf{C}$ . Then the pair  $(\mathbf{a}, \mathbf{u})$  is minus-admissible.

Before proving Proposition 3.1, we first give a lemma.

**Lemma 3.2.** *Let  $n \in \mathbb{N}$ ,  $c_1 \dots c_n \in \{0, 2\}^n$  and  $x \in \mathbb{R}$ . Then*

- (1) If  $c_n = 0$ , then

$$.c_1 \dots c_n + 3^{-n}x \notin \mathbf{C} \text{ if } x \in (-3, 0) \cup (1, 2).$$

- (2) If  $c_n = 2$ , then

$$.c_1 \dots c_n + 3^{-n}x \notin \mathbf{C} \text{ if } x \in (-1, 0) \cup (1, 4).$$

*Proof.* Notice that  $.c_1 \dots c_n$  is the left endpoint of an  $n$ -th basic interval  $I$  in defining  $\mathbf{C}$  (cf. the first paragraph in Section 2). Furthermore when  $c_n = 0$ , there is an open interval of length  $\geq 3^{-n+1}$  on the left hand side of  $I$  containing no points in  $\mathbf{C}$ , whilst there is an open interval of length  $3^{-n}$  on the right hand side of  $I$  containing no points in  $\mathbf{C}$ , from which (1) follows. Similarly we have (2).  $\square$

*Proof of Proposition 3.1.* We first prove (i). Assume that  $(\mathbf{a}, \mathbf{u}) \notin \Gamma$  and  $\pi(\mathbf{a}) + \pi(\mathbf{u}) \in \mathbf{C}$ . Suppose on the contrary that  $(\mathbf{a}, \mathbf{u})$  is not plus-admissible. Let  $n \geq 1$  be the smallest number such that  $a_n + u_n \notin \{0, 2\}$ . Then either  $(a_n, u_n) = (0, \bar{2})$ , or  $(a_n, u_n) = (2, 2)$ . Letting  $c_k = a_k + u_k$ , by the minimality of  $n$ , we have  $c_k \in \{0, 2\}$  for  $k = 1, \dots, n-1$ .

First consider the case when  $(a_n, u_n) = (0, \bar{2})$ . Since  $\mathbf{u} \in \Theta$ , we have  $u_1 \neq \bar{2}$  and thus  $n > 1$ . Furthermore, since  $\mathbf{u}$  does not end with  $2\bar{2}^\infty$  and  $\bar{2}2^\infty$ , we have

$$-3^{-n+1} = \sum_{k=n}^{\infty} (-2)3^{-k} \leq \sum_{k=n}^{\infty} (a_k + u_k)3^{-k} < (-2)3^{-n} + \sum_{k=n+1}^{\infty} 4 \cdot 3^{-k} = 0.$$

That is,

$$(3.5) \quad \sum_{k=n}^{\infty} (a_k + u_k)3^{-k} \in [-3^{-n+1}, 0).$$

Notice that

$$\pi(\mathbf{a}) + \pi(\mathbf{u}) = .c_1 \dots c_{n-1} + \sum_{k=n}^{\infty} (a_k + u_k)3^{-k} \in \mathbf{C}.$$

By (3.5) and Lemma 3.2, we have  $c_{n-1} = 2$  and  $\sum_{k=n}^{\infty} (a_k + u_k)3^{-k} = -3^{-n+1}$ , where the second equality implies that  $a_k = 0$  and  $u_k = \bar{2}$  for  $k \geq n$ . Notice that  $u_{n-1} \neq 2$  (for otherwise  $\mathbf{u}$  ends with  $\bar{2}2^{\infty}$ ), and  $a_{n-1} + u_{n-1} = c_{n-1} = 2$ . We have  $a_{n-1} = 2$  and  $u_{n-1} = 0$ . Therefore  $\sigma^{n-2}\mathbf{a} = 20^{\infty}$  and  $\sigma^{n-2}\mathbf{u} = 0\bar{2}^{\infty}$ . Hence  $(\mathbf{a}, \mathbf{u}) \in \Gamma$ , leading to a contradiction.

Next consider the case when  $(a_n, u_n) = (2, 2)$ . We have

$$3^{-n+1} = 4 \cdot 3^{-n} + \sum_{k=n+1}^{\infty} (-2)3^{-k} < \sum_{k=n}^{\infty} (a_k + u_k)3^{-k} \leq \sum_{k=n}^{\infty} 4 \cdot 3^{-k} = 2 \cdot 3^{-n+1}.$$

That is,

$$(3.6) \quad \sum_{k=n}^{\infty} (a_k + u_k)3^{-k} \in (3^{-n+1}, 2 \cdot 3^{-n+1}].$$

By (3.6) and Lemma 3.2, we have  $c_{n-1} = 0$  and  $\sum_{k=n}^{\infty} (a_k + u_k)3^{-k} = 2 \cdot 3^{-n+1}$ , where the second equality implies that  $a_k = 2$  and  $u_k = 2$  for  $k \geq n$ . Notice that  $u_{n-1} \neq \bar{2}$  (for otherwise  $\mathbf{u}$  ends with  $\bar{2}2^{\infty}$ ), and  $a_{n-1} + u_{n-1} = c_{n-1} = 0$ . We have  $a_{n-1} = 0$  and  $u_{n-1} = 0$ . Therefore  $\sigma^{n-2}\mathbf{a} = \sigma^{n-2}\mathbf{u} = 02^{\infty}$ . Hence  $(\mathbf{a}, \mathbf{u}) \in \Gamma$ , leading to a contradiction. This finishes the proof of (i).

Now we prove (ii). Assume that  $(\mathbf{a}, \mathbf{u}) \notin \Gamma'$  and  $\pi(\mathbf{a}) - \pi(\mathbf{u}) \in \mathcal{C}$ . Take  $\mathbf{c} \in \{0, 2\}^{\mathbb{N}}$  so that  $\pi(\mathbf{c}) = \pi(\mathbf{a}) - \pi(\mathbf{u})$ . Then  $\pi(\mathbf{c}) + \pi(\mathbf{u}) = \pi(\mathbf{a}) \in \mathcal{C}$ . Note that  $(\mathbf{a}, \mathbf{u}) \notin \Gamma'$  implies  $(\mathbf{c}, \mathbf{u}) \notin \Gamma$ . Hence by (i),  $(\mathbf{c}, \mathbf{u})$  is plus-admissible. By Lemma 3.1,  $\mathbf{a} = \mathbf{c} \oplus \mathbf{u}$ . By Remark 3.1,  $(\mathbf{c} \oplus \mathbf{u}, \bar{\mathbf{u}})$  is plus-admissible. Hence  $(\mathbf{a}, \bar{\mathbf{u}})$  is plus-admissible, i.e.,  $(\mathbf{a}, \mathbf{u})$  is minus-admissible. This proves (ii).  $\square$

In the end of this section, we give one more definition.

**Definition 3.2.** In general, for  $\mathbf{a} \in \{0, 2\}^{\mathbb{N}}$  and  $\mathbf{u}_1, \dots, \mathbf{u}_n \in \{0, 2, \bar{2}\}^{\mathbb{N}}$ , we say that  $(\mathbf{a}, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$  is plus-admissible if  $(\mathbf{a}_j, \mathbf{u}_{j+1})$  is plus-admissible for  $j = 0, \dots, n-1$ , where  $\mathbf{a}_j$ 's are defined inductively by  $\mathbf{a}_0 := \mathbf{a}$ ,  $\mathbf{a}_1 := \mathbf{a}_0 \oplus \mathbf{u}_1$ , ...,  $\mathbf{a}_{n-1} = \mathbf{a}_{n-2} \oplus \mathbf{u}_{n-1}$ .

**Remark 3.2.** Let  $\mathbf{a} \in \{0, 2\}^{\mathbb{N}}$  and  $\mathbf{u}_i = (u_{i,p})_{p=1}^{\infty} \in \{0, 2, \bar{2}\}^{\mathbb{N}}$ ,  $i = 1, \dots, n$ . It is easy to see that  $(\mathbf{a}, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$  is plus-admissible if and only if for each  $p \in \mathbb{N}$ , the following properties hold:

- (i) Neither two consecutive letters 2, nor two consecutive  $\bar{2}$  appear in the finite sequence (might be empty) obtained by deleting all letters 0 from  $(u_{i,p})_{i=1}^n$ ;

- (ii)  $a_p = 0$  if the first letter in the above sequence is 2, and  $a_p = 2$  if the first letter is  $\bar{2}$ .

#### 4. A MATRIX-VALUED FUNCTION $V$ AND MATCHING PROPERTIES

Let  $(a, \Phi)$  be a pair so that  $a \in \mathcal{C}$  and  $\Phi = \{\phi_i\}_{i=1}^k \in \mathcal{F}$ . In this section we define a matrix-valued function  $V$  over  $\{1, \dots, k\}^* := \bigcup_{n \geq 1} \{1, \dots, k\}^n$ , and show that  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$  if and only if certain matching properties hold for  $V$ . This criterion plays a key role in the proofs of our main results.

Suppose that the maps  $\phi_i$  are of the form  $\phi_i(x) = s_i 3^{-m_i} x + d_i$  as in (1.2). Let  $a = \sum_{n=1}^{\infty} a_n 3^{-n}$  and  $d_i = \sum_{n=1}^{\infty} u_{i,n} 3^{-n}$  ( $i = 1, \dots, k$ ) be the intrinsic ternary expansions of  $a, d_1, \dots, d_k$ . Write  $\mathbf{a} = (a_n)_{n=1}^{\infty}$  and  $\mathbf{u}_i = (u_{i,n})_{n=1}^{\infty}$ . Then  $\mathbf{a} \in \{0, 2\}^{\mathbb{N}}$  and  $\mathbf{u}_i \in \Theta$  for  $1 \leq i \leq k$ , where  $\Theta$  is defined as in (1.3).

For  $n \in \mathbb{N}$  and  $\mathbf{x} = x_1 \dots x_n \in \{1, \dots, k\}^n$ , define an  $n \times \infty$  matrix  $V(\mathbf{x}) := V_\Phi(\mathbf{x})$  by

$$(4.1) \quad V(\mathbf{x}) = (v_{j,p})_{1 \leq j \leq n, p \geq 1},$$

where  $\mathbf{v}_1 = (v_{1,p})_{p=1}^{\infty} := \mathbf{u}_{x_1}$  and for  $2 \leq j \leq n$ ,

$$\mathbf{v}_j = (v_{j,p})_{p=1}^{\infty} := \begin{cases} 0^{m_{x_1} + \dots + m_{x_{j-1}}} \mathbf{u}_{x_j}, & \text{if } s_{x_1} \dots s_{x_{j-1}} = 1, \\ 0^{m_{x_1} + \dots + m_{x_{j-1}}} \overline{\mathbf{u}}_{x_j}, & \text{if } s_{x_1} \dots s_{x_{j-1}} = -1. \end{cases}$$

The above matrix-valued function  $V(\cdot)$  plays an important role in our further analysis. Take the convention that  $|\mathbf{x}| = n$  if  $\mathbf{x} \in \{1, \dots, k\}^n$ . The following simple lemma just follows from the definition of  $V(\cdot)$ .

**Lemma 4.1.** *Let  $\mathbf{x} = x_1 \dots x_i$ ,  $\mathbf{z} = z_1 \dots z_j \in \{1, \dots, k\}^*$  and  $n \in \mathbb{N}$ . then*

- (i) *For  $1 \leq p \leq i$ , the  $(p, n)$ -entry of  $V(\mathbf{x})$  is equal to*

$$s_{x_1} \dots s_{x_{p-1}} u_{x_p, n - \sum_{t=1}^{p-1} m_{x_t}}.$$

- (ii) *If  $d_{x_p} = 0$ , then the  $(p, n)$ -entry of  $V(\mathbf{x})$  is zero.*
- (iii) *If  $s_{\mathbf{z}} := s_{z_1} \dots s_{z_j} = 1$ , then the  $n$ -th column of  $V(\mathbf{x})$  is a suffix of the  $(n+q)$ -th column of  $V(\mathbf{zx})$ , where  $q = \sum_{t=1}^j m_{z_t}$ ; that is, the last  $|\mathbf{x}|$ -entries in the  $(n+q)$ -th column of  $V(\mathbf{zx})$  coincide with the entries in the  $n$ -th column of  $V(\mathbf{x})$ . Moreover if  $s_{\mathbf{z}} = -1$ , then the  $n$ -th column of  $V(\mathbf{x})$  is a suffix of the  $(n+q)$ -th column of  $(-V(\mathbf{zx}))$ .*

Now we are ready to formulate our criterion.

**Theorem 4.1.**  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  if and only if for every  $\mathbf{x} \in \{1, \dots, k\}^*$  and  $p \in \mathbb{N}$ :

- (i) Neither two consecutive letters 2, nor two consecutive  $\bar{2}$  appear in the finite sequence (might be empty) obtained by deleting all zero entries from the  $p$ -th column of the matrix  $V(\mathbf{x})$ ;
- (ii)  $a_p = 0$  if the first letter in the above reduced sequence is 2, and  $a_p = 2$  if the first letter is  $\bar{2}$ .

**Remark 4.1.** Although this criterion depends on the matching properties of  $V(\mathbf{x})$  over all  $\mathbf{x} \in \{1, \dots, k\}^*$ , in the coming Sections 5-6 we will show that essentially it is enough to check the matching properties of  $V(\mathbf{x})$  for finitely many  $\mathbf{x}$ .

Before proving Theorem 4.1, we first give several lemmas.

**Lemma 4.2.**  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  if and only if  $a + \phi_{x_1 \dots x_n}(0) \in \mathbf{C}$  for every  $n \in \mathbb{N}$  and  $x_1 \dots x_n \in \{1, \dots, k\}^n$ , where  $\phi_{x_1 \dots x_n} := \phi_{x_1} \circ \dots \circ \phi_{x_n}$ .

*Proof.* Since  $\Phi \in \mathcal{F}$ , we have  $0 \in \mathbf{F}_\Phi$ . Therefore  $\mathbf{F}_\Phi$  is just the closure of the set  $\{\phi_{x_1 \dots x_n}(0) : x_1 \dots x_n \in \{1, \dots, k\}^n, n \in \mathbb{N}\}$ , from which the lemma follows.  $\square$

**Lemma 4.3.** Assume that  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Then the following properties hold:

- (i) For each  $1 \leq i \leq k$ ,  $(\mathbf{a}, \mathbf{u}_i)$  is plus-admissible.
- (ii) Let  $n > 1$  and  $x_1 \dots x_n \in \{1, \dots, k\}^n$ . Write

$$\alpha = a + \phi_{x_1 \dots x_{n-1}}(0), \quad \beta = 3^{-(m_{x_1} + \dots + m_{x_{n-1}})} d_{x_n}.$$

Let  $\alpha = \sum_{n=1}^{\infty} c_n 3^{-n}$  and  $\beta = \sum_{n=1}^{\infty} u_n 3^{-n}$  be the intrinsic ternary expansions of  $\alpha, \beta$ . Set  $\mathbf{c} = (c_n)_{n=1}^{\infty}$  and  $\mathbf{u} = (u_n)_{n=1}^{\infty}$ . Then the pair  $(\mathbf{c}, \mathbf{u})$  is plus-admissible if  $s_{x_1} \dots s_{x_{n-1}} = 1$ , and minus-admissible if  $s_{x_1} \dots s_{x_{n-1}} = -1$ .

*Proof.* We only prove (ii), since (i) can be viewed as the variant of (ii) corresponding to the particular case when  $n = 1$ .

First assume that  $s_{x_1} \dots s_{x_{n-1}} = 1$ . In this case,  $\alpha + \beta = a + \phi_{x_1 \dots x_n}(0)$ . By Lemma 4.2, we have  $\alpha \in \mathbf{C}$  and  $\alpha + \beta \in \mathbf{C}$ . By Proposition 3.1(i), to show that  $(\mathbf{c}, \mathbf{u})$  is plus-admissible, it suffices to show that  $(\mathbf{c}, \mathbf{u}) \notin \Gamma$ . Assume on the contrary,  $(\mathbf{c}, \mathbf{u}) \in \Gamma$ . Then it is direct to check that for  $t \in \mathbb{N}$ ,

$$(4.2) \quad \pi(\mathbf{c}) + 3^{-t} \pi(\mathbf{u}) \notin \mathbf{C} \text{ when } t \text{ is large enough.}$$

However, by Lemma 4.2,  $a + \phi_{x_1 \dots x_{n-1} 1^p x_n}(0) \in \mathcal{C}$  for each  $p \in \mathbb{N}$ . Notice that

$$\begin{aligned} a + \phi_{x_1 \dots x_{n-1} 1^p x_n}(0) &= a + \phi_{x_1 \dots x_{n-1}}(0) + s_{x_1} \dots s_{x_{n-1}} 3^{-pm_1} 3^{-(m_{x_1} + \dots + m_{x_{n-1}})} d_{x_n} \\ &= \alpha + 3^{-pm_1} \beta = \pi(\mathbf{c}) + 3^{-pm_1} \pi(\mathbf{u}). \end{aligned}$$

Hence we have  $\pi(\mathbf{c}) + 3^{-pm_1} \pi(\mathbf{u}) \in \mathcal{C}$  for any  $p \in \mathbb{N}$ . This contradicts (4.2).

Next assume that  $s_{x_1} \dots s_{x_{n-1}} = -1$ . A similar argument shows that  $(\mathbf{c}, \mathbf{u})$  is minus-admissible. We omit the details.  $\square$

*Proof of Theorem 4.1.* We first prove the ‘only if’ part. Let  $n \in \mathbb{N}$  and  $\mathbf{x} = x_1 \dots x_n \in \{1, \dots, k\}^n$ . Assume that the combinatoric properties (i)-(ii) hold for  $\mathbf{x}$ . Let  $\mathbf{v}_1, \dots, \mathbf{v}_n$  be the rows of  $V(\mathbf{x})$ . By Remark 3.2,  $(\mathbf{a}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$  is plus-admissible in the sense that  $(\mathbf{a}_{j-1}, \mathbf{v}_j)$  is plus-admissible for each  $1 \leq j \leq n$ , where  $\mathbf{a}_0 := \mathbf{a}$ ,  $\mathbf{a}_1 := \mathbf{a}_0 \oplus \mathbf{v}_1, \dots$ , and  $\mathbf{a}_j = \mathbf{a}_{j-1} \oplus \mathbf{v}_j$ . Applying Lemma 3.1 repeatedly, we have

$$(4.3) \quad a + d_{x_1} + s_{x_1} 3^{-m_{x_1}} d_{x_2} + \dots + (s_{x_1} \dots s_{x_{j-1}}) 3^{-(m_{x_1} + \dots + m_{x_{j-1}})} d_{x_j} \in \mathcal{C}$$

for  $j = 1, \dots, n$ . That is,  $a + \phi_{x_1 \dots x_j}(0) \in \mathcal{C}$  for  $1 \leq j \leq n$ . Letting  $\mathbf{x}$  vary over  $\{1, \dots, k\}^*$ , we have  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$  by Lemma 4.2.

Next we prove the ‘if’ part. Assume that  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ . Then for any  $n \in \mathbb{N}$  and  $\mathbf{x} = x_1 \dots x_n \in \{1, \dots, k\}^n$ , (4.3) holds for any  $1 \leq j \leq n$ . By Lemma 4.3,  $(\mathbf{a}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$  is plus-admissible. Hence by Remark 3.2, the combinatoric properties (i)-(ii) hold.  $\square$

## 5. THE PROOFS OF THEOREMS 1.2 AND 1.3

*Proof of Theorem 1.2.* We first prove the ‘if’ part of the theorem. Let  $\mathbf{x} \in \{1, \dots, k\}^*$  and  $p \in \mathbb{N}$ . The assumptions (i)-(iii) on  $U_\Phi$  guarantee that the  $p$ -th column of  $V(\mathbf{x})$  (cf. (4.1)) contains at most one element in  $\{2, \bar{2}\}$ . It follows that property (i) in Theorem 4.1 holds. Furthermore the assumption (iv) guarantees that property (ii) in Theorem 4.1 holds. Since  $(\mathbf{x}, p)$  is arbitrarily taken from  $\{1, \dots, k\}^* \times \mathbb{N}$ , by Theorem 4.1,  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ .

Next we prove the ‘only if’ part. Assume that  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ . Set

$$(5.1) \quad \mathcal{M} = \{n \in \mathbb{N} : U_\Phi(n) \neq \{0, 0, \dots, 0\}^T\}.$$

To prove that (i)-(iv) in Theorem 1.2 hold, it suffices to prove that the following two properties on  $U_\Phi = (u_{i,n})_{1 \leq i \leq k, n \geq 1}$  and  $\mathbf{a} = (a_n)_{n=1}^\infty$  hold, as they are equivalent to properties (i)-(iv):

- (1) If  $u_{i,p} = 2$  for some pair  $(i, p)$  satisfying  $1 \leq i \leq k$  and  $p \geq 1$ , then  $u_{j,p} \in \{0, 2\}$ ,  $u_{j,p+mn} = 0$  and  $a_{p+m(n-1)} = 0$  for any  $1 \leq j \leq k$  and  $n \geq 1$ .
- (2) If  $u_{i,p} = \bar{2}$  for some pair  $(i, p)$  satisfying  $1 \leq i \leq k$  and  $p \geq 1$ , then  $u_{j,p} \in \{0, \bar{2}\}$ ,  $u_{j,p+mn} = 0$  and  $a_{p+m(n-1)} = 2$  for any  $1 \leq j \leq k$  and  $n \geq 1$ .

Without loss of generality, we only prove (1). The proof of (2) is similar. Assume that  $u_{i,p} = 2$  for some pair  $(i, p)$  satisfying  $1 \leq i \leq k$  and  $p \geq 1$ . Let  $n \geq 1$ . Then the  $(p + m(n - 1))$ -th column of the matrix  $V(1^{n-1}i)$  is  $(0, \dots, 0, 2)^T$ , in which the first non-zero entry is 2. Hence by property (ii) in Theorem 4.1,  $a_{p+m(n-1)} = 0$ . This also implies that  $u_{j,p} \neq \bar{2}$  (and  $u_{j,p+mn} \neq \bar{2}$  as well) for any  $1 \leq j \leq k$ ; otherwise if  $u_{j,p} = \bar{2}$  for some  $j$ , then a similar argument shows that  $a_{p+m(n-1)} = 2$ , leading to a contradiction. Now notice that the  $(p + mn)$ -th column of the matrix  $V(j1^{n-1}i)$  is  $(u_{j,p+mn}, 0, \dots, 0, 2)$ . Since  $u_{j,p+mn} \neq \bar{2}$ , we get  $u_{j,p+mn} = 0$  from the combinatoric property (i) in Theorem 4.1. This finishes the proof of (1).  $\square$

*Proof of Theorem 1.3.* We first prove the ‘only if’ part of the theorem. Assume that  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Define

$$(5.2) \quad \mathcal{M} = \{n \in \mathbb{N} : U_\Phi(n) \neq \{0, 0, \dots, 0\}^T, \text{ but } U_\Phi(n') = \{0, 0, \dots, 0\}^T \\ \text{for all } n' < n \text{ with } n' \equiv n \pmod{m}\}.$$

Clearly, the elements in  $\mathcal{M}$  are incongruent modulo  $m$ . To prove the desired properties (ii)-(iv) for  $\mathcal{M}$  and  $U_\Phi$  in the theorem, it is equivalent to prove Properties 1-2 listed as below:

*Property 1: If  $u_{i,p} = 2$  for some pair  $(i, p) \in \{1, \dots, k\} \times \mathbb{N}$ , then*

$$a_{p+mn} = 0 \quad \text{for all } n \geq 0, \\ u_{j,p+mn} = \begin{cases} 0 & \text{or } 2, & \text{if } 1 \leq j \leq k \text{ and } n = 0, \\ 0, & \text{if } 1 \leq j \leq k' \text{ and } n \geq 1, \\ 2, & \text{if } k' < j \leq k \text{ and } n \geq 1. \end{cases}$$

Property 2: If  $u_{i,p} = \bar{2}$  for some pair  $(i,p) \in \{1, \dots, k\} \times \mathbb{N}$ , then

$$\begin{aligned} a_{p+mn} &= 2 \quad \text{for all } n \geq 0, \\ u_{j,p+mn} &= \begin{cases} 0 \quad \text{or} \quad \bar{2}, & \text{if } 1 \leq j \leq k \text{ and } n = 0, \\ 0, & \text{if } 1 \leq j \leq k' \text{ and } n \geq 1, \\ \bar{2}, & \text{if } k' < j \leq k \text{ and } n \geq 1. \end{cases} \end{aligned}$$

Without loss of generality, we prove Property 1 only. The proof of Property 2 is essentially identical to that of Property 1.

First we consider the case that  $u_{i,p} = 2$  for some pair  $(i,p)$  with  $i \leq k'$  and  $p \in \mathbb{N}$ . Let  $\Phi' = \{\phi_i\}_{i=1}^{k'}$ . Then  $\Phi' \in \mathcal{F}_m^+$  and  $a + \mathbf{F}_{\Phi'} \subseteq a + \mathbf{F}_{\Phi} \subseteq \mathbf{C}$ . Applying Theorem 1.2 to the pair  $(a, \Phi')$ , we have  $a_{p+mn} = 0$  for all  $n \geq 0$ ,  $u_{j,p} \in \{0, 2\}$  for  $j \leq k'$  and  $u_{j,p+mn} = 0$  if  $n \geq 1$ ,  $j \leq k'$ . Next suppose  $j \in (k', k]$ . Notice that the  $p$ -th column of  $V(j)$  is  $(u_{j,p})$ , and  $a_p = 0$ . By Theorem 4.1(ii), we get  $u_{j,p} \in \{0, 2\}$ . Let  $n \geq 1$ . Since the  $(p + mn)$ -th column of  $V(j1^{n-1}i)$  is

$$(u_{j,p+mn}, \underbrace{0, \dots, 0}_{n-1}, \bar{2})^T.$$

by Theorem 4.1(i),  $u_{j,p+mn} \in \{0, 2\}$ . If  $u_{j,p+mn} = 0$ , then the first non-zero letter in the above column vector is  $\bar{2}$ , which forces  $a_{p+mn} = 2$  by Theorem 4.1(ii), leading to a contradiction. Hence we have  $u_{j,p+mn} = 2$ .

Next we consider the case that  $u_{i,p} = 2$  for some pair  $(i,p)$  with  $k' < i \leq k$  and  $p \in \mathbb{N}$ . Let  $n \geq 0$ . The  $(p + mn)$ -th column of the matrix  $V(1^n i)$  is  $(\underbrace{0, \dots, 0}_n, 2)^T$ , hence by Theorem 4.1(ii),  $a_{p+mn} = 0$ . Let  $1 \leq j \leq k$ . The  $(p + mn)$ -th column of the matrix  $V(j)$  is  $(u_{j,p+mn})$ . Since  $a_{p+mn} = 0$ , by Theorem 4.1(ii), we have  $u_{j,p+mn} \in \{0, 2\}$ . Next assume  $n \geq 1$ . We consider the cases  $j \leq k'$  and  $j > k'$  separately. First suppose that  $j \leq k'$ . In this case, the  $(p + mn)$ -th column of the  $V(j1^{n-1}i)$  is  $(u_{j,p+mn}, \underbrace{0, \dots, 0}_{n-1}, 2)^T$ . Hence by Theorem 4.1(i),  $u_{j,p+mn} \in \{0, \bar{2}\}$ , which forces  $u_{j,p+mn} = 0$  (since we have proved  $u_{j,p+mn} \in \{0, 2\}$ ). Next suppose that  $j > k'$ . Since  $s_j = -1$ , the  $(p + mn)$ -th column of the  $V(j1^{n-1}i)$  is

$$(u_{j,p+mn}, \underbrace{0, \dots, 0}_{n-1}, \bar{2})^T.$$

If  $u_{j,p+mn} = 0$ , then by Theorem 4.1(ii),  $a_{p+mn} = 2$ , leads to a contradiction (since we have proved  $a_{p+mn} = 0$ ). Hence we have  $u_{j,p+mn} = 2$ . This finishes the proof of Property 1.

Next we prove the ‘if’ part of Theorem 1.3. Assume that there exists a finite non-empty set  $\mathcal{M} \subset \mathbb{N}$  so that properties (i)-(iv) in the theorem hold. Then it is direct to check that  $\mathcal{M}$  is given by (5.2) and Properties 1-2 hold. Let  $n \in \mathbb{N}$  and  $\mathbf{x} = x_1 \dots x_n \in \{1, \dots, k\}^*$ . We show below that Properties (i)-(ii) in Theorem 4.1 hold, from which  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$  follows.

Let  $p \in \mathbb{N}$  and consider the  $p$ -th column  $(v_{1,p}, v_{2,p}, \dots, v_{n,p})^T$  in the matrix  $V(\mathbf{x})$ . Assume that this vector contains at least one non-zero entry. Let  $s$  be the smallest integer so that  $v_{s,p} \neq 0$ . Without loss of generality, assume  $v_{s,p} = 2$ . We show below that  $a_p = 0$ , i.e., property (ii) in Theorem 4.1 holds. According to the definition of  $V(\mathbf{x})$ , we have  $u_{x_i,p-(i-1)m} = 0$  for  $i < s$  and is not equal to 0 for  $i = s$ . Now since  $u_{x_s,p-(s-1)m} \neq 0$ , by Properties 1-2, we have

$$u_{j,p-(i-1)m} = u_{x_s,p-(s-1)m} \neq 0$$

if  $i < s$  and  $j > k'$ . It forces that  $x_i \leq k'$  for each  $i < s$ . Hence  $u_{x_s,p-(s-1)m} = v_{s,p} = 2$ . By Properties 1-2, we have  $a_p = 0$ .

Now we show that property (i) in Theorem 4.1 also holds for the vector

$$(v_{1,p}, v_{2,p}, \dots, v_{n,p})^T,$$

i.e., if we delete all entries 0 from this vector, neither two consecutive letters 2 nor two consecutive letters  $\bar{2}$  appear in the new vector. Assume this is not true. Without loss of generality, assume that two consecutive letters 2 appear. That is, there exist  $1 \leq s < t \leq k$  such that  $v_{s,p} = v_{t,p} = 2$  and  $v_{i,p} = 0$  for all  $s < i < t$ . Hence  $u_{x_s,p-(s-1)m} \neq 0$ ,  $u_{x_t,p-(t-1)m} \neq 0$  and  $u_{x_i,p-(i-1)m} = 0$  for  $s < i < t$ . Since  $u_{x_t,p-(t-1)m} \neq 0$ , by Properties 1-2, we have

$$(5.3) \quad u_{j,p-(j'-1)m} = \begin{cases} 0, & \text{if } j \leq k' \text{ and } j' < t, \\ u_{x_t,p-(t-1)m}, & \text{if } j > k' \text{ and } j' < t. \end{cases}$$

Since  $u_{x_s,p-(s-1)m} \neq 0$ , by (5.3), we have  $x_s > k'$  and  $u_{x_s,p-(s-1)m} = u_{x_t,p-(t-1)m}$ . Furthermore, since  $u_{x_i,p-(i-1)m} = 0$  for  $s < i < t$ , by (5.3),  $x_i \leq k'$  for  $s < i < t$ . Since  $x_s > k'$  and  $x_i \leq k'$  for  $s < i < t$ , we have  $(v_{s,p}, v_{t,p})$  equals  $(u_{x_s,p-(s-1)m}, \overline{u_{x_t,p-(t-1)m}})$  or  $(\overline{u_{x_s,p-(s-1)m}}, u_{x_t,p-(t-1)m})$ , i.e.,  $(2, \bar{2})$  or  $(\bar{2}, 2)$ . It leads to a contradiction. This finishes the proof of the theorem.  $\square$

## 6. GENERAL SELF-SIMILAR SUBSETS OF $\mathcal{C}$

In this section we characterize all the self-similar subsets of  $\mathcal{C}$ . The main results are Theorems 6.1 and 6.2, which claim that for a given pair  $(a, \Phi) \in \mathcal{C} \times \mathcal{F}$ , there is a finite algorithm (depending on the contraction ratios of maps in  $\Phi$ ) to judge whether  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ .

Let  $a \in \mathcal{C}$  and  $\Phi = \{\phi_i\}_{i=1}^k \in \mathcal{F}$ , where the maps  $\phi_i$  are of the form  $\phi_i(x) = s_i 3^{-m_i} x + d_i$  as in (1.2). Throughout this section, we set

$$\begin{aligned}
 \ell &:= \gcd(m_i : 1 \leq i \leq k, d_i = 0), \\
 r &:= \gcd(m_1, \dots, m_k), \\
 (6.1) \quad L &:= \left( \min_{1 \leq i \leq k} \frac{m_i}{r} - 1 \right) \left( \max_{1 \leq j \leq k} \frac{m_j}{r} - 1 \right), \\
 D &:= \max_{1 \leq i \leq k} m_i,
 \end{aligned}$$

and

$$(6.2) \quad \Lambda := \left\{ \sum_{i=1}^k y_i m_i : y_i \in \mathbb{N} \cup \{0\} \text{ for } 1 \leq i \leq k \right\},$$

where gcd means greatest common divisor. By Lemma A.1,

$$(6.3) \quad r y \in \Lambda \text{ for any integer } y \geq L.$$

Let  $a = \sum_{n=1}^{\infty} a_n 3^{-n}$  and  $d_i = \sum_{n=1}^{\infty} u_{i,n} 3^{-n}$  ( $i = 1, \dots, k$ ) be the intrinsic ternary expansions of  $a, d_1, \dots, d_k$ . Then the intrinsic translation matrix of  $\Phi$  is of the form  $U_\Phi = (u_{i,n})_{1 \leq i \leq k, n \geq 1}$ .

**6.1. Necessary conditions.** In this subsection we give some necessary conditions so that  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ . We begin with the following lemma.

**Lemma 6.1.** *Assume that  $a + \mathbf{F}_\Phi \subseteq \mathcal{C}$ . Suppose that two entries  $u_{i,n}$  and  $u_{j,n'}$  of  $U_\Phi$  satisfy that*

$$(6.4) \quad u_{i,n} \neq 0, \quad u_{j,n'} \neq 0 \quad \text{and} \quad n - m_i - n' \in \Lambda.$$

*Then  $(u_{i,n+tm_1})_{t \geq 0}$  is a constant sequence.*

*Proof.* We assume, without loss of generality, that  $u_{i,n} = 2$ . (The case when  $u_{i,n} = \bar{2}$  can be handled similarly.) Since  $n - m_i - n' \in \Lambda$ , there exist non-negative integers  $y_1, y_2, \dots, y_k$  such that  $n - m_i - n' = y_1 m_1 + y_2 m_2 + \dots + y_k m_k$ .

For any  $t \geq 0$ , since the  $(n + tm_1)$ -column of the matrix  $V(1^t i)$  is

$$(0, \dots, 0, u_{i,n})^T = (0, \dots, 0, 2)^T,$$

we have  $a_{n+tm_1} = 0$  by Theorem 4.1(ii).

Let  $\mathbf{v} := (v_1, v_2, \dots, v_{2+y_1+y_2+\dots+y_k})^T$  be the  $n$ -th column of the matrix

$$V(i1^{y_1} 2^{y_2} \dots k^{y_k} j).$$

Then  $v_1 = u_{i,n} = 2$  and  $v_{2+y_1+y_2+\dots+y_k} = u_{j,n'}$  or  $\overline{u_{j,n'}}$ . By Theorem 4.1(i), the first non-zero entry in the vector  $(v_2, \dots, v_{2+y_1+y_2+\dots+y_k})^T$  should be  $\bar{2}$ . Now let  $t \geq 1$ . Notice that the  $(n + tm_1)$ -th column of the matrix  $V(i1^t 1^{y_1} 2^{y_2} \dots k^{y_k} j)$  is of the form

$$\mathbf{v}' := (u_{i,n+tm_1}, \underbrace{0, \dots, 0}_t, v_2, \dots, v_{2+y_1+y_2+\dots+y_k})^T,$$

where the last  $(y_1 + y_2 + \dots + y_k + 1)$ -entries of  $\mathbf{v}'$ , i.e.  $v_2, \dots, v_{2+y_1+y_2+\dots+y_k}$ , coincide that of  $\mathbf{v}$ . Since  $a_{n+tm_1} = 0$ , by Theorem 4.1(ii), we have  $u_{i,n+tm_1} \in \{0, 2\}$ . If  $u_{i,n+tm_1} = 0$ , then the first non-zero entry in the vector  $\mathbf{v}'$  is  $\bar{2}$ , the same as that in the vector  $(v_2, \dots, v_{2+y_1+y_2+\dots+y_k})^T$ ; hence by Theorem 4.1(ii), we have  $a_{n+tm_1} = 2$ , leading to a contradiction. Hence we must have  $u_{i,n+tm_1} = 2$ , which completes the proof of the lemma.  $\square$

**Proposition 6.1.** *Assume that  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Then the sequence  $(U_\Phi(n))_{n \geq 1}$  is eventually periodic with period  $\ell$ .*

*Proof.* Since  $\ell = \gcd(m_j : 1 \leq j \leq k, d_j = 0)$ , it is sufficient to show that for each  $j$  with  $d_j = 0$  (which forces that  $s_j = 1$  by the definition of  $\mathcal{F}$ ),

$$(U_\Phi(n))_{n \geq 1} \text{ is eventually periodic with period } m_j.$$

Without loss of generality, we prove the above statement for  $j = 1$ . Suppose on the contrary that this statement is false for  $j = 1$ . Then there exist  $i \in \{1, \dots, k\}$  and  $1 \leq p_0 \leq m_1$  such that the sequence  $(u_{i,p_0+nm_1})_{n=1}^\infty$  is not eventually periodic with period 1. Consequently, there exist infinitely many  $n$  such that  $u_{i,p_0+nm_1} \neq 0$ . Hence by the pigeon hole principle, there exist two integers  $n_1 < n_2$  such that

$$u_{i,j+n_1m_1} \neq 0, u_{i,j+n_2m_1} \neq 0 \text{ and } n_2 \equiv n_1 \pmod{m_i}.$$

Notice that  $(p_0 + n_2m_1) - m_i - (p_0 + n_1m_1) = (n_2 - n_1)m_1$  is a multiple of  $m_i$ , and hence it belongs to  $\Lambda$ . By Lemma 6.1,  $(u_{i,p_0+n_2m_1+tm_1})_{t \geq 0}$  is eventually periodic with

period 1, which contradicts the fact that  $(u_{i,p_0+nm_1})_{n=1}^\infty$  is not eventually periodic with period 1.  $\square$

**Lemma 6.2.** *Assume that  $a+\mathbf{F}_\Phi \subseteq \mathbf{C}$ . Suppose that for some  $(i, n) \in \{1, \dots, k\} \times \mathbb{N}$ ,*

$$(6.5) \quad u_{i,n} \neq u_{i,n+\ell} = u_{i,n+t\ell}, \quad \forall t \geq 2.$$

*Then  $U_\Phi(n') = (0, \dots, 0)^T$  for  $n' \geq 1$  satisfying  $n - m_i - n' \in \Lambda$ .*

*Proof.* Let  $n' \in \mathbb{N}$  so that  $n - m_i - n' \in \Lambda$ . We need to prove that  $u_{j,n'} = 0$  for every  $1 \leq j \leq k$ .

Assume on the contrary that  $u_{j,n'} \neq 0$  for some  $1 \leq j \leq k$ . First we have  $u_{i,n} = 0$ ; otherwise, if  $u_{i,n} \neq 0$  then (6.4) fulfils and hence by Lemma 6.1,  $u_{i,n} = u_{i,n+tm_1}$  for each  $t \geq 1$ , which contradicts the assumption (6.5) since  $m_1$  is a multiple of  $\ell$ .

As  $u_{i,n} = 0$ , by (6.5),  $u_{i,n+\ell} = u_{i,n+t\ell} \neq 0$  for all  $t \geq 2$ . Without loss of generality assume that  $u_{i,n+\ell} = 2$ . Then  $u_{i,n+m_1} = u_{i,n+m_1+2\ell m_i} = 2$ , again using the fact that  $m_1$  is a multiple of  $\ell$ .

Let  $(t_1, \dots, t_{2\ell+1})^T$  be the  $(n+m_1+2\ell m_i)$ -th column of the matrix  $V(i^{2\ell+1})$ . Since  $u_{i,n+m_1} = u_{i,n+m_1+2\ell m_i} = 2$  and  $s_i^{2\ell} = 1$ , we have  $t_1 = t_{2\ell+1} = 2$ . By Theorem 4.1, the last non-zero entry in  $(t_1, \dots, t_{2\ell})^T$  takes the value  $\bar{2}$ .

Since  $n - m_i - n' \in \Lambda$ , there exist non-negative integers  $y_1, \dots, y_k$  such that  $n - m_i - n' = y_1 m_1 + y_2 m_2 + \dots + y_k m_k$ . Let  $\mathbf{w}$  denote the word  $1^{y_1} 2^{y_2} \dots k^{y_k}$ . Let

$$\mathbf{v} = (v_1, \dots, v_{y_1+\dots+y_k+1})^T$$

be the  $(n' + \sum_{i'=1}^k y_{i'} m_{i'})$ -th column of the matrix  $V(\mathbf{w}j)$ . Since  $u_{j,n'} \neq 0$ , we have  $v_{y_1+\dots+y_k+1} \neq 0$ .

Observe that the  $(n+m_1)$ -th column of the matrix  $V(i1\mathbf{w}j)$  is of the form

$$(2, 0, \tilde{v}_1, \dots, \tilde{v}_{y_1+\dots+y_k+1})^T,$$

where  $(\tilde{v}_1, \dots, \tilde{v}_{y_1+\dots+y_k+1})^T = \mathbf{v}$  if  $s_i > 0$ , and  $-\mathbf{v}$  if  $s_i < 0$ . By Theorem 4.1(i), the first non-zero entry in  $(\tilde{v}_1, \dots, \tilde{v}_{y_1+\dots+y_k+1})^T$  must be  $\bar{2}$ . Next consider the  $(n+m_1+2\ell m_i)$ -th column of the matrix  $V(i^{2\ell}1i\mathbf{w}j)$ , which is of the form

$$(6.6) \quad (t_1, \dots, t_{2\ell}, 0, 0, \tilde{v}_1, \dots, \tilde{v}_{y_1+\dots+y_k+1})^T.$$

Since the last non-zero entry in  $(t_1, \dots, t_{2\ell})^T$  is  $\bar{2}$  and the first non-zero entry in

$$(\tilde{v}_1, \dots, \tilde{v}_{y_1+\dots+y_k+1})^T$$

is  $\bar{2}$ , the vector in (6.6) contains two consecutive entries  $\bar{2}$  after deleting all zero entries, leading to a contradiction with Theorem 4.1(i). This completes the proof of the lemma.  $\square$

**Remark 6.1.** *The conclusion of Lemma 6.2 holds if we replace the assumption  $n - m_i - n' \in \Lambda$  by the following (stronger) condition:*

$$(6.7) \quad n - n' \geq m_i + Lr \quad \text{and} \quad n' \equiv n \pmod{r}.$$

Indeed, according to (6.3), the condition (6.7) implies that  $n - m_i - n' \in \Lambda$ .

**Lemma 6.3.** *Assume that  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}$ . Suppose that  $u_{i,n} \neq 0$  for some  $(i, n) \in \{1, \dots, k\} \times \mathbb{N}$ . Then  $(a_{\tilde{n}+t\ell})_{t \geq 0}$  is a constant sequence for each  $\tilde{n}$  satisfying that*

$$(6.8) \quad \tilde{n} - n \geq Lr \quad \text{and} \quad \tilde{n} \equiv n \pmod{r}.$$

*Proof.* Since  $\ell = \gcd(m_j : 1 \leq j \leq k, d_j = 0)$ , it is sufficient to show that for each  $j$  with  $d_j = 0$  (which forces that  $s_j = 1$  by the definition of  $\mathcal{F}$ ),  $(a_{\tilde{n}+t\ell})_{t \geq 0}$  is periodic with period  $m_j/\ell$  for each  $\tilde{n}$  satisfying (6.8). This is equivalent to prove that

$$(6.9) \quad (a_{\tilde{n}+tm_j})_{t \geq 0} \text{ is a constant sequence}$$

for each  $\tilde{n}$  satisfying (6.8). Without loss of generality, below we prove (6.9) for  $j = 1$ .

Fix  $\tilde{n}$  so that (6.8) fulfils. By (6.3), there exist non-negative integers  $y_1, \dots, y_k$  such that  $\tilde{n} - n = y_1 m_1 + \dots + y_k m_k$ . Write  $\mathbf{w} := 1^{y_1} 2^{y_2} \dots k^{y_k}$ . Let

$$\mathbf{v} = (v_1, \dots, v_{y_1 + \dots + y_k + 1})^T$$

be the  $\tilde{n}$ -th column of the matrix  $V(\mathbf{w}i)$ . Since  $u_{i,n} \neq 0$ , we have  $v_{y_1 + \dots + y_k + 1} \neq 0$ . Let  $v'$  be the first non-zero entry in the vector  $\mathbf{v}$ . By Theorem 4.1(ii),  $a_{\tilde{n}}$  is determined by  $v'$ ; that is,  $a_{\tilde{n}} = 0$  if  $v' = 2$ , and 2 otherwise.

Now let  $t \geq 1$ . Notice that the  $(\tilde{n} + tm_1)$ -th column of the matrix  $V(1^t \mathbf{w}i)$  is of the form

$$\mathbf{v}_t := (\underbrace{0, \dots, 0}_t, v_1, v_2, \dots, v_{y_1 + y_2 + \dots + y_k + 1})^T.$$

The first non-zero entry in  $\mathbf{v}_t$  is  $v'$ . Again, by Theorem 4.1(ii),  $a_{\tilde{n} + tm_1}$  is determined by  $v'$ . It follows that  $a_{\tilde{n} + tm_1} = a_{\tilde{n}}$ . This completes the proof of the lemma.  $\square$

6.2. **Case  $r = 1$ .** Set

$$(6.10) \quad n_0 = \inf \{n : U_{\Phi}(n) \neq \{0, \dots, 0\}^T\}.$$

**Theorem 6.1.** *Assume that  $r = 1$ . Then  $a + \mathbf{F}_{\Phi} \subseteq \mathbf{C}$  if and only if the following properties hold:*

- (1) *The sequence  $(U_{\Phi}(n))_{n \geq n_0 + L + D}$  is periodic with period  $\ell$ .*
- (2) *The sequence  $(a_n)_{n \geq n_0 + L}$  is periodic with period  $\ell$ .*
- (3) *Properties (i)-(ii) in Theorem 4.1 hold for those pairs  $(p, \mathbf{x}) \in \mathbb{N} \times \{1, \dots, k\}^*$  satisfying that*

$$n_0 \leq p < n_0 + L + D + 2\ell D \quad \text{and} \quad |\mathbf{x}| \leq p - n_0 + 1.$$

*Proof.* We first prove the ‘only if’ part of the theorem. Assume that  $a + \mathbf{F}_{\Phi} \subseteq \mathbf{C}$ . By Theorem 4.1, (3) holds.

To show (1), suppose on the contrary that (1) is false. By Proposition 6.1, the sequence of the column vectors of  $U_{\Phi}$  is eventually periodic with period  $\ell$ . Hence there exist  $i \in \{1, \dots, k\}$  and  $n \geq n_0 + L + D$  such that

$$u_{i,n} \neq u_{i,n+\ell} = u_{i,n+t\ell}, \quad \forall t \geq 2.$$

Then we have  $U_{\Phi}(n_0) = (0, \dots, 0)^T$  by Remark 6.1, because  $n \equiv n_0 \pmod{\ell}$  and  $n - n_0 \geq L + D \geq m_i + Lr$ . This contradicts (6.10), and completes the proof of (1).

Next we prove (2). By the definition of  $n_0$ , there exists  $i \in \{1, \dots, k\}$  so that  $u_{i,n_0} \neq 0$ . Since  $r = 1$ , by Lemma 6.3,  $(a_{\tilde{n}+t\ell})_{t \geq 0}$  is a constant sequence for each  $\tilde{n} \geq n_0 + L$ , from which (2) follows. This completes the proof of the ‘only if’ part of the theorem.

Now we turn to the proof of the ‘if’ part of the theorem. Suppose that  $a + \mathbf{F}_{\Phi}$  is not a subset of  $\mathbf{C}$ . By Theorem 4.1, at least one of the following two scenarios occurs:

- (a) There exist  $p \in \mathbb{N}$  and  $\mathbf{x} \in \{1, \dots, k\}^*$  such that the  $p$ -th column of  $V(\mathbf{x})$  is of the form  $(2, \underbrace{0, \dots, 0}_{|\mathbf{x}|-2}, 2)^T$  or  $(\bar{2}, \underbrace{0, \dots, 0}_{|\mathbf{x}|-2}, \bar{2})^T$ .
- (b) There exist  $p \in \mathbb{N}$  and  $\mathbf{x} \in \{1, \dots, k\}^*$  such that the  $p$ -th column of  $V(\mathbf{x})$  is of the form  $(\underbrace{0, \dots, 0}_{|\mathbf{x}|-1}, u_p)^T$  with  $u_p \neq 0$  and  $(a_p, u_p)$  is not plus-admissible.

First suppose that (a) occurs for some pair  $(p, \mathbf{x})$ . We may assume that  $p$  is the smallest in the sense that if (a) happens for another pair  $(p', \mathbf{x}')$ , then  $p \leq p'$ . and  $\mathbf{x} = x_1 \dots x_t \in \{1, \dots, k\}^*$ . Then

$$u_{x_1, p} \neq 0 \text{ and } u_{x_t, p-v} \neq 0$$

with  $v := m_{x_1} + \dots + m_{x_{t-1}}$ . The second condition and (6.10) imply that  $p - (m_{x_1} + \dots + m_{x_{t-1}}) \geq n_0$  and thus

$$|\mathbf{x}| = t \leq m_{x_1} + \dots + m_{x_{t-1}} + 1 \leq p - n_0 + 1.$$

By (3), we have

$$(6.11) \quad p \geq n_0 + L + D + 2\ell D.$$

Below we show that there exist  $p - 2\ell D \leq q < p$  with  $q \equiv p \pmod{\ell}$  and  $\mathbf{z} \in \{1, \dots, k\}^*$  so that (a) happens for the pair  $(q, \mathbf{z})$  instead of  $(p, \mathbf{x})$ , leading to a contradiction with the minimality of  $p$ .

We assume, without loss of generality, that the  $p$ -th column of  $V(\mathbf{x})$  is of the form  $(2, \underbrace{0, \dots, 0}_{|\mathbf{x}|-2}, 2)^T$ . If  $|\mathbf{x}| \leq 2\ell$ , then by (6.11),

$$p - \ell - (m_{x_1} + \dots + m_{x_{t-1}}) \geq n_0 + L + D;$$

hence by (1), the  $(p - \ell)$ -column of  $V(\mathbf{x})$  coincides with its  $p$ -th column. Thus (a) happens for the pair  $(p - \ell, \mathbf{x})$ . Next consider the case that  $|\mathbf{x}| \geq 2\ell + 1$ . For  $1 \leq i \leq 2\ell + 1$ , the pair  $(s_{x_1} \dots s_{x_i}, m_{x_1} + \dots + m_{x_i} \pmod{\ell})$  take values in  $\{1, -1\} \times \{0, 1, \dots, \ell - 1\}$ , a set with cardinality  $2\ell$ . By the pigeon hole principle, there exist  $1 \leq i < j \leq 2\ell + 1$  such that

$$s_{x_1} \dots s_{x_i} = s_{x_1} \dots s_{x_j} \text{ and } m_{x_1} + \dots + m_{x_i} \equiv m_{x_1} + \dots + m_{x_j} \pmod{\ell}.$$

Set  $\mathbf{z} = x_1 \dots x_i x_{j+1} \dots x_t$ , and  $q = p - (m_{x_{i+1}} + \dots + m_{x_j})$ . Then  $q \geq p - 2\ell D$  and  $q \equiv p \pmod{\ell}$ . Since

$$\begin{aligned} q - (m_{x_1} + \dots + m_{x_{i-1}}) &= p - (m_{x_1} + \dots + m_{x_j}) + m_{x_i} \\ &\geq p - 2\ell D \geq n_0 + D + \ell, \end{aligned}$$

by the  $\ell$ -periodicity of  $(U_{\mathbb{F}}(n))_{n \geq n_0 + D + \ell}$ , the first  $i$  entries in the  $q$ -th column of  $V(\mathbf{z})$  are the same as that in the  $p$ -th column of  $V(\mathbf{x})$ . In the mean time, since  $s_{x_1} \dots s_{x_i} = s_{x_1} \dots s_{x_j}$ , applying Lemma 4.1(iii) we see that the last  $(t - j)$  entries in the  $q$ -th column of  $V(\mathbf{z})$  coincide with that in the  $p$ -th column of  $V(\mathbf{x})$ . Hence

the  $q$ -th column of  $V(\mathbf{z})$  is of the form  $(2, \underbrace{0, \dots, 0}_{|\mathbf{z}|-2}, 2)^T$ . This proves the above claim, and hence derives a contradiction if (a) occurs.

Next suppose that (b) occurs for some pair  $(p, \mathbf{x})$ . Again we assume that  $p$  is the smallest. Following essentially the same argument as in the above paragraph, we can find  $p - 2\ell D \leq q < p$  with  $q \equiv p \pmod{\ell}$  and  $\mathbf{z} \in \{1, \dots, k\}^*$ , such that the  $q$ -th column of  $V(\mathbf{z})$  is of the form  $(\underbrace{0, \dots, 0}_{|\mathbf{z}|-1}, u_p)^T$ . Since  $q \geq p - 2\ell D \geq n_0 + L + D$ , by (2) we have  $a_q = a_p$ , and thus  $(a_q, u_p)$  is not plus admissible. Therefore (b) occurs for the pair  $(q, \mathbf{z})$ , contradicting the minimality of  $p$ . This completes the proof of Theorem 6.1.  $\square$

6.3. **Case  $r > 1$ .** Analogous to Theorem 6.1, we have

**Theorem 6.2.** *Assume that  $r > 1$ . For  $i = 0, 1, \dots, r - 1$ , set*

$$n_i = \inf \{n \in \mathbb{N} : U_{\Phi}(rn - i) \neq \{0, \dots, 0\}^T\},$$

*with convention  $\inf \emptyset = \infty$ . Then  $a + \mathbf{F}_{\Phi} \subseteq \mathbf{C}$  if and only if the following properties hold for each  $i \in \{0, 1, \dots, r - 1\}$  with  $n_i \neq \infty$ :*

- (1) *The sequence  $(U_{\Phi}(rn - i))_{n \geq n_i + L + D/r}$  is periodic with period  $\ell/r$ .*
- (2) *The sequence  $(a_{rn-i})_{n \geq n_i + L}$  is periodic with period  $\ell/r$ .*
- (3) *Properties (i)-(ii) in Theorem 4.1 hold for those pairs  $(rp - i, \mathbf{x}) \in \mathbb{N} \times \{1, \dots, k\}^*$  satisfying that*

$$n_i \leq p < n_i + L + D/r + 2\ell D/r \quad \text{and} \quad |\mathbf{x}| \leq p - n_i + 1.$$

The proof of Theorem 6.2 is essentially identical to that of Theorem 6.1 and so is omitted.

## 7. GENERALIZATIONS AND REMARKS

In this section we give some generalizations of our main results. For  $0 < \alpha < 1/2$ , write

$$\mathbf{C}_{\alpha} = \left\{ \sum_{n=1}^{\infty} \epsilon_n \alpha^n : \epsilon_n = 0 \text{ or } 2 \right\},$$

and call  $\mathbf{C}_\alpha$  the central  $\alpha$ -Cantor set<sup>1</sup>. The following addition and subtraction principles hold for  $0 < \alpha < 1/3$ .

**Lemma 7.1.** *Let  $0 < \alpha < 1/3$ . Let  $a = \sum_{n=1}^{\infty} a_n \alpha^n$  with  $a_n \in \{0, 2\}$  and  $b = \sum_{n=1}^{\infty} u_n \alpha^n$  with  $u_n \in \{0, 2, -2\}$ . Then  $a + b \in \mathbf{C}_\alpha$  if and only if  $a_n + b_n \in \{0, 2\}$  for all  $n \in \mathbb{N}$ . Similarly,  $a - b \in \mathbf{C}_\alpha$  if and only if  $a_n - b_n \in \{0, 2\}$  for all  $n \in \mathbb{N}$ .*

*Proof.* Without loss of generality we only prove that  $a + b \in \mathbf{C}_\alpha$  if and only if  $a_n + b_n \in \{0, 2\}$  for all  $n \in \mathbb{N}$ . The ‘if’ part is trivial. To show the ‘only if’ part, suppose that  $a + b \in \mathbf{C}_\alpha$ . Then there exists  $(c_n) \in \{0, 2\}^{\mathbb{N}}$  such that

$$\sum_{n=1}^{\infty} (a_n + b_n) \alpha^n = \sum_{n=1}^{\infty} c_n \alpha^n.$$

It is sufficient to show that  $a_n + b_n = c_n$  for all  $n \in \mathbb{N}$ . Suppose this is not true. Let  $n_0$  be the smallest positive integer so that  $a_{n_0} + b_{n_0} \neq c_{n_0}$ . Then we have

$$(7.1) \quad a_{n_0} + b_{n_0} - c_{n_0} = \sum_{m=1}^{\infty} (c_{n_0+m} - a_{n_0+m} - b_{n_0+m}) \alpha^m.$$

Notice that  $|a_{n_0} + b_{n_0} - c_{n_0}| \geq 2$ , whilst  $-4 \leq c_{n_0+m} - a_{n_0+m} - b_{n_0+m} \leq 4$  for each  $m \in \mathbb{N}$ . Hence by (7.1), we have  $2 \leq 4 \sum_{n=1}^{\infty} \alpha^n = 4\alpha/1 - \alpha$ , which contradicts the assumption that  $\alpha < 1/3$ .  $\square$

Due to the above lemma, slightly modified versions of our main results (Theorems 1.2, 1.3, 6.1 and 6.2) remain valid for  $\mathbf{C}_\alpha$  ( $0 < \alpha < 1/3$ ). The proofs are essentially identical to that for the case  $\alpha = 1/3$ . To be concise, below we only formulate the modified version of Theorem 1.2.

**Theorem 7.1.** *Let  $\alpha \in (0, 1/3)$  and  $m \in \mathbb{N}$ . Let  $a = \sum_{n=1}^{\infty} a_n \alpha^n$  with  $a_n \in \{0, 2\}$  and  $\Phi = \{\alpha^m x + d_i\}_{i=1}^k$  with  $d_i = \sum_{i=1}^k u_{i,n} \alpha^n$ , where  $u_{1,n} = 0$  and  $u_{i,n} \in \{0, 2, -2\}$  for all  $n \in \mathbb{N}$  and  $1 \leq i \leq k$ . Set*

$$U_{\Phi, \alpha} = (u_{i,n})_{1 \leq i \leq k, n \geq 1},$$

and let  $U_{\Phi, \alpha}(n)$  denote the  $n$ -th column of  $U_{\Phi, \alpha}$ . Then  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}_\alpha$  if and only if there exists a finite non-empty set  $\mathcal{M} \subset \mathbb{N}$  such that the following properties (i)-(iv) hold:

- (i) Any two numbers in  $\mathcal{M}$  are incongruent modulo  $m$ ;

---

<sup>1</sup>Notice that the diameter of  $\mathbf{C}_\alpha$  is  $2\alpha/(1 - \alpha)$ .

- (ii) For each  $n \in \mathcal{M}$ ,  $U_{\Phi, \alpha}(n)$  is either positive or negative;
- (iii) For each  $n \in \mathbb{N} \setminus \mathcal{M}$ ,  $U_{\Phi, \alpha}(n) = (0, 0, \dots, 0)^T$ ;
- (iv) For any  $n \in \mathcal{M}$ , and any integer  $t \geq 0$ ,

$$a_{n+mt} = \begin{cases} 0, & \text{if } U_{\Phi, \alpha}(n) \text{ is positive,} \\ 2, & \text{if } U_{\Phi, \alpha}(n) \text{ is negative.} \end{cases}$$

Now it arises a natural question whether Theorem 1.1 can be extended to other values  $\alpha \in (0, 1/3)$ . After the completion of the first version of this paper, the following extension of Theorem 1.1 was given in [3] in a slightly variant version.

**Theorem 7.2.** [3, Theorem 1.4] *Let  $0 < \alpha < 1/2$ . Assume that  $\mathbf{F} \subseteq \mathbf{C}_\alpha$  is a non-trivial self-similar set, generated by a linear IFS  $\Phi = \{\phi_i\}_{i=1}^k$  on  $\mathbb{R}$ .*

- (1) *If  $\alpha \leq 1/4$ , then for each  $1 \leq i \leq k$ ,  $\phi_i$  has contraction ratio  $\pm\alpha^{-m_i}$ , where  $m_i \in \mathbb{N}$ .*
- (2) *If  $\alpha < \sqrt{2} - 1$  or  $1/\alpha$  is a Pisot number, then  $\phi_i$  has contraction ratio  $\pm\alpha^{t_i}$ , where  $t_i \in \mathbb{Q}_+$ .*

Recall that  $\lambda > 1$  is called a *Pisot number* if it is an algebraic integer whose algebraic conjugates are all less than 1 in modulus.

Furthermore it was observed in [3, Lemma 4.1] that there exist countably many points  $\alpha$  in  $((3 - \sqrt{5})/2, 1/2)$ , namely the positive roots of  $x - (x + x^2 + \dots + x^k)^2$  ( $k = 2, 3, \dots$ ), such that  $\mathbf{C}_\alpha$  contains a self-similar subset  $E$  which admits a contraction ratio  $\alpha^{q/2}$  for an odd integer  $q$  in one of the maps in a generating IFS of  $E$ . Nevertheless,  $(3 - \sqrt{5})/2 \approx 0.382 > 1/3$ .

For  $p \geq 2$ , let  $\Lambda_p$  denote the set of  $\alpha \in (0, 1/2)$  so that  $\mathbf{C}_\alpha$  contains a self-similar subset  $E$  which admits a contraction ratio  $\alpha^{q/p}$ , with  $q$  coprime to  $p$ , in one of the maps in a generating IFS of  $E$ . By Theorems 7.1 and 7.2, we can formulate a criterion for  $\alpha \in (1/4, 1/3) \cap \Lambda_p$ ; and in particular we can characterize the elements in the set  $(1/4, 1/3) \cap \Lambda_2$ .

**Theorem 7.3.** *Let  $p \in \mathbb{N}$  with  $p \geq 2$ . Then  $\alpha \in (1/4, 1/3) \cap \Lambda_p$  if and only if there exists  $b > 0$  such that  $b\alpha^{n/p}$  ( $n = 0, \dots, p - 1$ ) are of the forms*

$$(7.2) \quad b\alpha^{n/p} = \sum_{k=1}^{t_n} \epsilon_{k,n} \alpha^k$$

simultaneously, where  $t_n \in \mathbb{N}$  and  $\epsilon_{k,n} \in \{0, 2, -2\}$  for  $1 \leq n \leq p-1$  and  $1 \leq k \leq t_n$ . In particular,

$$(1/4, 1/3) \cap \Lambda_2 := \{0 < \alpha < 1/3 : \exists \text{ polynomials } A(x) \text{ and } B(x) \\ \text{of coefficients } 0, \pm 2 \text{ such that } \alpha^{1/2} = A(\alpha)/B(\alpha)\}.$$

*Proof.* We first prove the necessity part of the theorem. Suppose that  $\alpha \in (0, 1/3) \cap \Lambda_p$ . It is not hard to show that there exist  $a, c > 0$ ,  $m' \in \mathbb{N}$  such that  $\mathbf{C}_\alpha \supset a + \mathbf{F}_\Psi$ , where  $\Psi = \{\alpha^{m'+1/p}x, \alpha^{m'+1/p}x + c\}$ . Notice that

$$\Phi := \Psi^p = \left\{ \alpha^{pm'+1}x + c \sum_{n=0}^{p-1} \epsilon_n \alpha^{m'n+n/p} : \epsilon_n \in \{0, 1\} \text{ for } 0 \leq n \leq p-1 \right\},$$

and  $\mathbf{F}_\Phi = \mathbf{F}_\Psi$ . We have

$$\mathbf{C}_\alpha \supset a + \mathbf{F}_\Phi.$$

Set  $m = pm'+1$ . By Theorem (7.1), for each  $(\epsilon_0, \dots, \epsilon_{p-1}) \in \{0, 1\}^p$ ,  $c \sum_{n=0}^{p-1} \epsilon_n \alpha^{m'n+n/p}$  can be expanded into a polynomials of  $\alpha$  with coefficients in  $\{0, 2, -2\}$ . In particular, for each  $0 \leq n \leq p-1$ ,  $c\alpha^{m'n+n/p}$  can be expanded into a polynomials of  $\alpha$  with coefficients in  $\{0, 2, -2\}$ , which implies that (7.2) holds for  $b = c\alpha^{m'/p}$ .

Next we prove the sufficiency part. Suppose that (7.2) holds. Take an integer  $m'$  so that

$$(7.3) \quad m' \geq \max\{t_n : n = 0, \dots, p-1\}.$$

Let

$$\Phi = \left\{ \alpha^{pm'+1}x + b \sum_{n=0}^{p-1} \epsilon_n \alpha^{m'n+n/p} : \epsilon_n \in \{0, 1\} \text{ for } 0 \leq n \leq p-1 \right\}.$$

Denote

$$G := \left\{ b \sum_{n=0}^{p-1} \epsilon_n \alpha^{m'n+n/p} : \epsilon_n \in \{0, 1\} \text{ for } 0 \leq n \leq p-1 \right\}.$$

The assumption (7.2) and the condition (7.3) ensure that each  $g \in G$  has an expansion  $\sum_{k=0}^{m'/p} \epsilon_{k,g} \alpha^k$ , with  $\epsilon_{k,g} \in \{0, 2, -2\}$ ; and moreover, for each  $0 \leq k \leq m'/p$ , either  $\epsilon_{k,g} \in \{0, 2\}$  for all  $g \in G$ , or  $\epsilon_{k,g} \in \{0, -2\}$  for all  $g \in G$ . Rewrite  $\Phi$  as  $\{\alpha^m x + b_i\}_{1 \leq i \leq 2^p}$ , with  $m = pm' + 1$  and  $b_1 = 0$ . Then the matrix  $U_{\Phi, \alpha}$  satisfies the conditions (i)-(iii) of Theorem (7.2), with  $\mathcal{M} \subseteq \{1, \dots, m\}$ . Choose  $(a_n) \in \{0, 2\}^{\mathbb{N}}$  such that (iv) of Theorem 7.1 holds. Then by Theorem (7.2),  $a + \mathbf{F}_\Phi \subseteq \mathbf{C}_\alpha$ . This completes the proof of the sufficiency part, since  $\mathbf{F}_\Phi = \mathbf{F}_\Psi$  with  $\Psi = \{\alpha^{m'+1/p}x, \alpha^{m'+1/p}x + b\}$ .  $\square$

**Remark 7.1.** (i)  $1/4$  is an accumulation point of the set  $(1/4, 1/3) \cap \Lambda_2$ , since this set contains the positive roots of  $x^{1/2} = \frac{1-x-x^2-\dots-x^n}{1+x+x^2+\dots+x^n}$  ( $n \geq 2$ ).

(ii) By Theorem 7.3, for each  $p \geq 2$ ,  $(1/4, 1/3) \cap \Lambda_p$  does not contain any transcendental numbers. We do not know whether  $(1/4, 1/3) \cap \Lambda_p \neq \emptyset$  for some  $p \geq 3$ .

**Remark 7.2.** As our main results (Theorems 1.2, 1.3, 6.1 and 6.2) can be extended to  $\mathbf{C}_\alpha$  for  $0 < \alpha < 1/3$ , by Theorem 7.3 and the definition of  $\Lambda_p$ , we can characterize all the self-similar subsets of  $\mathbf{C}_\alpha$ , in a way similar to that for  $\mathbf{C}$ , for all those  $\alpha$  in the following set

$$(0, 1/4] \cup \left[ (1/4, 1/3) \setminus \left( \bigcup_{p \geq 2} \Lambda_p \right) \right].$$

## APPENDIX A

In this part we give some known or simple results that are needed in the previous sections. We begin with the following.

**Lemma A.1.** [1, Theorem 1] *Let  $p_1, \dots, p_n \in \mathbb{N}$  and  $h = \gcd(p_1, \dots, p_n)$ . Then for any integer  $y \geq \left( \min_{1 \leq i \leq n} \frac{p_i}{h} - 1 \right) \left( \max_{1 \leq i \leq n} \frac{p_i}{h} - 1 \right)$ , there exist  $y_1, \dots, y_n \in \mathbb{N} \cup \{0\}$  such that*

$$yh = \sum_{i=1}^n y_i p_i.$$

The next result is elementary.

**Lemma A.2.** *Suppose  $\mathbf{E} \subseteq [0, 1]$  with  $0 \in \mathbf{E}$  is a non-trivial self-similar set generated by an IFS  $\Phi = \{\phi_i\}_{i=1}^k$  of the form*

$$\phi_i(x) = s_i 3^{-m_i} x + d_i, \quad i = 1, \dots, k,$$

*where  $s_i = \pm 1, m_i \in \mathbb{N}$  and  $d_i \in \mathbb{R}$ . Then  $0 \leq d_i \leq 1$  for all  $1 \leq i \leq k$ , and  $d_i > 0$  whenever  $s_i = -1$ . Moreover, there exists at least one  $1 \leq i \leq k$  so that  $d_i > 0$ .*

*Proof.* Since  $0 \in \mathbf{E}$ , we have  $d_i = \phi_i(0) \in \mathbf{E} \subseteq [0, 1]$  for each  $1 \leq i \leq k$ . Now suppose  $s_i = -1$  for some  $i$ . Then  $\mathbf{E} \supset \phi_i(\mathbf{E}) = d_i - 3^{-m_i} \mathbf{E}$ . Since  $\mathbf{E} \subseteq [0, 1]$ , we must have  $d_i > 0$ , for otherwise,  $\phi_i(\mathbf{E}) \subseteq (-\infty, 0]$ , which is impossible because  $\phi_i(\mathbf{E})$  contains at least two points and  $\phi_i(\mathbf{E}) \subset \mathbf{E} \subseteq [0, 1]$ .

Next we show that there exists at least one  $i \in \{1, \dots, k\}$  so that  $d_i > 0$ . If this is not true, then by the above argument,  $s_i = 1$  and  $d_i = 0$  for all  $i$ , which implies that  $\mathbf{E} = \{0\}$ , leading to a contradiction.  $\square$

Now we are ready to prove the following result (Lemma 1.1 in Section 1).

**Lemma A.3.** *Let  $\mathbf{F} \subseteq \mathbf{C}$  be a non-trivial self-similar set. Write  $\mathbf{F}' = \{1 - x : x \in \mathbf{F}\}$ . Then either  $\mathbf{F}$  or  $\mathbf{F}'$  can be written as  $a + \mathbf{E}$ , where  $a \in \mathbf{C}$  and  $\mathbf{E} = \mathbf{F}_\Phi$  for some  $\Phi \in \mathcal{F}$ , where  $\mathbf{F}_\Phi$  denotes the attractor of  $\Phi$ .*

*Proof.* Let  $\{S_i\}_{i=1}^k$  be a linear generating IFS of  $\mathbf{F}$ . By Theorem 1.1, this IFS must be of the form

$$S_i(x) = s_i 3^{-m_i} x + d_i, \quad i = 1, \dots, k,$$

where  $s_i = \pm 1$ ,  $m_i \in \mathbb{N}$  and  $d_i \in \mathbb{R}$ . Let  $[a, b]$  be the convex hull of  $\mathbf{F}$ . Then  $\bigcup_{i=1}^k S_i([a, b]) \subseteq [a, b]$ , and  $\{a, b\} \subset \bigcup_{i=1}^k S_i([a, b])$ . Without loss of generality, we assume that  $a \in S_1([a, b])$  and  $b \in S_k([a, b])$ . There are 3 possibilities: (1)  $s_1 = 1$ ; (2)  $s_k = 1$ ; (3)  $s_1 = s_k = -1$ .

First assume that (1) occurs. Then  $S_1(a) = a$ . Let  $\mathbf{E} = \mathbf{F} - a := \{x - a : x \in \mathbf{F}\}$ . Then  $\mathbf{E} \subseteq [0, 1]$  and

$$\mathbf{E} = \mathbf{F} - a = \bigcup_{i=1}^k (S_i(\mathbf{F}) - a) = \bigcup_{i=1}^k (S_i(\mathbf{F} - a) + S_i(a) - a) = \bigcup_{i=1}^k \phi_i(\mathbf{E}),$$

where the maps  $\phi_i$  are defined by  $\phi_i(x) = S_i(x) + S_i(a) - a$ . Note that  $\phi_1(x) = 3^{-m_1} x$ . This together with Lemma A.2 yields  $\Phi := \{\phi_i\}_{i=1}^k \in \mathcal{F}$ . Now  $\mathbf{F} = a + \mathbf{F}_\Phi$ .

Next assume that (2) occurs. Then  $S_k(b) = b$ . Let  $\mathbf{E} = b - \mathbf{F}$ . Similarly, one has  $\mathbf{E} \subseteq [0, 1]$  and

$$\mathbf{E} = \bigcup_{i=1}^k \phi_i(\mathbf{E}),$$

where  $\phi_{k+1-i}(x) := S_i(x) + b - S_i(b)$  for  $i = 1, \dots, k$ . Again, we have  $\phi_1(x) = 3^{-m_k} x$ . Thus by Lemma A.2 we have  $\Phi := \{\phi_i\}_{i=1}^k \in \mathcal{F}$ . Now  $\mathbf{F}' = 1 - \mathbf{F} = (1 - b) + \mathbf{F}_\Phi$ .

In the end, assume that (3) occurs. Then  $a = S_1(b)$  and  $b = S_k(a)$ . Hence  $S_1 \circ S_k(a) = a$ . Let  $\mathbf{E} = \mathbf{F} - a$ . Then  $\mathbf{E} = \bigcup_{i=1}^k \bigcup_{j=1}^k \widetilde{\phi}_{i,j}(\mathbf{E})$ , where  $\widetilde{\phi}_{i,j}(x) = S_i \circ S_j(x) + S_i \circ S_j(a) - a$ . Recoding the IFS  $\{\widetilde{\phi}_{i,j} : 1 \leq i, j \leq k\}$  to  $\{\phi_i\}_{i=1}^{k^2}$  so that  $\phi_1 = \widetilde{\phi}_{1,k}$ , then  $\Phi := \{\phi_i\}_{i=1}^{k^2} \in \mathcal{F}$  and  $\mathbf{F} = a + \mathbf{F}_\Phi$ .  $\square$

**Acknowledgements.** The authors would like to thank Ranran Huang and Ying Xiong for helpful comments. The first author was partially supported by the RGC grant in CUHK. The second author was partially supported by NNSF (11225105, 11371339). The third author was supported in part by NSF Grant DMS-1043034 and DMS-0936830.

#### REFERENCES

- [1] A. Brauer, On a problem of partitions. *Amer. J. Math.* **64** (1942), 299–312.
- [2] M. Csörnyei, Open questions, *Period. Math. Hungar.* **37** (1998), 227–237.
- [3] D. J. Feng, W. Huang and H. Rao, Affine embeddings and intersections of Cantor sets, *J. Math. Pures Appl.* (2014). <http://dx.doi.org/10.1016/j.matpur.2014.03.003>
- [4] D. J. Feng and Y. Wang, On the structures of generating iterated function systems of Cantor sets. *Adv. Math.* **222** (2009), 1964–1981.
- [5] J. E. Hutchinson, Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), 713–747.
- [6] R. Salem, *Algebraic numbers and Fourier analysis*. D. C. Heath and Co., Boston, Mass. 1963
- [7] R. Salem and A. Zygmund, Sur un théorème de Piętetcki-Shapiro. *C. R. Acad. Sci. Paris* **240** (1955), 2040–2042.
- [8] A. Zygmund, *Trigonometric series*. Vol. I, II. Third edition. Cambridge University Press, Cambridge, 2002.

DEPARTMENT OF MATHEMATICS, THE CHINESE UNIVERSITY OF HONG KONG, SHATIN, HONG KONG

*E-mail address:* [djfeng@math.cuhk.edu.hk](mailto:djfeng@math.cuhk.edu.hk)

DEPARTMENT OF MATHEMATICS, CENTRAL CHINA NORMAL UNIVERSITY, WUHAN 430070, HUBEI, PEOPLES'S REPUBLIC OF CHINA

*E-mail address:* [hrao@mail.ccnu.edu.cn](mailto:hrao@mail.ccnu.edu.cn)

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48824, UNITED STATES

*E-mail address:* [ywang@math.msu.edu](mailto:ywang@math.msu.edu)