THE DIFFERENCE SET OF TWO CANTOR SETS REVISITED

by Chao Du
Master Thesis in Mathematics

Delft Institute of Applied Mathematics,
Delft University of Technology, the Netherlands,
July 2003

Acknowledgement

I am very grateful to my supervisors. I would like to thank my adviser Prof. F.M. Dekking for helpful suggestions and inspiring contribution to my research. I thank Prof. K. Simon and R.J. Fokkink for the stimulating discussions. I would send my gratitude to Prof. Roger Cooke and Dorota Kurowicka for their help in my master's study. I also thank all my colleges and friends during my study at TU Delft. Finally, I thank my family and my girlfriend for their great support.

1. Introduction

In this work we will discuss the thesis of Per Larsson (with advisor Lennart Carleson, 1991). The results in this very interesting thesis were never completely published in journals. There is the note for the *C.R. Acad. Sci. Paris Sér. IMath.*310(1990), but the reviewer of this note comments that "The main argument of the proof comes from branching process theory, but many details are not clear."

Per Larsson defines a family of two-parameter random Cantor sets where the parameters a and b are real numbers that have to satisfy

$$3a + 2b < 1 \quad \text{and} \quad 4a > 1. \tag{1.1}$$

For reasons that will become clear later, we impose

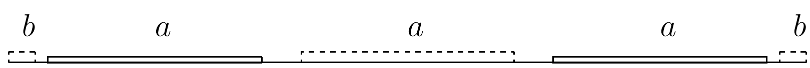
$$(1 - 3a - 2b)/4a > 2b(1 - a) \quad \text{and} \quad 1/4 < a < 1/3. \tag{1.2}$$

It's easy to check that condition (1.2) implies (1.1).

Given a and b the random Cantor set is defined as follows. Start with the unit interval. First remove the a-part (i.e., an interval of length a) centered around the midpoint of the unit interval and the b-parts at the two ends of the unit interval. Then choose two intervals E^1 and E^2 of length a in the remaining two intervals. This is possible because of the condition $3a + 2b < 1$. The location of the intervals E^1 and E^2 is random, distributed as two independent random variables $b + U^1$ and $1/2+a/2+U^2$ where $U^1 \stackrel{d}{=} U^2 \stackrel{d}{=} Uniform([0, t])$ with $t = 1/2-3a/2-b$. Repeat the procedure independently with the intervals E^1 and E^2 on the scale where the length of E^1 and E^2 is 1. Let C^2 be the union of the four intervals we obtain. Continue in this way and define C^3, C^4, Finally define the Cantor set C as the intersection of all C^n. See Figure 1.

The difference set of two sets C_1 and C_2 is defined as the set $C_1-C_2 = \{z = x-y \ : \ x \in C_1, \ y \in C_2\}$. Given two independent random Cantor sets defined as above with the same a and b, we will show that if the sum of the Hausdorff dimensions of the sets exceeds 1, then the difference set of these two sets will almost surely contain an interval.

Figure 1



Theorem 1.1. *Let C_1, C_2 be two independent random Cantor sets defined as above with the same a and b, then $C_1 - C_2$ almost surely contains an interval.*

The theorem will be proved by using a multi-type branching process, actually a branching process with an uncountable number of types. In Larsson's work, first, he proves that the point 0 is covered by a positive number of intervals with positive probability by using a branching process. Then these intervals will be used to start new branching processes giving intervals which cover a set of dense points close to 0. Since $C_1 - C_2$ is a close set, the fact that "a dense subset of an interval is contained in $C_1 - C_2$" gives that "$C_1 - C_2$ contains an interval". Larsson's proof of the validity of Condition (2.1) is missing. What is new in the present work is that we supply a proof of this condition (and our analysis will yield that Larsson's conditions are too weak - c.f. (1.1) and (1.2)).

2. Multi-type branching processes revisited

A branching process is a process where individuals give rise to new individuals according to a probabilistic law. Here we consider discrete time branching processes. At times $n \in \mathbb{N}$ an individual is replaced by a random number of new individuals. A multi-type branching process is a branching process where to each individual there is assigned a number: the type. We deal with *continuous* multi-type branching process where the type space is continuous. Following Harris's definition, we let Ω be the type space and let $A \subset \Omega$. Define $Z_n(x, A)$ as the number of individuals with type in the set A, in the n-th generation of a branching process starting from an individual with type x. Let

$$M_n(x, A) = E(Z_n(x, A)).$$

We say that M_n has a density if there exists a non-negative function $m_n(x, y)$ such that for each $x \in \Omega$ and a Borel set $A \subset \Omega$,

$$M_n(x, A) = \int_A m_n(x, y)dy.$$

If M_1 has a density m_1, and M_n has a density m_n, then M_{n+1} likewise has a density given by

$$m_{n+1}(x, y) = \int_\Omega m_n(x, z)m_1(z, y)dz.$$

To determine the convergence properties of the branching process we need to find the eigenvalues λ of the operator induced by m_1 defined by

$$\lambda\mu(x) = \int m_1(x,y)\mu(y)dy \text{ and } \lambda\nu(y) = \int \nu(x)m_1(x,y)dx,$$

where μ and ν are called the right and left eigenfunctions. The following theorem from [2] gives that under certain conditions there exists a largest eigenvalue which is positive.

Theorem 2.1. *Assume that $M_1(x,\Omega)$ is a bounded function of x, M_1 has a density m_1 and there is an integer n_0 and constants r_1 and r_2 such that for all $x,\ y \in \Omega$,*

$$0 < r_1 \le m_{n_0}(x,y) \le r_2 < \infty. \tag{2.1}$$

Then m_1 has a positive eigenvalue ρ, larger in magnitude than any other eigenvalue, corresponding to right and left eigenfunctions μ and ν that are bounded and uniformly positive. Moreover, μ and ν are the only nontrivial non negative right and left eigenfunctions. Furthermore, if we normalize μ and ν so that $\int \mu(x)\nu(x)dx = 1$ then for some $0 < \Delta < 1$,

$$m_n(x,y) = \rho^n\mu(x)\nu(y)[1 + O(\Delta^n)],\ n \to \infty,$$

where the bound Δ can be taken independently of x and y.

If we treat the functions $m_n(x,y)$ as transition functions from type x to type y, the condition (2.1) means that there exists an n_0 such that a branching process starting from an individual of type x gives rise to an individual of type y in the n_0-th generation, for any $x, y \in \Omega$, with uniformly positive "density".

Theorem 2.2. *([2]) Suppose the conditions of Theorem 2.1 are satisfied, and the largest eigenvalue ρ is larger than 1, and $E(Z_1(x,\Omega)^2)$ is a bounded function of x. For each $x \in \Omega$ and each subset A of Ω let $W_n(A) = Z_n(A)/\rho^n$, $n = 1, 2, \cdots$. Then we have*

$$P_x\left(\lim_{n\to\infty} W_n(A) = W(A)\right) = 1,$$

where $W(A)$ is a random variable such that $W(A)$ has positive probability of being positive for each $x \in \Omega$, $A \subset \Omega$ such that A has positive measure.

3. Application of multi-type branching process to random Cantor sets

First, we shall use a branching process to show that the point 0 is included in $C_1^n - C_2^n$ and has exponentially many coverings with positive probability. Note that the point 0 is in this difference set if and only if $C_1^n \bigcap C_2^n$ is not empty. The set $C_1^n \bigcap C_2^n$ consists of a number of intervals $I^k = I_1^k \bigcap I_2^k$, where $I_1^k \subset C_1^n$ and $I_2^k \subset C_2^n$. The length of I_1^k is a number between 0 and a^n. If we re-scale these lengths by dividing by a^n we get a number l^k such that $0 < l^k \leq 1$. We call this l^k the type of the individual I^k. Therefore the I^k's can be considered as individuals in the n-th generation of a continuous type multi-type branching process with type space $\Omega = (0, 1]$. The distributions of the number of offsprings of an individual I^k and the type of the offsprings depend only on the type of I^k. See Figure 2.

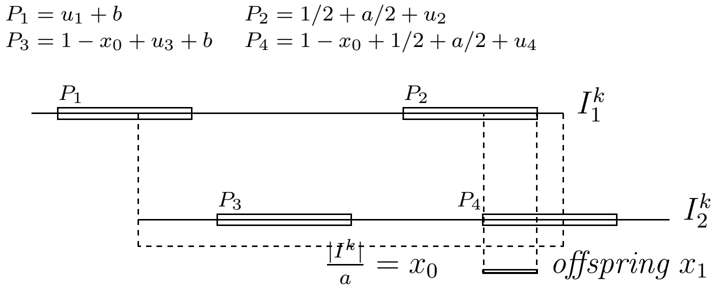
By the above definition of a random Cantor set, given an individual with type x_0, we assume that the intervals are located with respect to each other as in Figure 2. An individual can have 0, 1 or 2 children. We can calculate the types of its offspring as follows: if there is one child it can be of type X_1^1, or X_1^2, or X_1^3; if there are two children these will be of type X_1^2 and X_1^3, where

$$
\begin{aligned}
X_1^1 &= (1 - |1/2 - a/2 + b - x_0 - Y_1|/a)^+, \\
X_1^2 &= (1 - |1 - x_0 - Y_2|/a)^+, \\
X_1^3 &= (1 - |1 - x_0 - Y_3|/a)^+,
\end{aligned}
$$

and

$$
\begin{aligned}
U_i &\stackrel{d}{=} Uniform([0, t]),\ i = 1, 2, 3, 4, \\
Y_1 &= U_2 - U_3, \\
Y_2 &= U_1 - U_3, \\
Y_3 &= U_2 - U_4.
\end{aligned}
$$

Figure 2



We get that

$$\begin{aligned} M_1(x,A) &= E(Z_1(x,A)) \\ &= E(\mathbb{I}_{X_1^1\in A}+\mathbb{I}_{X_1^2\in A}+\mathbb{I}_{X_1^3\in A}) \\ &= P(X_1^1\in A)+P(X_1^2\in A)+P(X_1^3\in A) \\ &= \int_A [f_{X_1^1}(y|x)+f_{X_1^2}(y|x)+f_{X_1^3}(y|x)]dy. \end{aligned}$$

Hence $f_{X_1^1}(y|x)$, $f_{X_1^2}(y|x)$ and $f_{X_1^3}(y|x)$ are sub-density functions for $X_1^1>0$, $X_1^2>0$ and $X_1^3>0$. So M_1 has a density

$$m_1(x,y)=f_{X_1^1}(y|x)+f_{X_1^2}(y|x)+f_{X_1^3}(y|x).$$

We denote the support of the function $m_n(x,y)$ on $\Omega\times\Omega$ by

$$\Omega_n=\overline{\{(x,y):m_n(x,y)>0\}}.$$

We also define the *cross section* $E_n(x)$ at x in Ω:

$$E_n(x)=\{y:(x,y)\in\Omega_n\}.$$

Correspondingly, we define

$$f_n(x)=\inf E_n(x) \text{ and } g_n(x)=\sup E_n(x).$$

It follows from the earlier remarks that $E_1(x)$ is a connected interval (i.e., $E_1(x)=[f_1(x),g_1(x)]$), and that $f_1(x)$ and $g_1(x)$ are continuous functions. See Figure 3 [1].

We also give the explicit function recipe of $f_1(x)$ and $g_1(x)$ here:

$$f_1(x)=(x-x_i)k_i+y_i \ \ in \ \ [x_i,x_{i+1}), \ \ i=1,2,3,4,5,$$

where

$$\begin{aligned} &(x_1,y_1,k_1)=(0,0,0), \\ &(x_2,y_2,k_2)=(1-3a,0,1/a), \\ &(x_3,y_3,k_3)=((1-a+2b)/2,(5a+2b-1)/2a,-1/a), \\ &(x_4,y_4,k_4)=(2a+2b,0,0), \\ &(x_5,y_5,k_5)=((1-5a-2b)/2,0,1/a), \\ &x_6=1, \end{aligned}$$

and

[1] In this figure and all the following figures we have taken $a=0.3$ and $b=0.01$

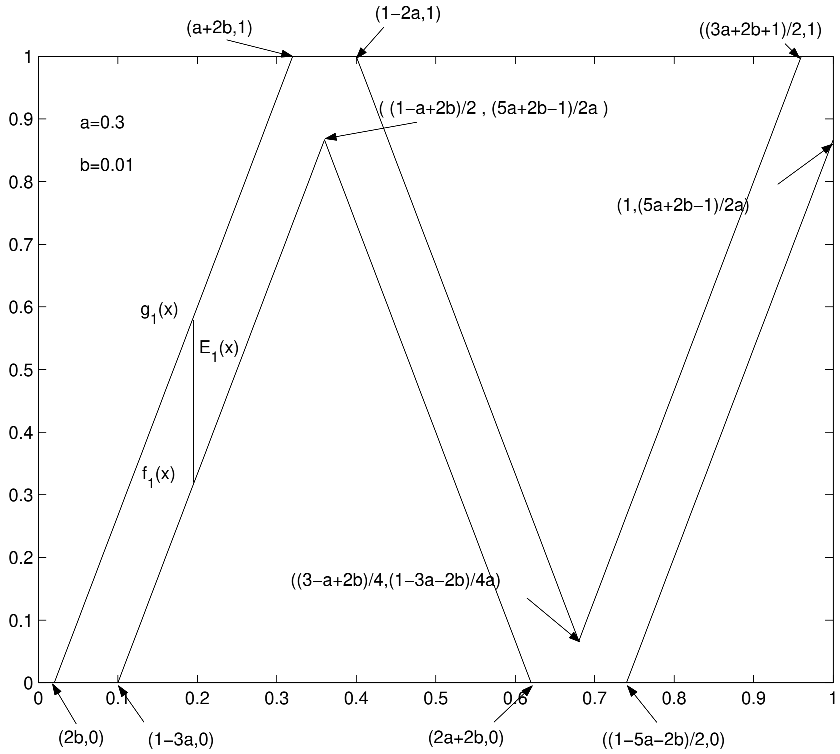


FIGURE 3

$$g_1(x) = (x - u_i)k_i + v_i \ \ in \ \ [u_i, u_{i+1}), \ \ i = 1, 2, 3, 4, 5, 6,$$

where

$$\begin{aligned}
&(u_1, v_1, k_1) = (0, 0, 0),\\
&(u_2, v_2, k_2) = (2b, 0, 1/a),\\
&(u_3, v_3, k_3) = (a + 2b, 1, 0),\\
&(u_4, v_4, k_4) = (1 - 2a, 1, -1/a),\\
&(u_5, v_5, k_5) = ((3 - a + 2b)/4, (1 - 3a - 2b)/4a, 1/a),\\
&(u_6, v_6, k_6) = ((3a + 2b + 1)/2, 1, 0),\\
&u_7 = 1.
\end{aligned}$$

We call the points $x_3 = (1 - a + 2b)/2$ and $u_5 = (3 - a + 2b)/4$ fold points (these are the x-coordinates of the positions where Ω_1 'changes direction', see Figure 3). Moreover, one can easily check that:

$$m_1(x, y) > 0 \text{ when } y \in (f_1(x), g_1(x)) \tag{3.1}$$

and

$$m_1(x,1) > 0 \text{ when } x \in ((3a+2b+1)/2, 1]. \tag{3.2}$$

In the following, we will prove a weaker property than the lower bound of the equation (2.1) in Theorem 2.1: We can replace Ω by a smaller type space, such that there exists an integer n_0 such that for the new type space Ω, $\Omega \times \Omega \subseteq \Omega_{n_0}$. It's easy to see that the support Ω_1 does not cover the whole square $\Omega \times \Omega$ (c.f. Figure 3). We will show that the support will "grow" to the whole square in a finite number of steps by the following two lemmas. The reader may also see Figure 4 to get an intuitive idea of the "growing" of the support.

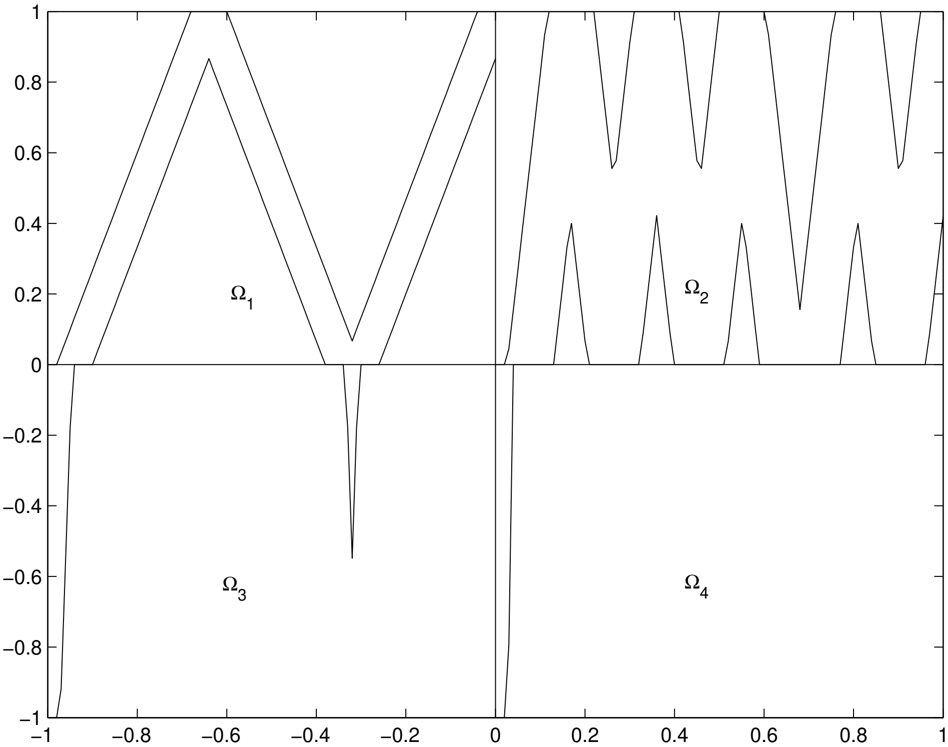


FIGURE 4

Lemma 3.1. *$E_n(x)$ is a connected interval.*

Proof. The proof is by induction. We already noted that $E_1(x)$ is a connected interval,

$$E_1(x) = [f_1(x), g_1(x)].$$

Assume that $E_n(x)$ is a connected interval. Recall that

$$m_{n+1}(x,y) = \int_\Omega m_n(x,z)m_1(z,y)dz,$$

so by looking at Figure 5, we can see that

$$E_{n+1}(x) = \bigcup_{y\in E_n(x)} E_1(y) = \bigcup_{y\in E_n(x)} [f_1(y), g_1(y)]$$

is a connected interval since $f_1(y)$ and $g_1(y)$ are continuous functions of y. □

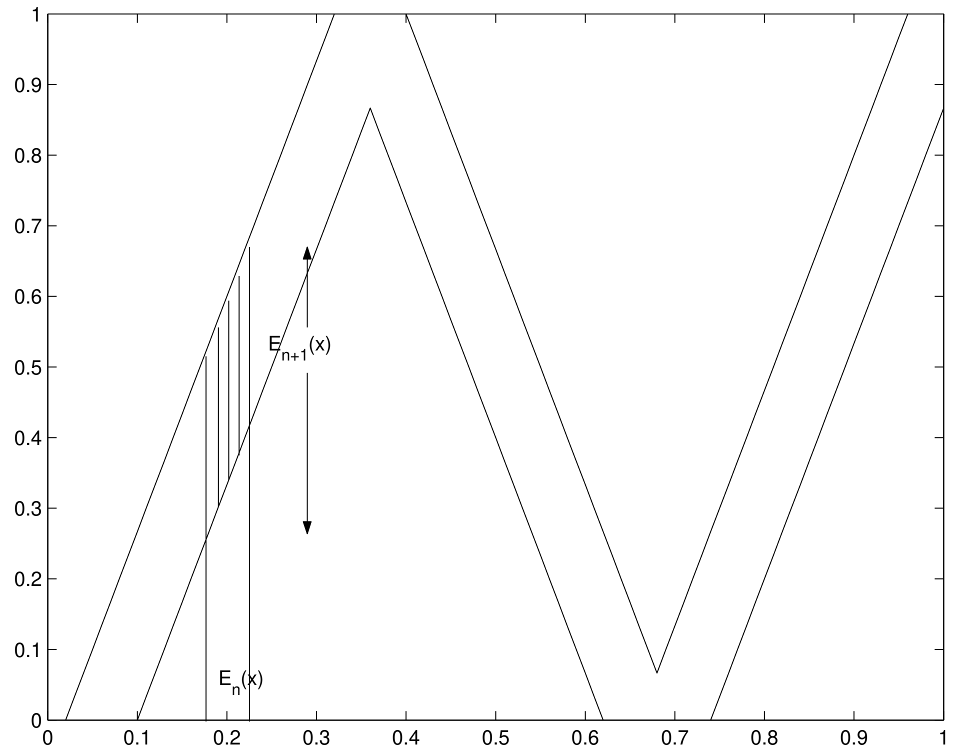


FIGURE 5

We note that if x is too small, the process will die out almost surely, so we let c be "the largest number such that an individual with type less than or equal to c cannot produce individuals with types larger than c". i.e.

$$c = \frac{2b}{1-a}.$$

Another interesting quantity is d defined by

$$d = (1 - 3a + 2b)/4a.$$

Indeed, if $d \leq c$, then the process will die out almost surely too (for yet another type of x's): the offspring type of x will be less than c if we choose $x = 3/4 - a/4 + b/2$. See Figure 6. Hence we *need* $d > c$. It's easily shown that $d > c$ implies $0 < 3a + 2b < 1$ given that $1/4 < a < 1/3$.

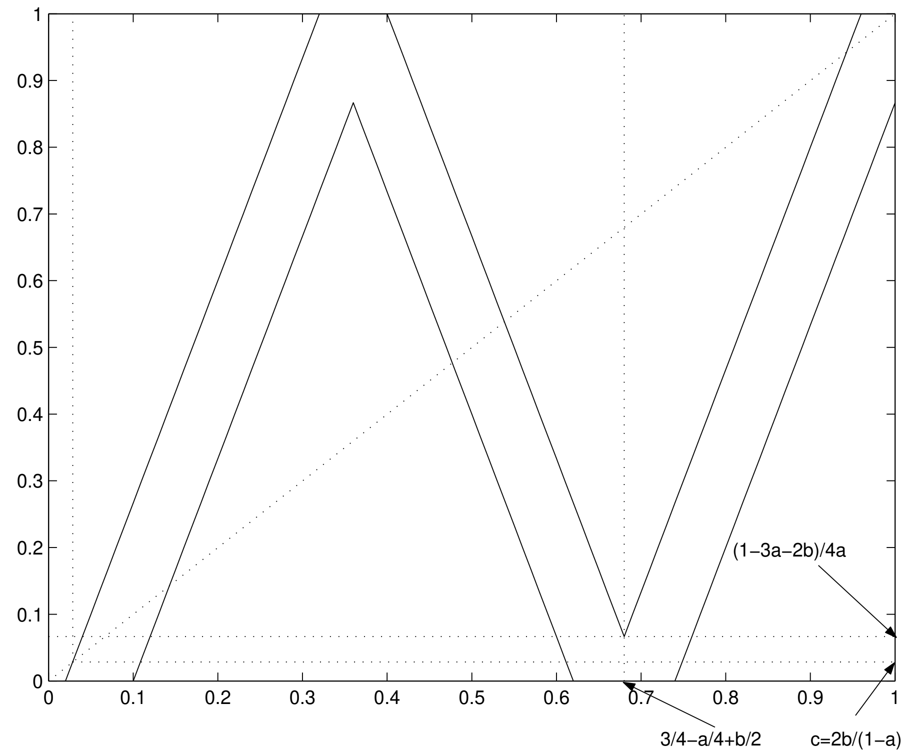


FIGURE 6

Lemma 3.2. *Let a and b be fixed real numbers with $1/4 < a < 1/3$ and $d > c$, then there exists n such that for arbitrary $x \in (c + \varepsilon, 1]$, $|E_n(x)| = 1$, where $0 < \varepsilon < d - c$.*

Proof. Since $x > c + \varepsilon$, we have that:

$$|E_1(x)| \geq \min\{E_1(c + \varepsilon), d, 2d\} > c + \varepsilon.$$

See Figure 7. Thus we truncate the first c-part of $E_1(x)$, i.e., we concentrate our attention on the interval $[f_1(x) + c, g(x)]$. The part $E_2^*(x) \subset E_2(x)$ which is generated by this interval can be estimated as follows. If there are two fold points in this interval, then certainly $|E_2^*(x)| = 1$, and if there is at most one fold point in the interval, then the contribution to $|E_2^*(x)|$ is at least one of $c + \varepsilon + (\varphi - f_1(x) - c) \cdot \frac{1}{a}$

and $c+\varepsilon+(g_1(x)-\varphi)\cdot\frac{1}{a}$, where φ is the fold point that drops in the interval $[f_1(x)+c, g(x)]$. If there are no fold points in the interval, we can let $\varphi=(g_1(x)-f_1(x)-c)/2$. Therefore,

$$|E_2^*(x)| > (c+\varepsilon)+\frac{|E_1(x)|-c}{2}\cdot\frac{1}{a}.$$

So, using that $a<1/3$ in the third step, we obtain

$$\begin{aligned} |E_2(x)|-|E_1(x)| &\geq |E_2^*(x)|-|E_1(x)| \\ &> c+(|E_1(x)|-c)/2a-|E_1(x)| \\ &> c+3(|E_1(x)|-c)/2-|E_1(x)| \\ &= (|E_1(x)|-c)/2 \\ &> \varepsilon/2. \end{aligned}$$

By induction, we obtain $|E_n(x)|\geq\min(1,|E_1(x)|+(n-1)\varepsilon/2)$. This ends the proof. (Actually it seems that the growth of $|E_n(x)|$ is geometrical rather than linear.) □

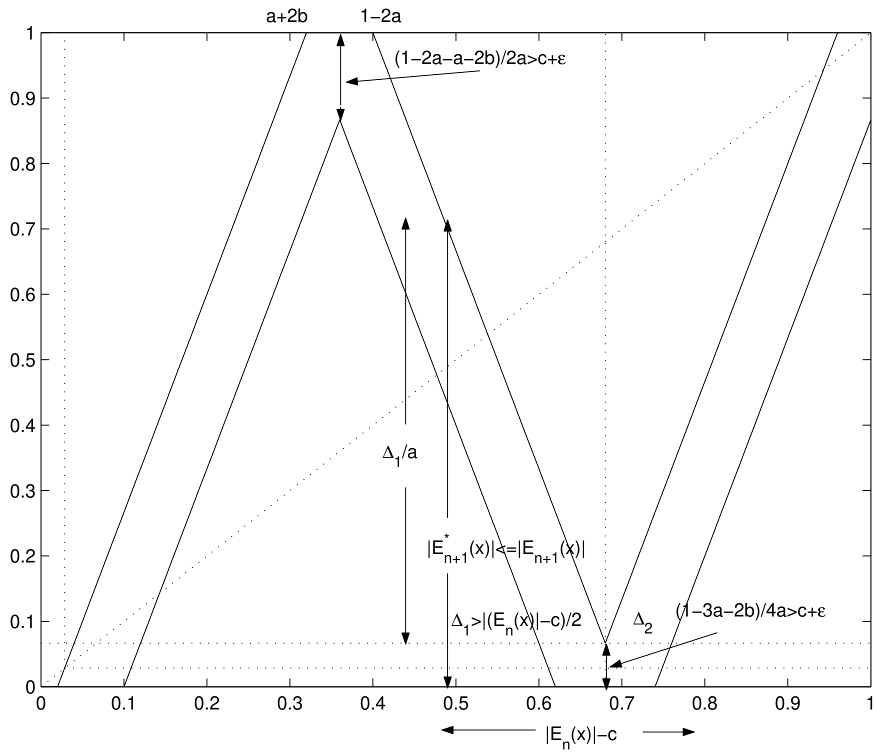


FIGURE 7

By using Lemma 3.2 and property (3.1), we can see that for any $\varepsilon > 0$, there is an integer n_0 such that for all $(x, y) \in (c+\varepsilon, 1] \times (c+\varepsilon, 1)$, $m_{n_0}(x, y) > 0$. Moreover, $m_{n_0}(x, 1) > 0$ where $x \in (c + \varepsilon, 1]$ by using property (3.2) and equation

$$m_{n+1}(x, y) = \int_{\Omega} m_n(x, z) m_1(z, y) dz.$$

So we can take

$$r_1 = \min_{(x,y) \in [c+\varepsilon,1] \times [c+\varepsilon,1]} m_{n_0}(x, y)$$

There still is a problem for x, $y < c$. The trick used by Larsson is now to change the branching process (which is counting *all* the intervals I in $C_1^n - C_2^n$) by considering only intervals that have length larger than $c+\varepsilon$. When we use $\Omega = (c+\varepsilon, 1]$ as the new type space, for this smaller Ω the left part of condition (2.1) holds.

Note that right part of condition (2.1) is trivially fulfilled, since one can choose $r_2 = 3t/a$. (Recall that $t = 1/2 - 3a/2 - b$, the largest shift of the intervals E^1 and E^2, so t is the largest value taken by the density of Y_i, $i = 1, 2, 3$.) Thus we are in a position to apply Theorem 2.1.

To calculate the largest eigenvalue ρ of the m_1-operator on the type space of $\Omega = (c + \varepsilon, 1]$ Larsson first calculates a certain eigenvalue of m_1-operator on the type space of $\Omega = (0, 1]$. Notice that

$$m_1(x, y) = f_{X_1^1}(y|x) + f_{X_1^2}(y|x) + f_{X_1^3}(y|x).$$

It will be shown that the constant function $\nu(x) = \mathbb{I}_{(0,1]}(x)$ is a left eigenfunction for each of the operators defined by

$$T_i g(y) = \int_{(0,1]} g(x) f_{X_1^i}(y|x) dx, \ i = 1, 2, 3.$$

To prove that ν is an eigenfunction of T_2, we must show that that

$$\int_{(0,1]} \nu(x) f_{X_1^2}(y|x) dx = \lambda \nu(y).$$

First, given $\nu(x) = \mathbb{I}_{(0,1]}(x)$ and $y \in [0,t] \subset [-1+t, 1-t]$,

$$\begin{aligned}
\int_{(0,1]} \nu(x) f_{|1-x+Y_2|}(y)dx &= \int_{(0,1]} f_{1-x+Y_2}(y)dx + \int_{(0,1]} f_{1-x+Y_2}(-y)dx \\
&= \int_{(0,1]} f_{Y_2}(y-x)dx + \int_{(0,1]} f_{Y_2}(y+x)dx \\
&= \int_{(0,1]} f_{Y_2}(y-x)dx + \int_{[-1,0)} f_{Y_2}(y-x)dx \\
&= \int_{[-1,1]} f_{Y_2}(y-x)dx \\
&= 1,
\end{aligned}$$

Thus $\int_{(0,1]} \nu(x) f_{X_1^2}(y|x)dx = \int_{(0,1]} \nu(x) f_{(1-|1-x+Y_2|/a)^+}(y|x)dx = a$ in $(0,1]$. Therefore μ is an eigenfunction of T_2 with eigenvalue a. In the same way ν is an eigenfunction of T_1 with eigenvalue $2a$. Summing them up, we have $\int_{\Omega=(0,1]} \nu(x) m_1(x,y)dx = 4a \cdot \nu(y)$. So ν is an eigenfunction of m_1 on the type space of $\Omega = (0,1]$ with eigenvalue $4a$. Note that since ν is strictly positive, $4a$ actually is the largest eigenvalue of the operator, this eigenvalue satisfies the condition of Theorem 2.2.

Next Larsson shows that if we change the type space from $\Omega = (0,1]$ to $\Omega = (c+\varepsilon, 1]$, the largest eigenvalue will not change much, so that we can arrange that it will still satisfy the condition of Theorem 2.2, i.e., it is bigger than 1.

Let T be the operator defined by

$$Tg(y) = \int_{\Omega=(0,1]} g(x) m_1(x,y)dx$$

and let

$$\begin{aligned}
T_0 = \mathbb{I}_{(0,c]} \cdot T, \quad T_1 = \mathbb{I}_{(c,1]} \cdot T, \\
\nu_0 = \mathbb{I}_{(0,c]} \cdot \nu, \quad \nu_1 = \mathbb{I}_{(c,1]} \cdot \nu.
\end{aligned}$$

Recall here that c is the largest number such that an individual with type less than or equal to c can not produce individuals with types larger than c. Using that for this choice of c, Tg has support in $(0,c]$ for all g with support in $(0,c]$ we get

$$\begin{aligned}
4a \cdot (\mu_0 + \mu_1) &= 4a \cdot \mu \\
&= T\mu \\
&= (T_0 + T_1)(\mu_0 + \mu_1) \\
&= T_0\mu_0 + T_0\mu_1 + T_1\mu_0 + T_1\mu_1.
\end{aligned}$$

Considering these functions on $(c, 1]$, we see that

$$4a \cdot \mu_1 = T_1 \mu_1,$$

since $T_1 \mu_0 = 0$. Therefore $4a$ is also an eigenvalue of T_1 with the new eigenfunction μ_1. The operator T_1 corresponds to the type space $\Omega = (c, 1]$.

Notice that if m_1 is a density with eigenvalue λ and m_1^* is a density close to m_1 in the L^2 sense, then the corresponding operator T_1^* has an eigenvalue close to λ. By using this, if we increase c slightly, we change the operator $T_1 = \mathbb{I}_{(c,1]} \cdot T$ to $T_2 = \mathbb{I}_{(c+\varepsilon,1]} \cdot T$. Since the change in the operator when we increase c is continuous, we can make the change ε small enough to ensure that the new operator T_2 still has an eigenvalue larger in magnitude than 1. Recall here that $M_1(x, \Omega)$ on the type space $\Omega = (c + \varepsilon, 1]$ satisfies the condition in Theorem 2.1.

One can now apply the results of Theorem 2.1 to conclude that the largest eigenvalue ρ of $T_2 = \mathbb{I}_{(c+\varepsilon,1]} \cdot T$ is larger than 1 if one chooses ε small enough.

4. Finding an interval

Larsson claims that Theorem 2.2 implies the existence of a integer $N = N(A)$ and a real number $k = k(A)$ that with a positive probability, say $q = q(A)$, we have

$$Z_n(x, A) \geq k\rho^n \ \textit{for all}\ n \geq N.$$

Then he claims to prove that: with positive probability, there is a certain interval I_0 contained in $C_1 - C_2$. Following a remark from Larsson we will show how to finish the proof starting from this.

Notice that the event $E = \{C_1 - C_2 \textit{ contains an interval}\}$ contains the event $\hat{E} = \{C_1 - C_2 \textit{ contains } I_0\}$, thus $p := P(E) \geq P(\hat{E}) > 0$. It's easy to verify that $P(x + C_1 - C_2 \textit{ contains an interval}) = P(E) = p$ where x is the shift. Now we will show that the probability p is actually 1. The self-similarity property of the Cantor sets yields that for all $n \geq 1$

$$C_i = \bigcup_{j=1}^{2^n} C_i^j, \ i = 1, 2.$$

where the C_i^j (apart from a scaling) have the same distribution as the C_i. Therefore the probability that "$C_1^k - C_2^l$ contains an interval" is also equal to p for all k and l, by shift invariance. Since all these intervals are independent, the probability that $C_1 - C_2$ contains an interval is

$$\begin{aligned} p &= 1 - P(\bigcap_{i=1}^{2^n}\bigcap_{j=1}^{2^n} C_1^i - C_2^j \textit{ does not contains an interval}) \\ &\geq 1 - P(\bigcap_{i=1}^{2^n} C_1^i - C_2^i \textit{ does not contains an interval}) \\ &= 1 - \prod_{i=1}^{2^n} P(C_1^i - C_2^i \textit{ does not contains an interval}) \\ &= 1-(1-p)^{2^n}. \end{aligned}$$

Since $1-(1-p)^{2^n}$ is arbitrarily close to 1 if n is big enough and $p > 0$, we obtain $p = 1$. This ends the proof of Theorem 1.1.

References

[1] Per Larsson, The difference set of two Cantor sets, Doctoral Dissertation in Mathematics, U.U.D.M. Report 1991:11

[2] T. Harris, The Theory of Branching Processes. Springer, Berlin, 1963