Translations, measure and dimension

Dissertation submitted to

The Hungarian Academy of Sciences

for the degree "MTA Doktora"

Tamás Keleti

Eötvös Loránd University

Budapest

2009

Contents

Introduction **5**

Notation **7**

1 Sets without given patterns **9**

2 Covering the real line with small sets **13**
2.1 Copies of the same set 13
2.2 Shuffle the plane 14

3 Density and coverings in $\mathbb{R}^n$ **17**
3.1 The key result 17
3.2 A direct application 18
3.3 A covering property 19
3.4 The minimal density property 20
3.5 A Besicovitch type covering property 22

4 The measure of the intersection of two copies of a self-similar or self-affine set **25**
4.1 Self-affine and self-similar sets with the strong separation condition. 27
4.2 A lemma about invariant extension of measures 28
4.3 Intersection of translates of a self-affine Sierpiński sponge 29

5 Periodic decomposition of measurable integer valued functions **33**

Bibliography **37**

Supplements **43**

Introduction

This thesis is about the relation between the additive and measure theoretic structure of $\mathbb{R}$ and more generally of $\mathbb{R}^n$. This is done via different types of questions. These problems lead us to other areas of mathematics as well.

One of the central concepts we study is smallness. A subset H of $\mathbb{R}$ or $\mathbb{R}^n$ can be small in various different ways. In geometric measure theory H is small if its measure or dimension is small. If we consider the additive structure of $\mathbb{R}$ or $\mathbb{R}^n$ then there are many natural possible ways to define small sets. For example, one can call a subset small if few translates of it cannot cover the real line. Or one can call a set small if it does not contain a given pattern, say, arithmetic progression of length 3. One of our main goals is to decide weather smallness in geometric measure theory sense implies smallness in the additive structure of $\mathbb{R}$ or $\mathbb{R}^n$, and vice versa.

One can get some results easily. For example, it is clear that if a set has Lebesgue measure zero then one cannot cover the real line with countably many of its translates. One of the main results of Chapter 2 is that one cannot cover the real line by less than continuum many translates of a compact set with packing dimension less than 1 (Theorem 2.3). An other result about smallness of this type leads to results in group theory in Section 2.2.

Using the classical Lebesgue's density theorem one can easily show that if a set has positive (Lebesgue) measure then it contains similar copies of any given finite set. The most important open problem of this area is a conjecture of Erdős that states that no infinite set has this property; in other words, for any infinite set one can construct a set of positive measure that contains no similar copy of the given infinite set. In Chapter 1 we will see that having large Hausdorff dimension is not enough even for guaranteeing finite patterns in $\mathbb{R}$.

We also study the following type of questions about smallness and coverings: If a measurable set is covered by some given type of sets such that its density is small in each of the covering sets, does it imply that the set has small measure? We will see in Chapter 3 that if we allow any rectangles in the covering then the answer is negative, however, if we allow only axis-parallel rectangles then the answer is positive. The positive result leads us to covering results that are connected to classical covering results, which are important in harmonic analysis. By studying those collections of sets for which the answer is positive we meet some problems in geometry and as a spin-off we also get for example an inverse isoperimetric inequality.

If K is a classical set in geometry then the measure of the intersection of K and its translate $K + t$ is close to the measure of K if t is small, and positive whenever the intersection is nonempty. If K is a fractal set then the situation is much more interesting and completely different. The study of the size of the intersection of Cantor type sets has been a central research area in geometric measure theory and dynamical systems lately.

In Chapter 4 we to study the measure of the intersection of two Cantor type sets which are (affine, similar, isometric or translated) copies of a self-similar or self-affine set in $\mathbb{R}^d$. By measure here we mean natural self-similar or self-affine measure on one of the two sets. We get instability results stating that the measure of the intersection is separated from the measure of one copy. This strong non-continuity property is in sharp contrast with the well known fact that for any Lebesgue measurable set $H \subset \mathbb{R}^d$ with finite measure the Lebesgue measure of $H \cap (H + t)$ is continuous in t. We get results stating that the intersection is of positive measure if and only if it contains a relative open set. This result resembles some recent deep results stating that for certain classes of sets having positive Lebesgue measure and nonempty interior is equivalent. As an application we also get isometry (or at least translation) invariant measures of $\mathbb{R}^n$ such that the measure of the given self-similar or self-affine set is 1.

In Chapter 5 we study the relation of the additive and the measure structure of $\mathbb{R}$ via studying decompositions of (Lebesgue) measurable integer valued functions into sum of periodic functions with given periods. The central question we study is whether the existence of real valued measurable periodic decomposition of an integer valued function implies integer valued (or at least almost everywhere integer valued) periodic measurable decomposition with the same periods. We will see that this is not always true and we will characterize those periods for which this holds. For this first we characterize those periods for which the decomposition of a measurable $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ function into the sum of periodic measurable $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ functions with these given periods is essentially unique.

This thesis is based on papers [Suppl-1],...,[Suppl-8], which are supplemented.

Acknowledgements. I would like to thank László Surányi and Miklós Laczkovich for introducing me to mathematics, real analysis and geometric measure theory. In fact, I learnt much more than mathematics from them. I also learnt a lot from David Preiss and Cliff Weil, and also from members of the younger generation, who used to be my students at some point and later became my collaborators: Marianna Csörnyei, Miklós Abért, Márton Elekes, Zoltán Ruzsa, Tamás Mátrai, András Máthé, Zoltán Gyenes, Viktor Harangi and Péter Maga. I am grateful to my all other coauthors as well: Udayan B. Darji, Vilmos Prokaj, Petr Holický, Gyula Károlyi, Mihalis Kolountzakis, Géza Kós, Imre Z. Ruzsa, Bálint Farkas, Szilárd Gy. Révész and Elliot Paquette.

I would like to thank all the help of Maarit and Esa Järvenpää and the hospitality of the University of Jyväskylä, where I completed this thesis. I am also grateful for the Eötvös Loránd University, to the Alfréd Rényi Institute of Mathematics, to the University College London, to the Michigan State University and to the University of Crete for providing me the opportunity to work and do research.

I acknowledge the financial support of many OTKA grants, the Széchenyi Professor Scholarship and the Bolyai János Research Scholarship.

I would like to thank the technical help of Margit Gémes, Árpád Tóth and Márton Elekes.

Finally, special thanks to my Mother, my Father, my wife Gabi and my children Doma and Hanga for their support and for the inspiration.

Notation

The following notations are used throughout the thesis. Most notions that are needed only in one of the chapters are defined there.

The sets of real, rational and integer numbers are dented by $\mathbb{R}, \mathbb{Q}$ and $\mathbb{Z}$, respectively.

By a *Borel measure* we mean a measure defined on the Borel sets. It is called a *continuous* Borel measure if the measure of any singleton is zero.

If not specified otherwise then by *measure* and *measurability* we always mean Lebesgue measure and Lebesgue meausarability. The Lebesgue measure of a set A is denoted by $|A|$. By the *density* of a set A in a set B we mean $\frac{|A\cap B|}{|B|}$, or if we consider some other measure μ, then $\frac{\mu(A\cap B)}{\mu(B)}$.

Let diam denote the diameter. The *s-dimensional Hausdorff measure* of a set $A \subset \mathbb{R}^n$ is defined as

$$\lim_{\delta\to 0+}\left(\inf\left\{\sum_{i=1}^{\infty}(\operatorname{diam}(E_i))^s : A \subset \bigcup_{i=1}^{\infty} E_i, \operatorname{diam}E_i < \delta\right\}\right).$$

The *s-dimensional packing measure* of a set $A \subset \mathbb{R}^n$ is defined as follows. Let

$$P^s(A) = \lim_{\delta\to 0+}\left(\sup\sum_i(\operatorname{diam}(B_i))^s\right),$$

where the supremum is taken over all disjoint families (packings) of closed balls $\{B_1, B_2, \ldots\}$ such that $\operatorname{diam}(B_i) < \delta$ and the centers of B_i's are in A. This P^s is not σ-additive and so the *s-dimensional packing measure* of a set $A \subset \mathbb{R}^n$ is defined as

$$\mathcal{P}^s(A) = \inf\left\{\sum_{i=1}^{\infty}P^s(A_i) : A = \cup_i A_i\right\}.$$

The *Hausdorff/packing dimension* of a set $A \subset \mathbb{R}^n$ is the infimum of those s-s for which the Hausdorff/packing measure of A is zero. The packing dimension will be denoted by $\dim_P$.

If we replace $(\operatorname{diam}(E_i))^s$ by $h(\operatorname{diam}(E_i))$ in the definitions of Hausdorff/ packing measure, where $h : [0,\infty) \to [0,\infty)$ is a nondecreasing function with $h(0) = 0$ then we get *generalized Hausdorff/packing measure* with *gauge-function* h. (See more on these notions e.g. in [Ma95].)

By an *interval* in $\mathbb{R}^n$ we mean an n-dimensional axis-parallel open rectangle: the Cartesian product of n open (1-dimensional) intervals.

For $1 \le q < \infty$ we denote the L_q norm of a function $f : \mathbb{R}^n \to \mathbb{R}$ by $\|f\|_q$; that is, $\|f\|_q = (\int_{\mathbb{R}^n}|f|^q)^{1/q}$. A measurable function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be in L_q if its L_q norm is finite, it is said to be locally L_q if $\int_B |f|^q < \infty$ for any bounded measurable set B. The L_∞ norm is the smallest number s such that $|f| \leq s$ holds almost everywhere. A measurable function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be in L_∞ if its L_∞ norm is finite, in other words, if it is *essentially bounded*.

By a *perfect* set we mean a closed set without isolated points. The *relative interior* of a set $A \subset B$ in a set B is denoted by $\mathrm{int}_B A$.

We denote by dist the Euclidean distance.

A mapping $g : \mathbb{R}^n \to \mathbb{R}^n$ is called a *similitude* if there is a constant $r > 0$, called *similarity ratio*, such that $\mathrm{dist}(g(a), g(b)) = r \cdot \mathrm{dist}(a, b)$ for any $a, b \in \mathbb{R}^d$. A set B is a *similar copy* of A if $B = f(A)$ for some similitude f.

The translate of a set H by a vector t is denoted by $H + t$; that is,

$$H + t = \{h + t : h \in H\}.$$

Chapter 1

Sets without given patterns

As we noted in the Introduction, a subset of the reals with positive Lebesgue measure contains similar copies of any given finite set. Knowing this one might hope that something similar might be true for subsets of $\mathbb{R}$ with sufficiently large Hausdorff dimension. In this chapter we show that this is not the case, we can construct in $\mathbb{R}$ compact sets with Hausdorff dimension 1 that avoid given patterns.

We call a set of 3 or 4 real numbers a *parallelogram* if it is of the form $\{a, a+u, a+v, a+u+v\}$, where $a \in \mathbb{R}$ and $0 < u \leq v$. First we want to avoid parallelograms; that is, we want to construct compact set $A \subset \mathbb{R}$ with Hausdorff dimension 1 such that A contains no parallelogram. (In particular, such an A clearly cannot contain any arithmetic progression of length at least 3.)

Note that in $\mathbb{R}$ a set does not contain parallelogram if and only if it intersects each of its (non-identical) translates by at most one point. Therefore the following theorem gives a set of Hausdorff dimension 1 that contains no parallelogram.

Theorem 1.1. *[Suppl-1, Theorem 1] There exists a compact set in $\mathbb{R}$ with Hausdorff dimension 1 that intersects each of its (non-identical) translates in at most one point.*

The first result of this type was obtained by P. Mattila in 1984 [Ma84], who constructed compact subsets A and B of $\mathbb{R}$ with Hausdorff dimension 1 such that the intersection of A and any translate of B contains at most one point. The above result shows that - if we allow only non-identical translations - one can also have $A = B$.

In Chapter 4 we will see an other peculiar property of the set constructed in Theorem 1.1: it is a compact set $C \subset \mathbb{R}$ with Hausdorff dimension 1 such that any continuous Borel measure μ on C can be extended to a translation invariant Borel measure on $\mathbb{R}$.

Finding or avoiding given patterns in a set of given size is also connected to the Erdős conjecture we mentioned in the introduction, which states that for any infinite set $A \subset \mathbb{R}$ there exists a set $E \subset \mathbb{R}$ of positive Lebesgue measure which does not contain any similar (i.e. translated and rescaled) copy of A. It is known that slowly decaying sequences are not counter-examples [Fa84, Bo87, Ko97] (see e.g. [HL98, Ko83, Sv00] for other related results) but nothing is known about any infinite sequence that converges to zero at least exponentially. On the other hand, as we already mentioned, it follows easily from Lebesgue's density theorem that any set $E \subset \mathbb{R}$ of positive Lebesgue measure contains similar copies of every finite sets.

Bisbas and Kolountzakis [BK06] gave an incomplete proof of the following related statement: For every infinite set $A \subset \mathbb{R}$ there exists a compact set $E \subset \mathbb{R}$ of Hausdorff dimension 1 such that E contains no similar copy of A. Kolountzakis asked whether the same holds for finite sets as well. Iosevich asked a similar question: if $A \subset \mathbb{R}$ is a finite set and $E \subset [0,1]$ is a set of given Hausdorff dimension, must E contain a similar copy of A?

I answered these questions by showing that for any set $A \subset \mathbb{R}$ of at least 3 elements there exists a 1-dimensional set that contains no similar copy of A. In fact, I proved a bit more by proving the following theorem, which immediately yields the following two corollaries.

Theorem 1.2. *[Suppl-2, Theorem 1] For any countable set $A \subset (1,\infty)$ there exists a compact set $E \subset \mathbb{R}$ with Hausdorff dimension 1 such that if $x < y < z, x, y, z \in E$ then $\frac{z-x}{z-y} \notin A$.*

Corollary 1.3. *[Suppl-2, Corollary 2] For any sequence $B_1, B_2, \ldots \subset \mathbb{R}$ of sets of at least three elements there exists a compact set $E \subset \mathbb{R}$ with Hausdorff dimension 1 that contains no similar copy of any of $B_1, B_2, \ldots$.*

Corollary 1.4. *[Suppl-2, Corollary 2] For any countable set $B \subset \mathbb{R}$ there exists a compact set $E \subset \mathbb{R}$ with Hausdorff dimension 1 that intersects any similar copy of B in at most two points.*

Laba and Pramanik [LP09] obtained a positive result by proving that if a compact set $E \subset \mathbb{R}$ has Hausdorff dimension sufficiently close to 1 and E supports a probability measure whose Fourier transform has appropriate decay at infinity then E must contain non-trivial 3-term arithmetic progressions. It would be interesting to know whether similar conditions could guarantee other finite patterns as well.

Perhaps one can even find conditions weaker than having positive measure that implies that a compact subset of $\mathbb{R}$ contains similar copies of all finite subsets. This is not impossible since Erdős and Kakutani [EK57] constructed a compact set of measure zero with this property. The Erdős-Kakutani set has Hausdorff dimension 1 but, using ideas from [ES04], András Máthé [MaA09] constructed such a set with Hausdorff dimension 0. (This example of Máthé will also appear at the end of Section 2.1.) However, the packing dimension of such a set must be 1, since the argument of the proof of Theorem 2.3 [Suppl-3, Theorem 2] (which we will discuss in Section 2.1) gives that if a compact set $C \subset \mathbb{R}$ contains similar copies of all sets of n points then C has packing dimension at least $(n-2)/n$.

Recently Péter Maga [MaP] has generalized some of the above results using similar arguments. Generalizing Theorem 1.1 he has constructed for any n a compact set of Hausdorff dimension n in $\mathbb{R}^n$ that intersects each of its (non-identical) translates by at most one point. He could also obtain results in the spirit of Corollary 1.3 by showing that in $\mathbb{R}^2$ for any set B of at least 3 elements there exists a compact set in $\mathbb{R}^2$ of Hausdorff dimension 2 that contains no similar copy to B. The method does not seem to work in higher dimension and it is an intriguing problem to decide for example how large can the Hausdorff dimension of a set in $\mathbb{R}^3$ be that does not contain three points that form a regular triangle. Embedding the above two-dimensional example to $\mathbb{R}^3$ we can reach 2 and some heuristics suggest that perhaps one cannot go further. Getting a result in the opposite direction, that would say that large Hausdorff dimension implies some patterns would be very interesting. There is ongoing research in this direction.

The proofs of Theorems 1.1 and 1.2, and also of the above mentioned generalizations of P. Maga, uses the same trick as the devil in the following infinite game.

Devil's game: *At each step you give one Euro coin to the devil and he gives you two Euro coins. But he can choose the coin you give to him and you have to play infinitely many steps.*

If you play this game against the devil then he will enumerate all coins and at each step he chooses your coin with the smallest number. This way, although you have more and more money, after infinitely many steps the devil will have all the coins.

Similar trick works in the proofs of the above theorems. We enumerate the configurations we have to exclude, then at each step we exclude one of them and may cause many bad configurations but, as in the Devil's game, eventually we exclude all bad configurations.

Chapter 2

Covering the real line with small sets

2.1 Copies of the same set

When is $\mathbb{R}$ the union of less than continuum many translates of a given compact subset of $\mathbb{R}$? Of course, if the compact set has non-empty interior, then $\mathbb{R}$ is easily seen to be the union of countably many translates of the compact set. On the other hand, if we assume the continuum hypothesis, then it follows from the Baire category theorem that there is no such nowhere dense compact set.

Gary Gruenhage observed that it is consistent with ZFC that given a compact set of positive Lebesgue measure one can find less than continuum many translates of it whose union is $\mathbb{R}$. Hence, for nowhere dense compact sets of positive Lebesgue measure the question whether $\mathbb{R}$ can be written as less than continuum many translates of the given set is independent of ZFC.

Gruenhage also showed that $\mathbb{R}$ is not the union of less than continuum many translates of the standard "middle 1/3 Cantor set". Motivated by these results, he asked the following natural question:

Problem 2.1. *Is it true that $\mathbb{R}$ is not the union of less than continuum many translates of any compact set of Lebesgue measure zero?*

Since continuum hypothesis implies positive answer, a negative answer to this problem would require some extra set-theoretic assumption.

Later, Daniel Mauldin asked a slightly modified question. Namely,

Problem 2.2. *Is it true that $\mathbb{R}$ is not the union of less than continuum many translates of any compact set of Hausdorff dimension less than 1?*

The main result of our paper [Suppl-3] with Udayan B. Darji is that if we consider packing dimension instead of Hausdorff dimension then the answer is affirmative:

Theorem 2.3. *[Suppl-3] Less than continuum many translated copies of a compact subset of $\mathbb{R}$ with packing dimension less than 1 cannot cover the real line.*

In fact, we proved the following stronger result, which also gives affirmative answer to a question of Ronnie Levy, who asked whether it is true that $\mathbb{R}$ is not the union of less than continuum many similar copies of the standard middle 1/3 Cantor set.

Theorem 2.4. *[Suppl-3, Theorem 2.5] Less than continuum many similar copies of a compact subset of* $\mathbb{R}$ *with packing dimension less than* 1 *cannot cover the real line.*

We proved Theorem 2.4 by constructing a nonempty perfect set P that intersects every similar copy of a given compact set C with packing dimension less than 1 in a finite set. Since any nonempty perfect set has cardinality continuum this gives that one cannot even cover P by less than continuum many similar copies of C.

The following property of the packing dimension (which does not hold for Hausdorff dimension) plays a crucial role in the proof: for any any Borel sets we have $\dim_p(A \times B) \leq \dim_p(A) + \dim_p(B)$ (see e.g. in [Ma95]).

As a possible way of attacking Problem 2.1 we posed the following question.

Problem 2.5. *[Suppl-3, Problem 3.1] Is there a compact set* C *of Lebesgue measure zero such that every perfect set intersects at least one of the translates of* C *in uncountably many points?*

A negative answer would clearly imply positive answer to Problem 2.1. Although a positive answer does not imply anything directly, at least it does not have to depend on the axioms.

Later this approach turned out be successful for answering Problem 2.1: Márton Elekes and Juris Steprāns [ES04] gave a positive answer to Problem 2.5 in ZFC and then they proved that a negative answer to Problem 2.1 is consistent with ZFC. In fact, what they showed was that the Erdős-Kakutani set, which we mentioned in the previous chapter, is a good example for both problems.

Recently, the question of Mauldin (Problem 2.2) has been also answered. András Máthé [MaA09], using the ideas of Elekes and Steprans, constructed a zero Hausdorff dimensional compact set for which it is consistent with ZFC that less than continuum many translates of it covers the real line. (This is the same set we mentioned in the previous chapter as an example of a compact set with zero Hausdorff dimension that contains similar copies of all finite subsets of $\mathbb{R}$.) Thus our result Theorem 2.3 is sharp in the sense that it is very far from being true for Hausdorff dimension.

2.2 Shuffle the plane

In the previous section we tried to cover the real line by few copies of a fixed small set. Now we want to cover the real line by few small sets. This time we consider a set "small" if it has continuum many pairwise disjoint translates. Although one may guess that less than continuum many small sets (in the above sense) cannot cover the real line either, we observed with Miklós Abért that even countably many is enough. In fact, we proved the following slightly stronger result.

Lemma 2.6. *[Suppl-4, Lemma 5] One can give a countable partition* $\cup_{n=1}^{\infty} A_n = \mathbb{R}$ *and continuum many translated copies of every* A_n *such that the collection* $\{A_n + t_{n,\alpha} : n \in \mathbb{N}, \alpha \in [0,1)\}$ *of all translated copies are pairwise disjoint.*

Somewhat surprisingly this lemma eventually led to a purely group-theoretic result. For this we studied those transformations that one can obtain by composing the following very simple ones:

Definition 2.7. By a *vertical* (resp. *horizontal*) *slide* we mean an $\mathbb{R}^2 \to \mathbb{R}^2$ map of the form $(x,y) \to (x, y + f(x))$ (resp. $(x,y) \to (x + g(y), y)$), where f (resp. g) is an arbitrary $\mathbb{R} \to \mathbb{R}$ function.

By a *slide* we mean a vertical or horizontal slide.

Note that geometrically a vertical (resp. horizontal) slide means a transformation of the plane in which we translate vertical (resp. horizontal) lines vertically (resp. horizontally).

Clearly any slide is a permutation of the plane, so the question is which permutations we can get by using (finitely many) slides. One can also ask the following (weaker) question: When can a subset of the plane be transformed to an other subset using (finitely many) slides? Clearly, the sets must have the same cardinality and their complements must have the same cardinality, too - so the question is whether these conditions are sufficient or there exist other invariants of these maps.

Our main result is the following:

Theorem 2.8. *[Suppl-4, Theorem 2] Any permutation of the plane can be obtained by a fixed number* (209) *of slides. That is, for any permutation p of the plane there exist $\mathbb{R} \to \mathbb{R}$ functions $f_1, \ldots, f_{105}$ and $g_1, \ldots, g_{104}$ such that we have $p = F_1 \circ G_1 \circ \cdots \circ F_{104} \circ G_{104} \circ F_{105}$, where $F_i(x,y) = (x, y + f_i(x))$ and $G_i(x,y) = (x + g_i(y), y)$.*

Therefore the only invariants are the cardinality and the cardinality of the complement; a set can be mapped to an other set by finitely many slides if and only if they have the same cardinality and their complements have the same cardinality, too. In particular, there is no finitely additive non-negative function from the set of all subsets of the plane that agrees with ordinary area on squares and invariant under both vertical and horizontal slides.

Since both the vertical and the horizontal slides form (isomorphic) Abelian subgroups of the group of all permutations of $\mathbb{R}^2$, we also get the following (purely group-theoretic) result:

Corollary 2.9. *[Suppl-4, Corollary 3] The full symmetric group acting on a set of continuum cardinal is a product of finitely many* (209) *copies of two isomorphic Abelian subgroups.*

This is where our original motivation of this investigation came from. In [Ab02] the same result (excluding the constant 209) is proved for the full symmetric group acting on a countable set via the analogous result about slides on $\mathbb{Z} \times \mathbb{Z}$.

It is also proved in [Ab02] that the full symmetric group acting on any set is a product of finitely many Abelian subgroups. There - in the non-trivial infinite case - three Abelian subgroups were used and one of them was non-isomorphic to the other two.

Later Péter Komjáth [Ko02] extended Theorem 2.8 to arbitrary infinite abelian groups and he also showed that it is enough to use much less slides.

Lastly, we shed some light on how covering with small sets (Lemma 2.6) is used for constructing slides for a permutation of the plane (Theorem 2.8). The proof of Theorem 2.8 uses Lemma 2.6 via the following statement:

Claim 2.10. *[Suppl-4, Claim 6] The horizontal strip $S = \mathbb{R} \times [0, 1)$ can be mapped into the line* $\mathrm{e} = \mathbb{R} \times \{0\}$ *by 3 slides.*

If we have a construction like in Lemma 2.6 then first by a vertical slide we lift up each $A_n \times [0, 1)$ by n, then by a horizontal slide we can translate each $A_n \times \{n + \alpha\}$ by $t_{n,\alpha}$ ($n \in \mathbb{N}, \alpha \in [0, 1)$). Since the sets $\{A_n + t_{n,\alpha} : n \in \mathbb{N}, \alpha \in [0, 1)\}$ are pairwise disjoint we can map (in fact, project) these sets into $e = \mathbb{R} \times \{0\}$ by a vertical slide. Therefore Lemma 2.6 indeed implies Claim 2.10.

Using this claim and ideas from the proof of the above mentioned analogous result of M. Abért [Ab02] for $\mathbb{Z} \times \mathbb{Z}$ one gets Theorem 2.8.

Chapter 3

Density and coverings in $\mathbb{R}^n$

This chapter contains a result about the connection of the additive structure of $\mathbb{R}^n$ and smallness in measure, applications in different areas and some related results.

3.1 The key result

While the author was working on a modified problem of A. Carbery, the following question arose:

Question 3.1. *If a measurable subset of the unit square is covered by axis-parallel rectangles (contained in the unit square) such that its density is small in each rectangle, can we conclude that the set itself must have small measure?*

(Recall that by the *density* of A in B (with $|B| > 0$) we mean $\frac{|A \cap B|}{|B|}$, where $|.|$ means the (Lebesgue) measure.)

First we claim that if we allowed any (not necessary axis-parallel) rectangles then the answer to Question 3.1 would be negative. For this we recall a classical construction of Otto M. Nikodym (see e.g. [Gu75]). He constructed a set N in the unit square with measure 1 such that for each point $p \in N$ there is a straight line l_p so that $l_p \cap N = \{p\}$. Let N be such a Nikodym set and let H be a closed subset of N with measure at least $1 - \varepsilon$. Then, using that H is closed, for each $p \in H \subset N$ we can find a very narrow small rectangle R_p inside the unit square in the direction of l_p that contains p and in which the density of N is less than ε. Therefore H can be covered by rectangles (contained in the unit square) so that its density is less than ε in each rectangle, but still the measure of H is at least $1 - \varepsilon$.

The above observation explains why the answer is not as clear as first one might think and also that Question 3.1 is a problem about the connection of the additive structure of $\mathbb{R}^2$ and smallness in measure.

The key result of this chapter is an affirmative answer to Question 3.1, even in n-dimension:

Theorem 3.2. *[Suppl-5, Theorem 2.1] If H is a measurable subset of the open unit cube $(0,1)^n$ with $|H| > h$ and $\mathcal{R}$ is a class of intervals in $(0,1)^n$ that covers*

H, then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $(\frac{h}{2n})^n$; that is,

$$\frac{|H \cap R|}{|R|} > \left(\frac{h}{2n}\right)^n .$$

(Recall that by an *interval* of $\mathbb{R}^n$ we mean an n-dimensional axis-parallel open rectangle: the Cartesian product of n open (1-dimensional) intervals.)

This theorem and many other measure theoretic results of this chapter can be equivalently formulated as combinatorial ones, in the sense that the measurable sets and the intervals may be assumed to be finite unions of dyadic cubes and the coverings may be assumed to be finite. Nevertheless, the proof of this key result (Theorem 3.2) uses methods of analysis. A minimal operator analogous to the well known Hardy-Littlewood maximal operator (see e.g. [Gu75] or [Gu81]) is introduced:

The classical maximal operator for the class $\mathcal{I}^n$ of all intervals of $\mathbb{R}^n$ is defined as

$$M_n f(x) = \sup\left\{ \frac{1}{|R|} \int_R |f| : x \in R \in \mathcal{I}^n \right\}$$

for any locally L_1 function f on $\mathbb{R}^n$, while the minimal operator introduced and used in [Suppl-5] is defined as

$$m_n f(x) = \inf\left\{ \frac{1}{|R|} \int_R |f| : x \in R \in \mathcal{I}_0^n \right\},$$

where $\mathcal{I}_0^n$ denotes the class of all subintervals of $[0,1]^n$. A similar notion of minimal operator was also introduced in [CN95].

3.2 A direct application

A. Carbery asked the following question (see in [CCW]), which is still open:

For which functions $a : [0,1] \to [0,1]$ is it true that

($*$) *if H is a measurable subset of I^2 then one can always find 4 points of H such that they are the vertices of an axis-parallel rectangle with area at least $a(|H|)$?*

This question led I. Gyöngy to ask the following question:

For which functions $f : [0,1] \to [0,1]$ is it true that

($**$) *if H is a measurable subset of I^2 then one can always find 4 points of H such that they are the vertices of an axis-parallel rectangle R such that $|R \cap H| \geq f(|H|)$?*

Clearly it is harder to satisfy ($**$) then ($*$). However, using Theorem 3.2, it is easy to obtain a function satisfying ($**$) from a function that satisfies ($*$):

Proposition 3.3. *[Suppl-5, Proposition 3.4] If the function a satisfies ($*$) then $f(h) = \rho_2(h/2)a(h/2)$ satisfies ($**$), (where $\rho_2(h) = h^2/16$ is the function that appeared in Theorem 3.2 for $n = 2$).*

Since A. Carbery, M. Christ and J. Wright [CCW] proved that $a(h) = ch^2/\log(1/h)$ (for a suitable $c > 0$ and h small enough) satisfies ($*$) we get the following partial result for the question of I. Gyöngy:

Corollary 3.4. *[Suppl-5, Corollary 3.5] The function* $f(h) = c'h^4/\log(1/h)$ *(if* $h \le \delta < 1$ *and* $f(h) = f(\delta)$ *if* $h > \delta$*) satisfies* $(**)$*, where* c' *depends only on* δ*.*

Remarks 3.5. The following simple example shows that a function that satisfies $(*)$ cannot be greater than u^2. Let H_m be the union of the diagonal squares of the regular $m \times m$ subdivision of the unit square. Then clearly $|H_m| = 1/m$ and each axis-parallel rectangle with vertices in H_m has area at most $\frac{1}{m^2}$. It is unknown weather $a(u) = cu^2$ satisfies $(*)$ (for a sufficiently small $c > 0$).

Using a finite geometry construction of I. Reiman [Re58] it was shown in [Suppl-5, Example 3.6] that $(**)$ does not hold for the function $h^3 + h^4$ ($\sim h^3$). Therefore the best exponent (or the infimum of the exponents) for functions satisfying $(**)$ is in the interval [3,4]. This is the best we currently know.

All positive results of this section can be easily generalized to n-dimensional spaces and one gets that $a_n(u) = c_n u^{2^{n-1}+\alpha}$ satisfies the n-dimensional version of $(*)$ (for proper $c_n > 0$ depending only on n and α), while $f_n(h) = c'_n h^{n+2^{n-1}+\alpha}$ satisfies the n-dimensional version of $(**)$.

However, it is considerably more difficult to construct examples showing that we cannot have much better results than the above mentioned. The natural n-dimensional generalization of the example for $(*)$ (e.g. the union of those cubes of the regular $m \times \ldots \times m$ subdivision of the unit cube for which the sum of the coordinates is divisible by m) shows only that a function satisfying the n-dimensional version of $(*)$ cannot be greater than u^n. No natural generalization of the finite geometry example for $(**)$ seems to be known.

By standard probabilistic method, it is easy to prove the following combinatorial result:

One can select $O(m^{n-n/2^{n-1}})$ *points of the regular* n*-dimensional* $m \times \ldots \times m$ *lattice such that no* 2^n *of them are the vertices of an* n*-dimensional interval. Moreover, we can assume that we chose* $O(m^{n-1-n/2^{n-1}})$ *points of each* $n-1$*-dimensional* $m \times \ldots \times m$ *sublattice.*

Then, taking the union of the corresponding open cubes of a regular subdivision of the unit cube, we get a set H with measure $O(1/m^{n/2^{n-1}})$ such that if the vertices of an n-dimensional interval R are in H then $|R| < 1/m$ and $|R \cap H| < O(1/m^{1+n/2^{n-1}})$. Thus we get $O(u^{2^{n-1}/n})$ and $O(u^{(2^{n-1}/n)+1})$ functions that do not satisfy the n-dimensional versions of $(*)$ and $(**)$, respectively; which are still quite far from our positive results.

One possible way to obtain better examples is to show that, as Erdős [Er64] conjectured, one can also select $O(m^{n-1/2^{n-1}})$ points of the regular n-dimensional $m \times \ldots \times m$ lattice such that no 2^n of them are the vertices of an n-dimensional interval.

Then we would have $O(u^{2^{n-1}})$ and $O(u^{2^{n-1}+1})$ functions that do not satisfy the n-dimensional versions of $(*)$ and $(**)$, respectively, which would be quite close to our positive results.

3.3 A covering property

Although the key result is only about subsets of the unit cube of $\mathbb{R}^n$, it is not hard to apply it to get an analogous density result for an arbitrary measurable subset of $\mathbb{R}^n$:

Theorem 3.6. *[Suppl-5, Theorem 2.4] Suppose that H is a measurable subset of $\mathbb{R}^n$ with finite measure, $\mathcal{R}$ is a class of intervals of $\mathbb{R}^n$ that covers H and the density of H in $\cup\mathcal{R}$ is greater than $h > 0$. Then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $\rho_n(h) = \left(\frac{h}{2n}\right)^n$; that is,*

$$\frac{|H \cap R|}{|R|} > \rho_n(h) = \left(\frac{h}{2n}\right)^n.$$

Then, using greedy algorithm, this leads to the following covering result:

Theorem 3.7. *[Suppl-5, Theorem 2.5] For each $n \in \mathbb{N}$ there is a function $C_n : \mathbb{R}^+ \to \mathbb{R}^+$ such that for any collection $\mathcal{R}$ of intervals in $\mathbb{R}^n$ with $|\cup\mathcal{R}| < \infty$ and $\varepsilon > 0$ there exist $R_1, \ldots, R_m \in \mathcal{R}$ for which*

$$(i) \qquad \frac{|\cup\mathcal{R} \setminus \cup_{k=1}^m R_k|}{|\cup\mathcal{R}|} < \varepsilon$$

and

$$(ii) \qquad \frac{\sum_{k=1}^m |R_k|}{|\cup\mathcal{R}|} < C_n(\varepsilon).$$

Remark 3.8. The proofs give $C_n(\varepsilon) = D_n(1/\varepsilon)^{n-1}$, where $D_n = \frac{n}{4(n-1)}(8n)^n$ for $n \geq 2$ and $D_1 = 2$. This result is sharp in the sense that only the constants D_n can be improved, and similarly in Theorems 3.2 and 3.6 the exponent of h cannot be lowered (see [Suppl-5, Example 2.7]).

3.4 The minimal density property

In this section we compare the results of the previous section with the classical notions and results and we shall also see that density results can be used to sharpen covering results for more general classes of covering sets.

Recall that we denote the L_q norm of a function $f : \mathbb{R}^n \to \mathbb{R}$ by $\|f\|_q$; that is, $\|f\|_q = (\int_{\mathbb{R}^n} |f|^q)^{1/q}$, and the characteristic function of a set $A \subset \mathbb{R}^n$ is denoted by χ_A.

In the sequel let $\mathcal{B}$ be a class of nonempty open bounded subsets of $\mathbb{R}^n$ and $1 \leq q \leq \infty$.

Cordoba and Fefferman [CF75] introduced the following notion:

Definition 3.9. The class $\mathcal{B}$ is said to have the *covering property* V_q if there exist constants $C < \infty$ and $c > 0$ such that for any $\mathcal{R} \subset \mathcal{B}$ with $|\cup\mathcal{R}| < \infty$ we can find $R_1, \ldots, R_m \in \mathcal{R}$ such that

$$\text{(i')} \quad |\cup_{k=1}^m R_k| \geq c\,|\cup\mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad \|\sum_{k=1}^m \chi_{R_k}\|_q \leq C\,|\cup\mathcal{R}|^{1/q}.$$

It was proved in [CF75] that the class $\mathcal{I}^n$ of all intervals of $\mathbb{R}^n$ has the covering property V_q for any $1 \leq q < \infty$. Note that Theorem 3.7 states the following stronger covering property of the class $\mathcal{I}^n$ for $q = 1$:

Definition 3.10. [Suppl-5, Definition 4.2] We say that $\mathcal{B}$ has the *complete covering property* V_q (CV_q) if there exists a function $C : \mathbb{R}^+ \to \mathbb{R}^+$ such that for any $\varepsilon > 0$ and $\mathcal{R} \subset \mathcal{B}$ with $|\cup\mathcal{R}| < \infty$ we can find $R_1, \ldots, R_m \in \mathcal{R}$ such that

$$\text{(i)} \quad |\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup\mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad \|\sum_{k=1}^m \chi_{R_k}\|_q \leq C(\varepsilon)|\cup\mathcal{R}|^{1/q}.$$

Theorem 3.6 can be also expressed as the following property of the class $\mathcal{I}^n$ of all intervals of $\mathbb{R}^n$:

Definition 3.11. [Suppl-5, Definition 4.2] We say that $\mathcal{B}$ has the *minimal density property* (MDP) if there exists a function $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ such that if $H \subset \mathbb{R}^n$ is measurable with finite measure, $\mathcal{R} \subset \mathcal{B}$ covers H and the density of H in $\cup\mathcal{R}$ is $d > 0$ then one can find an $R \in \mathcal{R}$ in which the density of H is greater than $\rho(d)$; that is,

$$\frac{|R \cap H|}{|R|} > \rho\left(\frac{|H|}{|\cup\mathcal{R}|}\right).$$

As Theorem 3.6 implied Theorem 3.7 using greedy algorithm, one can similarly prove that MDP always implies CV_1. In fact, the converse also holds:

Theorem 3.12. *[Suppl-5, Theorem 4.6]*

$$MDP \Leftrightarrow CV_1.$$

That is, for any class $\mathcal{B}$ of nonempty open bounded subsets of $\mathbb{R}^n$ the minimal density property and the CV_1 property are equivalent.

Remark 3.13. Covering properties play essential role in the theory of differentiation of integrals (see e.g. in [Gu75] and [Gu81]). The basic question in this theory is whether the integral average $\frac{1}{|R_n|}\int_{R_n} f$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ converges to $f(x)$ if all R_n are from a given class $\mathcal{B}(x)$ of measurable sets that contain x and the diameter of R_n converges to 0. The classes $\mathcal{B}(x)$ form a so called *differentiation basis*. If for each function f from a function class $\mathcal{F}$ the above property holds for almost every $x \in \mathbb{R}^n$ then it is said that the differential basis *differentiates* $\mathcal{F}$. If a differential basis differentiates the class of characteristic functions of measurable sets then it is said to have the *difference property*. A differential basis is called to be a *Busemann-Feller differential basis* (or shortly *B-S basis*) if every set in $\cup_{x\in\mathbb{R}^n}\mathcal{B}(x)$ is open and $x \in R \in \cup_{y\in\mathbb{R}^n}\mathcal{B}(y)$ implies $R \in \mathcal{B}(x)$. If there exists a $c > 0$ so that for all sets R of all $\mathcal{B}(x)$ there exists a cube $Q \supset R$ so that $|R| > c|Q|$ then it is said to be a *regular* B-S basis.

If we assume that $\mathcal{B}$ is a class of nonempty open bounded subsets of $\mathbb{R}^n$ and each $x \in \mathbb{R}^n$ is contained in sets $R \in \mathcal{B}$ with arbitrarily small diameter then $\mathcal{B}$ clearly gives Busemann-Feller differentiation basis with $\mathcal{B}(x) = \{R : x \in R \in \mathcal{B}\}$. The basis we get for $\mathcal{B} = \mathcal{I}^n$ (the collection of intervals in $\mathbb{R}^n$) is called the *strong basis*.

It is a standard argument that the V_1 property (which clearly follows from the CV_1 property) of a B-F basis $\mathcal{B}$ implies that $\mathcal{B}$ differentiates the L_∞ functions, which clearly implies the density property of the basis $\mathcal{B}$. (In fact, as Busemann and Feller proved, differentiating L_∞ is equivalent to the density property). Therefore the minimal density property implies the density property. On the other hand, as we proved the minimal density property of $\mathcal{I}^n$, we have an alternative proof of Saks' strong maximal theorem. (For the definitions and results of this remark and the next remark see e.g. [Gu75] or [Gu81].)

Remark 3.14. Let $\mathcal{R}$ consist of sets that are the union of an open disc and an open sector with the same centre and twice larger radius.

Then $\mathcal{R}$ is clearly a regular B-F basis, so it has several standard nice properties (e.g. weak 1-1 property of the maximal operator, density property, it differentiates L_1 functions).

However $\mathcal{R}$ does not have the minimal density property. Indeed, we can cover an annulus by sets of $\mathcal{R}$ (with the same centre and radius) such that the density of the annulus is arbitrary small in each set.

Therefore

1. The minimal density property is strictly stronger than the density property.
2. The minimal density property and the CV_q properties of a class cannot be proved by using only the standard methods (e.g. properties of the maximal operator).

For any $1 \leq q < \infty$ the CV_q property clearly implies the V_q property and the CV_1 property. Somewhat surprisingly the converse also holds:

Theorem 3.15. *[Suppl-5, Corollary 4.12] If $\mathcal{B}$ has the MDP (or the equivalent CV_1) then*

$$V_q \Leftrightarrow CV_q \quad (1 \leq q < \infty).$$

Combining Theorem 3.6 and Theorem 3.15 we get that the class of intervals (in other words the strong basis) has the CV_q covering property, which in some sense, the strongest covering result for this basis:

Corollary 3.16. *[Suppl-5, Corollary 4.13] The class $\mathcal{I}^n$ of all intervals of $\mathbb{R}^n$ has the CV_q property for any $1 \leq q < \infty$.*

That is, for any $n \in \mathbb{N}$, $1 \leq q < \infty$ and $\varepsilon > 0$ there exists a constant $C(n, q, \varepsilon)$ such that if $\mathcal{R}$ is a family of n-dimensional intervals and $|\cup \mathcal{R}| < \infty$ then there is a finite sequence $R_1, \dots, R_m \in \mathcal{R}$ such that

$$\text{(i)} \quad |\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup \mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad \|\sum_{k=1}^m \chi_{R_k}\|_q \leq C(n,q,\varepsilon)|\cup \mathcal{R}|^{1/q}.$$

3.5 A Besicovitch type covering property

The results of the previous section show that the minimal density property of collection of open subsets of $\mathbb{R}^n$ might be very useful. On the other hand, so far we saw this property only for the collection $\mathcal{I}^n$ of all intervals of $\mathbb{R}^n$, and it is very far from trivial even for the simplest classes like for example the class of all balls. In this section we will give sufficient necessary geometric conditions for the minimal density property. In Remark 3.14 we saw that the class of sets that are the union of an open disc and an open sector with the same centre and twice larger radius does not have the minimal density theorem. The main result is the theorem below that shows that if the sets of $\mathcal{B}$ are "non-thorny" in the below defined sense then $\mathcal{B}$ has a much stronger property than the MDP or the CV_q properties: instead of (ii) of Definition 3.10, in this case, we have a better (Besicovitch type) control for the overlapping.

Definition 3.17. By a *drop* we mean the interior of the convex hull of a ball and a point (not contained in the ball). The *angle of the drop* is the angle between the line through the point and the center of the ball and any tangent line.

Let $0 < d < 1$ and $0 < \alpha < \pi/2$. We say that a bounded open set $H \subset \mathbb{R}^n$ is (d, α)*-non-thorny* if H is the union of drops with angle at least α and diameter at least $d \cdot \mathrm{diam} H$.

Theorem 3.18. *[Suppl-6, Theorem 3] Let $\mathcal{R}$ be a family of (d, α)-non-thorny sets in $\mathbb{R}^n$ with bounded diameter. Then for any $\varepsilon > 0$ one can choose sets $R_1, \dots, R_m \in \mathcal{R}$ such that*

(i)
$$|\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup \mathcal{R}| \qquad \textit{and}$$

(ii) the sequence $R_1, \dots, R_m$ can be distributed in M families of disjoint sets, where M depends only on n, d, α and ε.

The following statement follows immediately from Theorem 3.18.

Corollary 3.19. *[Suppl-6, Corollary 5] For any* $0 < d < 1$ and $0 < \alpha < \pi/2$, *any class of (d, α)-non-thorny sets in $\mathbb{R}^n$ has the CV_∞ property and consequently the CV_q property for any $1 \leq q < \infty$ and the minimal density property as well.*

Therefore this non-thornity is a sufficient condition for the MDP but it is in fact too strong. However, as we shall see below, quite large and important classes satisfy it.

Definition 3.20. A set $H \subset \mathbb{R}^n$ is said to be *star-shaped at x* if $\overline{xy} \subset H$ for every $y \in H$, where $\overline{xy}$ denotes the closed segment between x and y.

The *hub of H* ($\mathrm{hub}(H)$) is the set of all points at which H is star-shaped.

Let $r > 0$. We say that H is *r-star-shaped* if $\mathrm{hub}(H)$ contains an open ball with radius $r \cdot \mathrm{diam} H$.

Definition 3.21. A set $H \subset \mathbb{R}^n$ is *r-regular* if there exists a cube Q that contains H such that $|H|/|Q| > r$.

It is not hard to see (and probably well-known) that if H is a convex open r-regular set in $\mathbb{R}^n$ then H is r'-star-shaped, where r' depends only on n and r. It is easy to see that any r-star-shaped set is (d, α)-non-thorny, where d and α depend only on r. Thus Theorem 3.18 has the following consequences:

Corollary 3.22. *If $\mathcal{R}$ is a class of convex open r-regular sets or a class of r-star-shaped sets then for any $\varepsilon > 0$ one can select M subclasses of disjoint sets such that the selected sets cover the $1 - \varepsilon$ part of $\cup\mathcal{R}$, where M depends only on n, r and ε.*

Corollary 3.23. *Any class of convex open r-regular sets or of r-star-shaped sets in $\mathbb{R}^n$ has the CV_∞ property and consequently the CV_q property for any $1 \leq q < \infty$ and the minimal density property as well.*

The covering property of Theorem 3.18 (and Corollary 3.22) is similar to the Besicovitch property, the only difference is that, instead of all the centers, we cover a big part of the union. But, as the earlier mentioned example showed, in our case the Besicovitch property itself is not enough. However, in the proof of Theorem 3.18 we use the classical Besicovitch covering theorem (for balls) but we also need estimate for the "edge" of the union of drops. As a spin-off, this estimate also gives us the following reverse isoperimetric inequality for the union of star-shaped sets, which is interesting in itself:

Corollary 3.24. *[Suppl-6, Corollary 12] If E is the union of r-star-shaped sets in $\mathbb{R}^n$ with diameter D then we have*

$$\frac{\tilde{A}_+(E)}{|E|} \leq \frac{C(n,r)}{D},$$

where $\tilde{A}_+(E)$ denotes the upper outer surface area in the sense of Minkowski, that is

$$\tilde{A}_+(E) = \limsup_{\delta \to 0+} \frac{|S(E,\delta)| - |E|}{\delta},$$

where $S(E,\delta)$ is the open δ-neighborhood of E.

Remark 3.25. As a special case of Corollary 3.24, for example, we have that the ratio of the perimeter and the area of any finite union of (not necessary axis-parallel) unit squares is at most an absolute constant.

This special case was also posed by the author at the Schweitzer Miklós Mathematical Contest in 1998. In [Suppl-6] it was asked whether the best constant is 4 (which can be achieved by taking just one unit square). Currently the best result is due to Zoltán Gyenes [Gy09], who proved that the best constant is less than 5.6.

Chapter 4

The measure of the intersection of two copies of a self-similar or self-affine set

The study of the size of the intersection of Cantor sets has been a central research area in geometric measure theory and dynamical systems lately, see e.g. the works of Igudesman [Ig03], Li and Xiao [LX99], Moreira [Mo96], Moreira and Yoccoz [MY01], Nekka and Li [NL02], Peres and Solomyak [PS98]. For instance J-C. Yoccoz and C. G. T. de Moreira [MY01] proved that if the sum of the Hausdorff dimensions of two regular Cantor sets exceeds one then, in the typical case, there are translations of them stably having intersection with positive Hausdorff dimension.

In this chapter we study the measure of the intersection of two Cantor sets which are translated (sometimes affine, similar or isometric) copies of a self-similar or self-affine set in $\mathbb{R}^d$. By measure here we mean a self-similar or self-affine measure on one of the two sets, see the definitions later.

We get instability results stating that the measure of the intersection is separated from the measure of one copy. This strong non-continuity property is in sharp contrast with the well known fact that for any Lebesgue measurable set $H \subset \mathbb{R}^d$ with finite measure the Lebesgue measure of $H \cap (H + t)$ is continuous in t.

We get results stating that the intersection is of positive measure if and only if it contains a relative open set. This result resembles some recent deep results (e.g. in [LW96], [MY01]) stating that for certain classes of sets having positive Lebesgue measure and nonempty interior is equivalent. In the special case when the self-similar set is the classical Cantor set our above mentioned results were obtained by F. Nekka and Jun Li [NL02]. For other related results see also the work of Falconer [Fa85], Feng and Wang [FW09], Furstenberg [Fu70], Hutchinson [Hu81], Järvenpää [Ja99] and Mattila [Ma82], [Ma84], [M85].

As an application we also get isometry (or at least translation) invariant measures of $\mathbb{R}^d$ such that the measure of the given self-similar or self-affine set is 1.

The following notation and notions are used throughout this chapter.

Notation 4.1. We shall denote by $\cup^*$ the disjoint union.

Definition 4.2. Recall that a mapping $g : \mathbb{R}^d \to \mathbb{R}^d$ is called a *similitude* if there is a constant $r > 0$, called *similarity ratio*, such that $\text{dist}(g(a), g(b)) = r \cdot \text{dist}(a, b)$ for any $a, b \in \mathbb{R}^d$.

The *affine maps* of $\mathbb{R}^d$ are of the form $x \mapsto Ax + b$, where A is a $d \times d$ matrix and $b \in \mathbb{R}^d$ is a translation vector. Thus the set of all affine maps of $\mathbb{R}^d$ can be considered as $\mathbb{R}^{d^2+d}$ and so it can be considered as a metric space.

Definition 4.3. A $K \subset \mathbb{R}^d$ compact set is *self-similar* if $K = \phi_1(K) \cup \ldots \cup \phi_r(K)$, where $r \geq 2$ and $\phi_1, \ldots, \phi_r$ are contractive similitudes.

A $K \subset \mathbb{R}^d$ compact set is *self-affine* if $K = \phi_1(K) \cup \ldots \cup \phi_r(K)$, where $r \geq 2$ and $\phi_1, \ldots, \phi_r$ are injective affine maps, and there is a norm in which they are all contractions.

By the *n-th generation elementary pieces of* K we mean the sets of the form $(\phi_{i_1} \circ \ldots \circ \phi_{i_n})(K)$, where $n = 0, 1, 2, \ldots$.

We shall use multi-indices. By a *multi-index* we mean a finite sequence of indices; for $I = (i_1, i_2, \ldots, i_n)$ let $\phi_I = \phi_{i_1} \circ \ldots \circ \phi_{i_n}$.

Note that the elementary pieces of K are the sets of the form $\phi_I(K)$. These sets are also self-similar/self-affine; and if h is a similitude / injective affine map then $h(K)$ is also self-similar/self-affine and its elementary pieces are the sets of the form $h(\phi_I(K))$.

Definition 4.4. Let $K = \phi_1(K) \cup \ldots \cup \phi_r(K)$ be a self-similar/self-affine set, and let $p_1 + \ldots + p_r = 1$, $p_i > 0$ for all i. Consider the symbol space $\Omega = \{1, \ldots, r\}^{\mathbb{N}}$ equipped with the product topology and let ν be the Borel measure on Ω which is the countable infinite product of the discrete probability measure $p(\{i\}) = p_i$ on $\{1, \ldots, r\}$. Let

$$\pi : \Omega \to K, \quad \{\pi(i_1, i_2, \ldots)\} = \cap_{n=1}^{\infty} (\phi_{i_1} \circ \ldots \circ \phi_{i_n})(K)$$

be the continuous addressing map of K. Let μ be the image measure of ν under the projection π; that is,

$$\mu(H) = \nu\big(\pi^{-1}(H)\big) \quad \text{for every Borel set } H \subset K. \tag{4.1}$$

Such a μ is called a *self-similar/self-affine measure* on K.

Definition 4.5. A self-similar/self-affine set $K = \phi_1(K) \cup \ldots \cup \phi_r(K)$ (or more precisely, the collection $\phi_1, \ldots, \phi_r$ of the representing maps) satisfies the

- *strong separation condition* (SSC) if the union $\phi_1(K) \cup^* \ldots \cup^* \phi_r(K)$ is disjoint;
- *open set condition* (OSC) if there exists a nonempty bounded open set $U \subset \mathbb{R}^d$ such that $\phi_1(U) \cup^* \ldots \cup^* \phi_r(U) \subset U$;
- *convex open set condition* (COSC) if there exists a nonempty bounded open convex set $U \subset \mathbb{R}^d$ such that $\phi_1(U) \cup^* \ldots \cup^* \phi_r(U) \subset U$.

4.1 Self-affine and self-similar sets with the strong separation condition.

Our first nonstability result states that small affine perturbations of a self-affine set K with the strong separation property cannot intersect a very large part of K:

Theorem 4.6. *[Suppl-7, Theorem 3.2] Let $K = \phi_1(K) \cup^* \ldots \cup^* \phi_r(K)$ be a self-affine set satisfying the strong separation condition and let μ be a self-affine measure on K. Then there exists a $c < 1$ and an open neighborhood U of the identity map in the space of injective affine maps from the affine span of K into itself such that $g \in U \setminus \{identity\} \Longrightarrow \mu\big(K \cap g(K)\big) < c$.*

Using Theorem 4.6 we can prove that an isometric but nonidentical copy of K cannot intersect a very large part of K:

Theorem 4.7. *[Suppl-7, Theorem 3.5] Let $K \subset \mathbb{R}^d$ be a self-affine set with the strong separation condition and let μ be a self-affine measure on K. Then there exists a constant $c < 1$ such that for any isometry g we have $\mu\big(K \cap g(K)\big) < c$ unless $g(K) = K$.*

One of our main goals is proving results of the type $\mu(g(K) \cap K) > 0 \Longleftrightarrow \mathrm{int}_K(g(K) \cap K) \neq \emptyset$. One possibility is combining the above type of result with some kind of density theorem, which says that for a positive measure subset of K there exists an elementary piece of K in which its density is very close to 1. This elementary piece is a similar/ affine copy of K, so for such an application we would need to allow similitudes / affine maps in Theorem 4.7. We cannot prove this in the self-affine case, but we could in the self-similar case:

Theorem 4.8. *[Suppl-7, Theorem 4.1] Let $K = \phi_1(K) \cup^* \ldots \cup^* \phi_r(K)$ be a self-similar set satisfying the strong separation condition and μ be a self-similar measure on it. There exists $c < 1$ such that for every similitude g either $\mu\big(g(K) \cap K\big) < c$ or $K \subset g(K)$.*

Then, in the way explained above, we can prove the following, which is the main result of this section:

Theorem 4.9. *[Suppl-7, Theorem 4.5] Let $K = \phi_1(K) \cup^* \ldots \cup^* \phi_r(K)$ be a self-similar set satisfying the strong separation condition, μ be a self-similar measure on it, and g be a similitude. Then $\mu\big(g(K) \cap K\big) > 0$ if and only if the interior (in K) of $g(K) \cap K$ is nonempty. Moreover, $\mu\big(int_K(g(K) \cap K)\big) = \mu\big(g(K) \cap K\big)$.*

As an immediate consequence we get the following.

Corollary 4.10. *[Suppl-7, Corollary 4.6] Let $K \subset \mathbb{R}^d$ be a self-similar set satisfying the strong separation condition, and let μ_1 and μ_2 be self-similar measures on K. Then for any similitude g of $\mathbb{R}^d$,*

$$\mu_1\big(g(K) \cap K\big) > 0 \Longleftrightarrow \mu_2\big(g(K) \cap K\big) > 0.$$

We also get the following fairly easily.

Corollary 4.11. *[Suppl-7, Corollary 4.7] Let $K \subset \mathbb{R}^d$ be a self-similar set satisfying the strong separation condition, let A_K be the affine span of K and let μ be a self-similar measure on K. Then the set of those similitudes $g : A_K \to \mathbb{R}^d$ for which $\mu\big(g(K) \cap K\big) > 0$ is countably infinite.*

4.2 A lemma about invariant extension of measures

The following simple lemma can be interesting in itself. It says that a measure on a set can be always extended as an invariant measure to the whole $\mathbb{R}^n$ unless there is a clear obstacle inside the set. It also holds in a more abstract setting (see [Suppl-7, Lemma 2.17]) but we need only the following special case.

Lemma 4.12. *[Suppl-7, Lemma 2.18] Let μ be a Borel measure on a Borel set $A \subset \mathbb{R}^n$ and let G be a group of affine transformations of $\mathbb{R}^n$. Suppose that*

$$\mu\big(g(B)\big) = \mu(B) \text{ whenever } g \in G,\ B, g(B) \subset A \text{ and } B \text{ is a Borel set.} \tag{4.2}$$

Then there exists a G-invariant Borel measure $\tilde{\mu}$ on $\mathbb{R}^n$ such that $\tilde{\mu}(B) = \mu(B)$ for any $B \subset A$ Borel set.

Remark 4.13. The extension we get in the proof does not always give the measure we expect – it may be infinity for too many sets. For example, if $A \subset \mathbb{R}$ is a Borel set of first category with positive Lebesgue measure, G is the group of translations and μ is the restriction of the Lebesgue measure to A then the Lebesgue measure itself would be the natural translation invariant extension of μ, however the extension $\tilde{\mu}$ as defined in the proof is infinity for every Borel set of second category. This also shows that the extension is far from being unique.

As an illustration of Lemma 4.12 we mention the following special case with a peculiar consequence. (Recall that a Borel measure is said to be *continuous* if the measure of any singleton is zero.)

Lemma 4.14. *[Suppl-7, Lemma 2.21] Let $A \subset \mathbb{R}^n$ ($n \in \mathbb{N}$) be a Borel set such that $A \cap (A + t)$ is at most countable for any $t \in \mathbb{R}^n \setminus \{0\}$. Then any continuous Borel measure μ on A can be extended to a translation invariant Borel measure on $\mathbb{R}^n$.* □

Note that although the condition that $A \cap (A + t)$ is at most countable for any $t \in \mathbb{R}^n \setminus \{0\}$ seems to imply that A is very small, such a set can be still fairly large. Recall that by Theorem 1.1 there exists a compact set $C \subset \mathbb{R}$ with Hausdorff dimension 1 such that $C \cap (C + t)$ contains at most one point for any $t \in \mathbb{R} \setminus \{0\}$. Combining this with Lemma 4.14 we get the following.

Corollary 4.15. *[Suppl-7, Corollary 2.22] There exists a compact set $C \subset \mathbb{R}$ with Hausdorff dimension 1 such that any continuous Borel measure μ on C can be extended to a translation invariant Borel measure on $\mathbb{R}$.*

Lemma 4.12 also guarantees that the following definition makes sense since (by Lemma 4.12) exactly the isometry invariant measures on K can be extended to isometry invariant measures on $\mathbb{R}^n$ in the usual sense.

Definition 4.16. [Suppl-7, Definition 2.20] Let μ be a Borel measure on a compact set K. We say that μ is *isometry invariant* if given any isometry g and a Borel set $B \subset K$ such that $g(B) \subset K$, then $\mu(B) = \mu\big(g(B)\big)$.

An application of the results of the previous section is the following characterization of isometry invariant measures on a self-similar set with the strong separation condition.

Theorem 4.17. *[Suppl-7, Theorem 5.3] Let $K = \phi_1(K) \cup^* \ldots \cup^* \phi_r(K)$ be a self-similar set with the strong separation condition and μ a self-similar measure on K for which congruent elementary pieces are of equal measure. Then μ is an isometry invariant measure on K.*

Remark 4.18. Using this theorem it is relatively easy to decide whether a self-similar measure is isometry invariant or not. Denote the similarity ratio of the similitude ϕ_i by α_i. It is clear that two elementary pieces are congruent if and only if they are images of K by similitudes of equal similarity ratio. Thus a self-similar measure μ is isometry invariant if and only if the equality $p_{i_1}p_{i_2}\ldots p_{i_n} = p_{j_1}p_{j_2}\ldots p_{j_m}$ holds (for the weights of the measure μ) whenever $\alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_n} = \alpha_{j_1}\alpha_{j_2}\ldots\alpha_{j_m}$. By switching from the similarity ratios α_i and weights p_i to the negative of their logarithm we get a system of linear equations for the variables $-\log p_i$. The solutions of this system and the additional equation $\sum_i p_i = 1$ give those weight vectors which define isometry invariant measures on K.

For example, it is easy to see that if the positive numbers $-\log \alpha_i$ ($i = 1, \ldots, r$) are linearly independent over $\mathbb{Q}$, then every self-similar measure is isometry invariant.

An easy consequence of Theorem 4.17 is the following.

Corollary 4.19. *[Suppl-7, Corollary 5.8] Let $K = \phi_1(K) \cup^* \ldots \cup^* \phi_r(K)$ be a self-similar set with strong separation condition, μ be a self-similar measure on K. Then if μ is invariant under orientation preserving isometries, then it is invariant under all isometries.*

4.3 Intersection of translates of a self-affine Sierpiński sponge

In the results of the previous sections we needed the strong separation condition and some of the theorems were only about self-similar sets. In this chapter we will study a class of self-affine sets that also include sets without the strong separation condition.

Take the unit cube $[0,1]^n$ in $\mathbb{R}^n$ and subdivide it into $m_1 \times \ldots \times m_n$ boxes of the same size ($m_1, \ldots, m_n \geq 2$) and cut out some of them. Then do the same with the remaining boxes using the same pattern as in the first step and so on. What remains after infinitely many steps is a self-affine set, which is called *self-affine Sierpiński sponge*. The more precise definition is the following.

Definition 4.20. By *self-affine Sierpiński sponge* we mean self-affine sets of the following type. Let $n, r \in \mathbb{N}$, $m_1, m_2, \ldots, m_n \geq 2$ integers, M be the linear transformation given by the diagonal $n \times n$ matrix

$$M = \begin{pmatrix} m_1 & & 0 \\ & \ddots & \\ 0 & & m_n \end{pmatrix},$$

and let

$$D = \{d_1, \dots, d_r\} \subset \{0, 1, \dots, m_1 - 1\} \times \dots \times \{0, 1, \dots, m_n - 1\}$$

be given. Let $\phi_j(x) = M^{-1}(x + d_j)$ $(j = 1, \dots, r)$. Then the self-affine set $K(M, D) = K = \phi_1(K) \cup \dots \cup \phi_r(K)$ is a Sierpiński sponge.

By the *natural* probability measure on a self-affine sponge $K = K(M, D)$ we shall mean the self-affine measure on K obtained by using equal weights $p_j = \frac{1}{r}$ $(j = 1, \dots, r)$.

For $n = 2$ these sets were studied in several papers (in which they were called self-affine carpets or self-affine carpets of Bedford and McMullen). Bedford [Be84] and McMullen [Mu84] determined the Hausdorff and Minkowski dimensions of these self-affine carpets. (The Hausdorff and Minkowski dimension of self-affine Sierpiński sponges was determined by Kenyon and Peres [KP96]). Gatzouras and Lalley [GL92] proved that except in some relatively simple cases such a set has zero or infinity Hausdorff measure in its dimension (and so in any dimension). Peres extended their results by proving that (except in the same rare simple cases) for any gauge function neither the Hausdorff [Pe94H] nor the packing [Pe94P] measure of a self-affine carpet can be positive and finite (in fact, the packing measure cannot be σ-finite either), and remarked that these results extend to self-affine Sierpiński sponges of higher dimensions.

With Márton Elekes we showed [EK06] that some nice sets – among others the set of Liouville numbers – have zero or non-σ-finite Hausdorff and packing measure for any gauge function by proving that these sets have zero or non-σ-finite measure for any translation invariant Borel measure. (Much earlier Davies [Da71] constructed a compact subset of $\mathbb{R}$ with this property.) So it was natural to ask whether the self-affine carpets of Bedford and McMullen have this stronger property.

We tried to get a translation invariant Borel measure for a self-affine sponge by extending the natural self-affine measure μ on it to a translation invariant measure. As Lemma 4.12 shows, it is enough to check the translation invariance inside the self-affine sponge K. Note that if $B, B+t \subset K$ then $B \subset K \cap (K-t)$ and $B + t \subset K \cap (K+t)$, so we have $\mu(B) = 0 = \mu(B+t)$ unless

$$\mu\big(K \cap (K+t)\big) > 0 \quad \text{or} \quad \mu\big(K \cap (K-t)\big) > 0. \tag{4.3}$$

Therefore, if we could prove that the translate of a self-affine sponge intersects in itself in a set of measure zero unless the intersection is very simple then by Lemma 4.12 we would be able to extend μ to a translation invariant measure to $\mathbb{R}^n$.

This was our original motivation for studying when the intersection of a self-affine sponge and its translate can have positive measure.

The following structure theorem is the key result of this section. It states that we can have positive measure intersection indeed only in exceptional cases.

Theorem 4.21. *[Suppl-7, Theorem 7.4] Let μ be the natural probability measure on a self-affine Sierpiński sponge $K = K(M, D) \subset \mathbb{R}^n$ (as described in Definition 4.20) and let $t \in \mathbb{R}^n$.*

Then $\mu\big(K \cap (K+t)\big) = 0$ holds except in the following two trivial exceptional cases:

(i) There exists two elementary pieces S_1 and S_2 of K such that $S_2 = S_1 + t$.
(ii) K is of the form $K = L \times K_0$, where L is a diagonal of a cube $[0, 1]^l$, where $l \in \{1, 2, \ldots, n\}$ and K_0 is a smaller dimensional self-affine Sierpiński sponge.

Then, as we indicated above, an easy application of Theorem 4.21 and Lemma 4.12 gives the following.

Theorem 4.22. *[Suppl-7, Theorem 8.1] For any self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ ($n \in \mathbb{N}$) there exists a translation invariant Borel measure ν on $\mathbb{R}^n$ such that $\nu(K) = 1$.*

Actually, later we found a more direct proof for the above theorem, which does not use Theorem 4.21 and which works for a slightly larger class of self-affine sets. This more general result is the following:

Theorem 4.23. *[Suppl-7, Theorem 8.2] Let ϕ be an injective affine map which is contractive (in some norm), $t_1, \ldots, t_r \in \mathbb{R}^n$ and $K \subset \mathbb{R}^n$ the compact self-affine set such that $K = \cup_{i=1}^r \phi(K) + t_i$. Suppose that the natural probability measure on K has the property that*

$$\mu\Big(K \cap \Big(\big((\phi(K)+t_i) \cap (\phi(K)+t_j)\big)+u\Big)\Big) = 0 \quad (\forall\, 1 \le i < j \le r,\ u \in \mathbb{R}^n). \tag{4.4}$$

(a) Then for any $t \in \mathbb{R}^n$ and elementary piece S of K we have

$$\mu\big(K \cap (S + t)\big) \le \mu(S).$$

(b) There exists a translation invariant Borel measure ν on $\mathbb{R}^n$ such that $\nu(K) = 1$. In fact, ν is an extension of μ.

After showing that the convex open set condition implies (4.4) we get the following nicer special case, which is still more general than Theorem 4.22, which we originally proved via the much deeper structure theorem Theorem 4.21.

Corollary 4.24. *[Suppl-7, Corollary 8.3] Let $K = \phi_1(K) \cup \ldots \cup \phi_r(K)$ be a self-affine set with the convex open set condition and suppose that $\phi_1(K), \ldots, \phi_r(K)$ are translates of each other.*

Then the natural probability measure on K can be extended as a translation invariant measure on $\mathbb{R}^n$.

So, although the motivation of Theorem 4.21 was proving Theorem 4.22, the following applications give deeper and probably more important results.

From Theorem 4.21 it is fairly easy to prove the following result, which is analogous to Theorem 4.9.

Corollary 4.25. *[Suppl-7, Corollary 7.7] Let μ be the natural probability measure on a self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ (as described in Definition 4.20) and let $t \in \mathbb{R}^n$.*

The set $K \cap (K + t)$ has positive μ-measure if and only if it has non-empty interior (relative) in K.

The analogous statement to the stability result Theorem 4.7 does not hold for all self-affine sponges. The sponges that were exceptional in Theorem 4.21 are counter-examples: indeed, if K is of the form $K = L \times K_0$, where L is a diagonal of a cube $[0,1]^l$, where $l \in \{1,2,\dots,n\}$ and K_0 is a smaller dimensional self-affine Sierpiński sponge, then a small translation in the direction of L clearly gives intersection with measure arbitrary close to 1. The following result, which is also an application of Theorem 4.21 says that there are no other counter-examples:

Theorem 4.26. *[Suppl-7, Theorem 7.9] Let μ be the natural probability measure on a self-affine Sierpiński sponge $K = K(M,D) \subset \mathbb{R}^n$ (as described in Definition 4.20) and let $t \in \mathbb{R}^n$.*

Then $\mu\big(K \cap (K+t)\big) \le 1 - \frac{1}{r^2}$ holds (where r denotes the number of elements in the pattern D) except in the following two trivial exceptional cases:
(i) $t = 0$.
(ii) K is of the form $K = L \times K_0$, where L is a diagonal of a cube $[0,1]^l$, where $l \in \{1,2,\dots,n\}$ and K_0 is a smaller dimensional self-affine Sierpiński sponge.

Chapter 5

Periodic decomposition of measurable integer valued functions

One can also study the relation of the additive and the measure structure of $\mathbb{R}$ via studying periodic (Lebesgue) measurable sets, or more generally periodic (Lebesgue) measurable functions. The most basic result in this direction is the well known fact that if a measurable set $H \subset \mathbb{R}$ has arbitrary small periods then either H or its complement must have zero Lebesgue measure. More generally, if a measurable function $\mathbb{R} \to \mathbb{R}$ is periodic with respect to arbitrary small periods then it is almost everywhere constant. This clearly implies that if a measurable function has two incommensurable periods (which simply means that the ratio of the periods is irrational) then it must be almost everywhere constant. The picture becomes more interesting if we study sums of periodic measurable functions with given periods. If we want to be closer to the study of measurable sets then we may require that the functions are characteristic functions, or at least integer valued functions.

Definition 5.1. Let $a_1, \dots, a_k \in \mathbb{R} \setminus \{0\}$ be fixed periods.

We say that a function $f : \mathbb{R} \to \mathbb{R}$ has an $(a_1, \dots, a_k)$-*decomposition* if it can be written as $f = f_1 + \ldots + f_k$, where f_j is a_j-periodic for each $j = 1, \dots, k$.

This is called an *integer valued / measurable / bounded /... decomposition* if each f_j is integer valued / measurable / bounded /... .

The characterization of those functions that has a periodic decomposition with given periods has a long history. It started in the seventies with some unpublished work of I. Z. Ruzsa. If f has an $(a_1, \dots, a_n)$-periodic decomposition of f then

$$\Delta_{a_1} \Delta_{a_2} \cdots \Delta_{a_n} f = 0, \quad \text{where} \quad \Delta_{a_j} f(x) = f(x + a_j) - f(x), \tag{5.1}$$

since the difference operators Δ_{a_j} commute.

Definition 5.2. A class of functions $\mathcal{F}$ is said to have the *decomposition property* if every $f \in \mathcal{F}$ that satisfies (5.1) has an $(a_1, \dots, a_k)$-periodic decomposition in $\mathcal{F}$.

Since for the identity function $f(x) = x$ we clearly have $\Delta_1 \Delta_1 f = 0$ but it is not the sum of two 1-periodic functions, many natural classes of functions (e.g. all $\mathbb{R} \to \mathbb{R}$ functions, continuous $\mathbb{R} \to \mathbb{R}$ functions) do not have the decomposition property. However, many classes of functions do have the decomposition property: for example the class of all bounded continuous $\mathbb{R} \to \mathbb{R}$ functions [LR89], the class of all bounded $\mathbb{R} \to \mathbb{R}$ functions [Ga92], [LR90], the class of all bounded measurable $\mathbb{R} \to \mathbb{R}$ functions [LR90] and the class of all bounded real valued functions on an arbitrary Abelian group [LR90].

For integer valued functions it was proved in [KKKR] that the class of bounded $\mathbb{Z} \to \mathbb{Z}$ functions has the decomposition property but the class of bounded $\mathbb{R} \to \mathbb{Z}$ functions does not have the decomposition property.

As we described above, in this chapter we focus on integer valued (or at least almost everywhere integer valued) functions.

One may hope that the existence of a *real* valued $(a_1, \dots, a_k)$-decomposition of an $\mathbb{R} \to \mathbb{Z}$ function implies the existence of an *integer* valued (or at least almost everywhere integer valued) measurable $(a_1, \dots, a_k)$-decomposition. However, the following result shows that this is not the case even for $k = 3$.

Theorem 5.3. *[Suppl-8, Theorem 1.2] There exists an integer valued bounded Lebesgue measurable function on the real line that can be written as a sum of three real valued bounded measurable periodic functions but cannot be written as a sum of three almost everywhere integer valued measurable periodic functions with the same periods.*

Since the existence of the real valued $(a_1, \dots, a_k)$-decomposition already implies (5.1) we get the following:

Corollary 5.4. *[Suppl-8, Corollary 1.3] The following classes of functions do not have the decomposition property:* $\{f : \mathbb{R} \to \mathbb{Z} \mid f \text{ is bounded and measurable}\}$, $\{f : \mathbb{R} \to \mathbb{R} \mid f \in L_\infty \text{ and } f \text{is almost everywhere integer valued}\}$, $\{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is bounded, measurable and almost everywhere integer valued}\}$, *and* $\{f : \mathbb{R} \to \mathbb{Z} \mid f \in L_\infty\}$.

The construction of Theorem 5.3 is fairly simple. Let $t \in \mathbb{R} \setminus \mathbb{Q}$ be arbitrary and let

$$f(x) = \{tx\} + \{(1-t)x\} + \{-x\},$$

where $\{.\}$ denotes the fractional part; that is, $\{a\} = a - [a]$. Then f is clearly measurable and it is the sum of a $\frac{1}{t}$-periodic a $\frac{1}{1-t}$-periodic and a 1-periodic bounded measurable function. Noting that f can be also written as

$$f(x) = tx - [tx] + (1-t)x - [(1-t)x] - x - [-x] = -[tx] - [(1-t)x] - [-x]$$

we get that f is integer valued.

The claim that f cannot be written as a sum of three almost everywhere integer valued measurable periodic functions with the same periods easily follows from the following known fact (see e.g. in [Ke97]).

Lemma 5.5. *Let $a_1, \dots, a_k \in \mathbb{R} \setminus \{0\}$ such that $a_i/a_j \notin \mathbb{Q}$ for any $i \neq j$ and suppose that $f_1 + \ldots + f_k = g_1 + \ldots + g_k$ and for each j, f_j and g_j are a_j-periodic measurable $\mathbb{R} \to \mathbb{R}$ functions.*

Then $f_j - g_j$ is almost everywhere constant for every $j = 1, \dots, k$.

It is easy to see that in the above lemma the condition $a_i/a_j \notin \mathbb{Q}$ for $i \neq j$ is also necessary. We will see (Corollary 5.8) that for $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ functions the necessary and sufficient condition for $a_1, \dots, a_k$ is stronger.

Our main goal is characterizing those k-tuples of periods for which the existence of a real valued measurable $(a_1, \dots, a_k)$-decomposition of an $\mathbb{R} \to \mathbb{Z}$ function implies the existence of an integer valued measurable $(a_1, \dots, a_k)$-decomposition.

First we want to get almost integer valued measurable decomposition.

Theorem 5.6. *[Suppl-8, Theorem 2.5] For any $a_1, \dots, a_k \in \mathbb{R}\backslash\{0\}$ the following seven statements are equivalent.*

$(i)/(i')$ *If an everywhere/almost everywhere integer valued measurable function f has a measurable real valued $(a_1, \dots, a_k)$-decomposition then f has a measurable almost everywhere integer valued $(a_1, \dots, a_k)$-decomposition.*

$(ii)/(ii')$ *If an everywhere/almost everywhere integer valued measurable function f has a bounded measurable real valued $(a_1, \dots, a_k)$-periodic decomposition then it also has a bounded measurable almost everywhere integer valued $(a_1, \dots, a_k)$-periodic decomposition.*

$(iii)/(iii')$ *An everywhere/almost everywhere integer valued bounded measurable function f has a bounded measurable almost everywhere integer valued $(a_1, \dots, a_k)$-decomposition if and only if $\Delta_{a_1} \dots \Delta_{a_k} f = 0$, where Δ_a denotes the difference operator defined as $\Delta_a f(x) = f(x + a) - f(x)$.*

(iv) *If $B_1, \dots, B_n$ are the equivalence classes of $\{a_1, \dots, a_k\}$ with respect to the relation $a \sim b \Leftrightarrow a/b \in \mathbb{Q}$, and b_j denotes the smallest common multiple of the numbers in B_j (for each $j = 1, \dots, n$) then $\frac{1}{b_1}, \dots, \frac{1}{b_n}$ are linearly independent over $\mathbb{Q}$.*

A few words about the proof:

The arguments proving that $(i), (i'), (ii), (ii')$ imply (iv) are similar to the proof of Theorem 5.3.

The implications $(ii) \Leftrightarrow (iii)$ and $(ii)' \Leftrightarrow (iii)'$ follow from the already mentioned Theorem of Laczkovich and Révész [LR90], which states that the class of bounded measurable functions has the decomposition property; that is, a bounded measurable function $f : \mathbb{R} \to \mathbb{R}$ has a bounded measurable real valued $(a_1, \dots, a_k)$-decomposition if and only if $\Delta_{a_1} \dots \Delta_{a_k} f = 0$.

The proofs for the sufficiency of condition (iv) are based on the following result, which might be also interesting in its own right.

Theorem 5.7. *[Suppl-8, Theorem 2.3] Let $a_1, \dots, a_k \in \mathbb{R} \setminus \{0\}$ such that $\frac{1}{a_1}, \dots, \frac{1}{a_k}$ are linearly independent over $\mathbb{Q}$. Suppose that $f_j : \mathbb{R} \to \mathbb{R}$ is an a_j-periodic measurable function for each $j = 1, \dots, k$ and that $f = f_1 + \ldots + f_k$ is an almost everywhere integer valued function.*

Then each fractional part $\{f_j\}$ is constant almost everywhere.

As a spin-off, using this result, we can easily characterize those k-tuples of periods for which the measurable decomposition of an $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ function is almost everywhere unique up to additive constants. Note that, by Lemma 5.5, the characterization is different for $\mathbb{R} \to \mathbb{R}$ functions.

Corollary 5.8. *[Suppl-8, Corollary 2.4] For any $a_1, \ldots, a_k \in \mathbb{R} \setminus \{0\}$ the following two statements are equivalent.*

(i) If $f_1 + \ldots + f_k = g_1 + \ldots + g_k$ and for each j, f_j and g_j are a_j-periodic measurable $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ functions then $f_j - g_j$ is almost everywhere constant for every $j = 1, \ldots, k$.

(ii) $\frac{1}{a_1}, \ldots, \frac{1}{a_k}$ are linearly independent over $\mathbb{Q}$.

Remark 5.9. If we want to get everywhere integer valued measurable decomposition of a given integer valued function, which has a real valued measurable decomposition we may try the following two steps:

Step 1. first we make an almost integer valued decomposition using Theorem 5.6,

Step 2. then we try to make an arbitrary integer valued decomposition on the exceptional nullset.

This means that for getting the analogous results for everywhere integer valued measurable functions we need to understand the analogous nonmeasurable cases, which were studied in [KKKR], [FHKR], [Ha07] and [Ha09]. We mention only the results we need here.

By a result of Károlyi, Keleti, Kós and Ruzsa [KKKR], the existence of a bounded real valued decomposition of an integer valued function does *not* always imply the existence of a bounded integer valued decomposition. Thus Step 2 above does not work for arbitrary periods, which leads to the following result:

Proposition 5.10. *[Suppl-8, Proposition 3.4] There exists $a_1, a_2, a_3 \in \mathbb{R}$ such that $\frac{1}{a_1}$, $\frac{1}{a_2}$ and $\frac{1}{a_3}$ are linearly independent over $\mathbb{Q}$ and a function $f : \mathbb{R} \to \{0, 1\}$ that has a bounded measurable real valued (a_1, a_2, a_3)-periodic decomposition but does not have a bounded measurable integer valued (a_1, a_2, a_3)-periodic decomposition.*

Consequently one cannot replace "almost everywhere integer valued" by "integer valued" in (ii) and (iii) of Theorem 5.6.

For getting a characterization of those periods for which the everywhere integer valued versions of (ii) and (iii) of Theorem 5.6 hold, first a characterization of those periods was needed for which the existence of a bounded real valued decomposition of an integer valued function implies the existence of a bounded integer valued decomposition. Viktor Harangi [Ha09] (see also [Ha07]) proved that this latter statement holds if and only if whenever $\frac{a_l}{a_m}, \frac{a_m}{a_n}, \frac{a_n}{a_l} \notin \mathbb{Q}$ (for some $l, m, n \in \{1, \ldots, k\}$) then a_l, a_m and a_n are linearly independent over $\mathbb{Q}$. Then, combining these results, he could also characterize those k-tuples of periods for which the everywhere integer valued versions of (ii) and (iii) of Theorem 5.6 holds.

For getting the everywhere integer valued analogous to (i) of Theorem 5.6, Remark 5.9 were made precise:

Proposition 5.11. *[Suppl-8, Proposition 3.3] The following two questions are equivalent:*

$(*)$ Can one replace "almost everywhere integer valued" by "integer valued" in (i) of Theorem 5.6?

(**) Is it true for any $a_1, \dots, a_k \in \mathbb{R}$ *that if an integer valued function $f : \mathbb{R} \to \mathbb{Z}$ has a real valued $(a_1, \dots, a_k)$-periodic decomposition, then f also has an integer valued $(a_1, \dots, a_k)$-periodic decomposition?*

Later an affirmative answer was given to $(**)$ in [FHKR], and so also to $(*)$. Therefore we get the following characterizations of the good k-tuples of periods:

Theorem 5.12. *[FHKR, Corollary 4.3] The following two assertions are equivalent for any $a_1, \dots, a_k \in \mathbb{R} \setminus \{0\}$.*

(i'') *If an integer valued function has a measurable real valued $(a_1, \dots, a_k)$-decomposition then it also has a measurable integer valued $(a_1, \dots, a_k)$-decomposition.*

(iv) *If $B_1, \dots, B_n$ are the equivalence classes of $\{a_1, \dots, a_k\}$ with respect to the relation $a \sim b \Leftrightarrow a/b \in \mathbb{Q}$, and b_j denotes the smallest common multiple of the numbers in B_j (for each $j = 1, \dots, n$) then $\frac{1}{b_1}, \dots, \frac{1}{b_n}$ are linearly independent over $\mathbb{Q}$.*

Bibliography

[Suppl-1] T. Keleti, A 1-dimensional subset of the reals that intersects each of its translates in at most a single point, *Real Anal. Exchange* **24** (1998/99), 843–844.

[Suppl-2] T. Keleti, Construction of one-dimensional subsets of the reals not containing similar copies of given patterns, *Analysis & PDE* **1** (2008), 29–33.

[Suppl-3] U.B. Darji and T. Keleti, Covering $\mathbb{R}$ with translates of a compact set, *Proc. Amer. Math. Soc.* **131** (2003), no. 8, 2598–2596.

[Suppl-4] M. Abért and T. Keleti, Shuffle the plane, *Proc. Amer. Math. Soc.* **130** (2002), 549–553.

[Suppl-5] T. Keleti T, Density and covering properties of intervals of $\mathbb{R}^n$, *Mathematika* **47** (2000), 229-242.

[Suppl-6] T. Keleti, A covering property of some classes of sets in $\mathbb{R}^n$, *Acta Univ. Carol., Math. Phys.* **39** (1998), 111-118.

[Suppl-7] M. Elekes, T. Keleti, A. Máthé, Self-similar and self-affine sets; measure of the intersection of two copies, *Ergodic Theory Dynam. Systems*, to appear.

[Suppl-8] T. Keleti, Periodic decomposition of measurable integer valued functions, *J. Math. Anal. Appl.* **337** (2008), 1394-1403.

[Ab02] M. Abért, Symmetric groups as products of abelian subgroups, *Bull. London Math. Soc.* **34** (2002), no. 4, 451–456.

[Be84] T. Bedford, *Crinkly curves, Markov partitions and box dimension in self-similar sets*, Ph. D. Thesis, University of Warwick, 1984.

[BK06] A. Bisbas, M.N. Kolountzakis, Avoiding affine copies of infinite sequences, unpublished manuscript, 2006.

[Bo87] J. Bourgain, Construction of sets of positive measure not containing an affine image of a given infinite structure, *Israel J. Math.* **60** (1987), 3, 333–344.

[CCW] A. Carbery, M. Christ and J. Wright, Multidimensional van der Corput and sublevel set estimates, *J. Amer. Math. Soc.* **12** (1999), no. 4, 981-1015.

[CF75] A. Córdoba, R. Fefferman, A geometric proof for the strong maximal theorem, *Ann. of Math.* **102** (1975), 95-100.

[CN95] D. Cruz-Uribe, SFO and C. J. Neugebauer, The structure of the reverse Hölder classes, *Trans. Amer. Math. Soc.* **347** (1995), 2941-2960.

[Da71] R. O. Davies, Sets which are null or non-sigma-finite for every translation-invariant measure, *Mathematika* **18** (1971), 161–162.

[EK06] M. Elekes and T. Keleti, Borel sets which are null or non-σ-finite for every translation invariant measure, *Adv. Math.* **201** (2006), 102-115.

[ES04] M. Elekes and J. Steprāns, Less than 2^ω many translates of a compact nullset may cover the real line, *Fund. Math.* **181** (2004), 89–96.

[Er64] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183-190.

[Er74] P. Erdős, Remarks on some problems in number theory, *Math. Balkanica* **4** (1974), 197-202.

[EK57] P. Erdős and S. Kakutani, On a perfect set, *Colloq. Math.* **4** (1957), 195–196.

[Fa84] K.J. Falconer, On a problem of Erdős on sequences and measurable sets, *Proc. Amer. Math. Soc.* **90** (1984), 77–78.

[Fa85] K. J. Falconer, Classes of sets with large intersection, *Mathematika* **32** (1985), no. 2, 191–205.

[Fa90] K.J. Falconer, *Fractal Geometry*, John Wiley & Sons, 1990.

[FHKR] B. Farkas, V. Harangi T. Keleti, Sz. Gy. Révész, Invariant decomposition of functions with respect to commuting invertible transformations, *Proc. Amer. Math Soc.* **136** (2008), 1325-1336.

[FW09] D. Feng and Y. Wang, On the structures of generating iterated function systems of Cantor sets, *Adv. Math.* **222** (2009), 1964-1981.

[Fu70] H. Furstenberg, Intersections of Cantor sets and transversality of semigroups, *Problems in analysis (Sympos. Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*, pp. 41-59, Princeton Univ. Press, 1970.

[GL92] D. Gatzouras and S. Lalley (1992), Hausdorff and box dimensions of certain self-affine fractals, *Indiana University Math. J.* **41** (1992), 533–568.

[Ga92] Z. Gajda, Note on decomposition of bounded functions into the sum of periodic terms, *Acta Math. Hungar.* **59** (1992), no. 1-2, 103–106.

[Gu74] M. de Guzmán, An inequality for the Hardy-Littlewood maximal operator with respect to a product of differentiation bases, *Studia Math.* **49** (1974), 185-194.

[Gu75] M. de Guzmán, *Differentiation of Integrals in* $\mathbb{R}^n$, Springer, Lecture Notes in Mathematics Vol. 481, Berlin, 1975.

[Gu81] M. de Guzmán, *Real Variable Methods in Fourier Analysis*, North-Holland, Mathematics Studies Vol. 46, Amsterdam, 1981.

[Gy09] Z. Gyenes, The ratio of the perimeter and area of unions of copies of a fixed set, *Discrete Comput. Geom.*, to appear.

[Ha09] V. Harangi, *On the uniqueness of periodic decomposition*, submitted.

[Ha07] V. Harangi, Periodic decomposition of functions, *Real Anal. Exchange* 2007, 31st Summer Symposium Conference, 185-188.

[HL98] P.D Humke and M. Laczkovich, A visit to the Erdős problem, *Proc. Amer. Math. Soc.* **126** (1998), 3, 819–822.

[Hu81] J. E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747.

[Ig03] K. Igudesman, Lacunary self-similar fractal sets and its application to intersection of Cantor sets, *Lobachevskii J. Math.* **12** (2003), 41–50.

[Ja99] M. Järvenpää, Hausdorff and packing dimensions, intersection measures, and similarities, *Ann. Acad. Sci. Fenn. Math.* **24** (1999), no. 1., 165–186.

[KKKR] Gy. Károlyi, G. Kós, T. Keleti, I. Z. Ruzsa, Periodic decomposition of integer valued functions, *Acta Math. Hung.* **119** (2008), no. 3, 227-242.

[Ke97] T. Keleti On the differences and sums of periodic measurable functions, *Acta Math. Hung.* **75** (4) (1997), 279–286.

[KP96] R. Kenyon and Y. Peres, Measures of full dimension on affine-invariant sets, *Ergodic Theory Dynamical Syst.* **16** (1996), 307–323.

[Ko97] M.N. Kolountzakis, Infinite patterns that can be avoided by measure, *Bull. London Math. Soc.* **29** (1997), 4, 415–424.

[Ko83] P. Komjáth, Large sets not containing images of a given sequence, *Canad. Math. Bull.* **26** (1983), 41–43.

[Ko02] P. Komjáth, Five degrees of separation, *Proc. Amer. Math. Soc.* **130** (2002), 2413–2417.

[Ku06] G. Kun, A surprising covering of the real line, *Proc. Amer. Math. Soc.* **134** (2006), 3555–3559.

[LP09] I. Laba and M. Pramanik, Arithmetic progressions in sets of fractional dimension, *Geom. Funct. Anal.* **19** (2009), 429-456.

[LR89] M. Laczkovich, Sz.Gy. Révész, Periodic decompositions of continuous functions, *Acta Math. Hungar.* **54** (1989), no. 3-4, 329–341.

[LR90] M. Laczkovich, Sz. Gy. Révész, Decompositions into periodic functions belonging to a given Banach space, *Acta Math. Hung.* **55** (3-4) (1990), 353–363.

[LW96] J. C. Lagarias and Y. Wang, Self-affine tiles in R^n, *Adv. Math.* **121** (1996), no. 1, 21–49.

[LX99] W. Li and D. Xiao, Intersection of translations of Cantor triadic set, *Acta Math. Sci. (English Ed.)* **19** (1999), no. 2, 214–219.

[MaP] P. Maga, *Full dimensinal sets without given patterns*, preprint.

[Ma82] P. Mattila, On the structure of self-similar fractals, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **7** (1982), no. 2., 189–195.

[Ma84] P. Mattila, Hausdorff dimensions and capacities of intersections of sets in n-space, *Acta Math.* **152** (1984), no. 1-2, 77–105.

[M85] P. Mattila, On the Hausdorff dimension and capacities of intersections, *Mathematika* **32** (1985), no. 2, 213–217.

[Ma95] P. Mattila, *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability* (Cambridge Studies in Advanced Mathematics, 44), Cambridge University Press, Cambridge, 1995.

[MaA09] A. Máthé, Covering the real line with translates of a zero-dimensional set, *Fund. Math.*, to appear.

[Mu84] C. McMullen, The Hausdorff dimension of general Sierpiński carpets, *Nagoya Math. J.* **96** (1984), 1–9.

[Mo96] C. G. T. de A. Moreira, Stable intersections of Cantor sets and homoclinic bifurcations, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13** (1996), 741–781.

[MY01] C. G. T. de A. Moreira and J-C. Yoccoz, Stable intersections of regular Cantor sets with large Hausdorff dimensions, *Ann. of Math. (2)* **154** (2001), no. 1, 45–96.

[NL02] F. Nekka and J. Li, Intersection of triadic Cantor sets with their translates. I. Fundamental properties, *Chaos Solitons Fractals* **13** (2002), no. 9, 1807–1817.

[PS98] Y. Peres and B. Solomyak, Self-similar measures and intersections of Cantor sets, *Trans. Amer. Math. Soc.* **350** (1998), no. 10, 4065–4087.

[Pe94P] Y. Peres, The packing measure of self-affine carpets, *Math. Proc. Cambridge Philos. Soc.* **115** (1994), no. 3, 437–450.

[Pe94H] Y. Peres, The self-affine carpets of McMullen and Bedford have infinite Hausdorff measure, *Math. Proc. Cambridge Philos. Soc.* **116** (1994), 513–526.

[Re58] I. Reiman, Über ein problem von K. Zarankiewicz, *Acta Math. Acad. Sci. Hung.* **9** (1958), 269-273.

[Sv00] R.E. Svetic, The Erdős similarity problem: a survey, *Real Anal. Exchange* **26** (2000/01), 525–539.

Supplements

Supplement 1: T. Keleti, A 1-dimensional subset of the reals that intersects each of its translates in at most a single point, *Real Anal. Exchange* **24** (1998/99), 843–844.

Supplement 2: T. Keleti, Construction of one-dimensional subsets of the reals not containing similar copies of given patterns, *Analysis & PDE* **1** (2008), 29–33.

Supplement 3: U.B. Darji and T. Keleti, Covering $\mathbb{R}$ with translates of a compact set, *Proc. Amer. Math. Soc.* **131** (2003), no. 8, 2598–2596.

Supplement 4: M. Abért and T. Keleti, Shuffle the plane, *Proc. Amer. Math. Soc.* **130** (2002), 549–553.

Supplement 5: T. Keleti T, Density and covering properties of intervals of $\mathbb{R}^n$, *Mathematika* **47** (2000), 229-242.

Supplement 6: T. Keleti, A covering property of some classes of sets in $\mathbb{R}^n$, *Acta Univ. Carol., Math. Phys.* **39** (1998), 111-118.

Supplement 7: M. Elekes, T. Keleti, A. Máthé, Self-similar and self-affine sets; measure of the intersection of two copies, *Ergodic Theory Dynam. Systems*, to appear.

Supplement 8: T. Keleti, Periodic decomposition of measurable integer valued functions, *J. Math. Anal. Appl.* **337** (2008), 1394-1403.

Tamás Keleti, *Department of Analysis, Eötvös Loránd University, Múzeum krt. 6-8, Budapest, H-1088, Hungary (e-mail: elek@cs.elte.hu)

A 1-dimensional subset of the reals that intersects each of its translates in at most a single point

Abstract

We construct a compact subset of $\mathbf{R}$ with Hausdorff dimension 1 that intersects each of its non-identical translates in at most one point. Moreover, one can make the set to be linearly independent over the rationals.

In 1984 P. Mattila [2] constructed compact subsets A and B of $\mathbf{R}$ with Hausdorff dimension 1 such that the intersection of A and any translate of B contains at most one point. In this note we show that - if we allow only non-identical translations - one can also have $A = B$.

We call a set of 3 or 4 real numbers $x_1 < x_2 \leq x_3 < x_4$ a *rectangle* if $x_2 - x_1 = x_4 - x_3$.

Note that a set intersects each of its translates in at most one point if and only if the set does not contain a rectangle. (Here and in the sequel by set we will always mean a subset of $\mathbf{R}$ and by translate a non-identical translate.)

Theorem 1 *There exists a compact set with Hausdorff dimension 1 that intersects each of its translates in at most one point.*

Proof. Let $\delta_m = 1/(6^{m-1}m!)$. We define inductively compact sets A_m as disjoint unions of the closed intervals $[n_{i_1\ldots i_m}\delta_m, (n_{i_1\ldots i_m}+1)\delta_m]$ for $1 \leq i_k \leq k$, $1 \leq k \leq m$. We will denote by $I_1^m, I_2^m, \ldots, I_{m!}^m$ the intervals of A_m and by $(J_1, J_2, \ldots)$ the sequence $(I_1^1, I_1^2, I_2^2, I_1^3, \ldots, I_{3!}^3, \ldots)$.

Let $n_1 = 0$. (Then $A_1 = I_1^1 = J_1 = [0,1]$.) Assume that $A_1, \ldots, A_m$ have already been defined. If $n_{i_1\ldots i_m}\delta_m \notin J_m$ then let

$$n_{i_1\ldots i_m i} = 6(m+1)n_{i_1\ldots i_m} + 6i - 6 \qquad (i = 1, \ldots, m+1), \tag{1}$$

Key Words: Hausdorff dimension, translation, linearly independent
Mathematical Reviews subject classification: 28A78
*This research was done while the author was visiting the University College London having a Royal Society/NATO Postdoctoral Fellowship award.

and if $n_{i_1\ldots i_m}\delta_m \in J_m$ then let

$$n_{i_1\ldots i_m i} = 6(m+1)n_{i_1\ldots i_m} + 6i - 3 \qquad (i = 1,\ldots,m+1). \tag{2}$$

Thus

$$[n_{i_1\ldots i_m i}\delta_{m+1}, (n_{i_1\ldots i_m i}+1)\delta_{m+1}] \subset [n_{i_1\ldots i_m}\delta_m, (n_{i_1\ldots i_m}+1)\delta_m]$$

for $i = 1,\ldots,m+1$, which means that the intervals of A_{m+1} are contained in the intervals of A_m.

Let $A = \cap_{l=1}^{\infty} A_m$. Then A has Hausdorff dimension 1, cf. [1] Example 4.6. Hence, by our previous remark, it is enough to show that A does not contain a rectangle.

Let $x_1 < x_2 \leq x_3 < x_4$ be points of A. Take an m such that $\delta_m < x_2 - x_1$. Then if $x_1 \in I_j^m = J_M$ then none of x_2, x_3 and x_4 is in I_j^m. Thus, when we defined A_{M+1}, we used (2) for defining the interval that contains x_1 and (1) for defining the intervals that contain x_2, x_3 and x_4. This implies that x_1 is of the form $(6N_1+3)\delta_M + \varepsilon_1$ but x_2, x_3 and x_4 are of the form $6N_j\delta_M + \varepsilon_j$, where $N_1,\ldots,N_4$ are integers and $0 \leq \varepsilon_i \leq \delta_M$ for $i = 1,\ldots,4$. Thus $x_2 - x_1 \neq x_4 - x_3$, which means that (x_1, x_2, x_3, x_4) is not a rectangle. $\square$

Remark 2 Slightly modifying the above construction (by replacing 6 with a slowly increasing sequence of even numbers) one can also get a compact set with Hausdorff dimension 1 which is linearly independent over the rationals. (The existence of a linearly independent perfect set is well known, even in any non-discrete locally compact abelian group, see e. g. [3].)

References

[1] K. Falconer, *Fractal Geometry*, John Wiley & Sons, 1990.

[2] P. Mattila, Hausdorff dimension and capacities of intersections of sets in n-space, *Acta Math.*, **152** (1984), 77-105.

[3] W. Rudin, *Fourier Analysis on Groups*, Interscience Publishers, New York - London, 1962.

ANALYSIS & PDE

Volume 1 No. 1 2008

CONSTRUCTION OF ONE-DIMENSIONAL SUBSETS OF THE REALS NOT CONTAINING SIMILAR COPIES OF GIVEN PATTERNS

TAMÁS KELETI

mathematical sciences publishers

ANALYSIS AND PDE
Vol. 1, No. 1, 2008

CONSTRUCTION OF ONE-DIMENSIONAL SUBSETS OF THE REALS NOT CONTAINING SIMILAR COPIES OF GIVEN PATTERNS

TAMÁS KELETI

For any countable collection of sets of three points we construct a compact subset of the real line with Hausdorff dimension 1 that contains no similar copy of any of the given triplets.

1. Introduction

An old conjecture of Erdős [1974] (also known as the Erdős similarity problem) states that for any infinite set $A \subset \mathbb{R}$ there exists a set $E \subset \mathbb{R}$ of positive Lebesgue measure which does not contain any similar (that is, translated and rescaled) copy of A. It is known that slowly decaying sequences are not counterexamples [Falconer 1984; Bourgain 1987; Kolountzakis 1997] (see for example [Humke and Laczkovich 1998; Komjáth 1983; Svetic 2000] for other related results) but nothing is known about any infinite sequence that converges to zero at least exponentially. On the other hand, it follows easily from Lebesgue's density theorem that any set $E \subset \mathbb{R}$ of positive Lebesgue measure contains similar copies of every finite set.

Bisbas and Kolountzakis [2006] gave an incomplete proof of a related statement: For every infinite set $A \subset \mathbb{R}$ there exists a compact set $E \subset \mathbb{R}$ of Hausdorff dimension 1 such that E contains no similar copy of A. Kolountzakis asked whether the same holds for finite sets as well. Iosevich asked a similar question: if $A \subset \mathbb{R}$ is a finite set and $E \subset [0, 1]$ is a set of given Hausdorff dimension, must E contain a similar copy of A?

In this paper we answer these questions by showing that for any set $A \subset \mathbb{R}$ of at least 3 elements there exists a 1-dimensional set that contains no similar copy of A. In fact, we obtain a bit more by proving the following theorem, which immediately yields the two subsequent corollaries.

Theorem 1.1. *For any countable set $A \subset (1, \infty)$ there exists a compact set $E \subset \mathbb{R}$ with Hausdorff dimension 1 such that if $x < y < z$ and $x, y, z \in E$, then*

$$\frac{z-x}{z-y} \notin A.$$

Corollary 1.2. *For any sequence $B_1, B_2, \ldots \subset \mathbb{R}$ of sets of at least three elements there exists a compact set $E \subset \mathbb{R}$ with Hausdorff dimension 1 that contains no similar copy of any of $B_1, B_2, \ldots$.*

Corollary 1.3. *For any countable set $B \subset \mathbb{R}$ there exists a compact set $E \subset \mathbb{R}$ with Hausdorff dimension 1 that intersects any similar copy of B in at most two points.*

MSC2000: 28A78.
Keywords: Hausdorff dimension, avoiding pattern, Erdős similarity problem, similar copy, affine copy.
Partially supported by OTKA grant 049786.

The method of the construction is similar to the method used in [Keleti 1998], where a compact set A of Hausdorff dimension 1 is constructed such that A does not contain any set of the form

$$\{a, a+b, a+c, a+b+c\}$$

for any $a, b, c \in \mathbb{R}$, $b, c \neq 0$, so in particular A does not contain any nontrivial 3-term arithmetic progression.

Laba and Pramanik [2007] obtained a positive result by proving that if a compact set $E \subset \mathbb{R}$ has Hausdorff dimension sufficiently close to 1 and E supports a probability measure whose Fourier transform has appropriate decay at infinity then E must contain nontrivial 3-term arithmetic progressions. It would be interesting to know whether similar conditions could guarantee other finite patterns as well.

Perhaps one can even find conditions weaker than having positive measure that implies that a compact subset of $\mathbb{R}$ contains similar copies of all finite subsets. This is not impossible since Erdős and Kakutani [1957] constructed a compact set of measure zero with this property. The Erdős–Kakutani set has Hausdorff dimension 1 but, using the ideas from [Elekes and Steprāns 2004], Máthé [≥ 2008] constructed such a set with Hausdorff dimension 0. However, the packing dimension of such a set must be 1, since the argument of the proof of [Darji and Keleti 2003, Theorem 2] gives that if a compact set $C \subset \mathbb{R}$ contains similar copies of all sets of n points then C has packing dimension at least $\frac{n-2}{n}$.

2. Proof of Theorem 1.1

Fix a sequence $\alpha_1, \alpha_2, \ldots \subset A$ so that each element of A appears infinitely many times in the sequence (α_k). Let

$$\beta_k = \max\Big(6\alpha_k, \frac{6\alpha_k}{\alpha_k - 1}\Big), \qquad (k \in \mathbb{N}). \tag{1}$$

Since $A \subset (1, \infty)$, the number β_k is defined and $\beta_k > 6$ for every k. We can clearly choose a sequence $m_1, m_2, \ldots \subset \{3, 4, 5, \ldots\}$ so that

$$\lim_{k\to\infty} \frac{\log(\beta_1 \cdots \beta_k)}{\log(m_1 \cdots m_{k-1})} = 0. \tag{2}$$

Let

$$\delta_k = \frac{1}{\beta_1 \cdots \beta_k \cdot m_1 \cdots m_k}. \tag{3}$$

By induction we shall define sets

$$E_0 \supset E_1 \supset E_2 \supset \ldots$$

such that for each $k \in \mathbb{N}$

($*$) E_k consists of $m_1 \cdots m_k$ closed intervals of length δ_k which are separated by gaps of at least δ_k and each interval of E_{k-1} contains m_k intervals of E_k.

We will denote by

$$I_1^k, I_2^k, \ldots, I_{m_1 \cdots m_k}^k$$

the intervals of E_k ordered from left to right, and by

$$(J_n, K_n, L_n)_{n\in\mathbb{Z}}$$

an enumeration of the set

$$\Gamma = \left\{(I_a^k, I_b^k, I_c^k) : a, b, c, k \in \mathbb{N}, a < b < c \leq m_1 \cdots m_k\right\}$$

such that if $n > 1$ and $(J_n, K_n, L_n) = (I_a^k, I_b^k, I_c^k)$ then $n > k$. Since each element of A appears infinitely many times in the sequence (α_k), by repeating each element of Γ infinitely many times we can also guarantee that for all $a \in A$ and for all $(J, K, L) \in \Gamma$, there exists $n \in \mathbb{N}$ such that

$$\alpha_n = a, \quad \text{and} \quad (J_n, K_n, L_n) = (J, K, L). \tag{4}$$

Let $E_0 = [0, 1]$ and choose E_1 so that $(*)$ holds for $k = 1$. Suppose that $k \geq 2$ and $E_1, \ldots, E_{k-1}$ are already defined so that $(*)$ holds for $1, \ldots, k-1$. Then (J_k, K_k, L_k) is already defined and each interval of E_{k-1} is either contained in exactly one of J_k, K_k and L_k or disjoint from them.

We shall define E_k so that

$$x \in E_k \cap J_k, \quad y \in E_k \cap K_k \quad \text{and} \quad z \in E_k \cap L_k$$

will imply that

$$\frac{z-x}{z-y} \neq \alpha_k.$$

Let I be an interval of E_{k-1} which is contained in J_k. Since I has length δ_{k-1} and using (3) and (1) we have

$$\frac{\delta_{k-1}}{3\alpha_k\delta_k} = \frac{m_k\beta_k}{3\alpha_k} \geq 2m_k > m_k + 1,$$

and I contains more than m_k points of the form $3\alpha_k\delta_k i$ for $i \in \mathbb{Z}$. Hence we can choose the m_k intervals of E_k in I as segments of the form

$$\delta_k(3i\alpha_k + [0, 1]) \quad (i \in \mathbb{Z}).$$

If I is an interval of E_{k-1} which is contained in K_k, then similarly, since

$$\frac{\delta_{k-1}}{3\delta_k} = \frac{m_k\beta_k}{3} \geq 2m_k > m_k + 1,$$

we can choose the m_k intervals of E_k in I as segments of the form

$$\delta_k(3j + [0, 1]) \quad (j \in \mathbb{Z}).$$

If I is an interval of E_{k-1} which is contained in L_k, then, since by (3) and (1) we have

$$\frac{\delta_{k-1}}{\frac{3\alpha_k}{\alpha_k-1}\delta_k} = \frac{m_k\beta_k}{\frac{3\alpha_k}{\alpha_k-1}} \geq 2m_k > m_k + 1,$$

we can choose the m_k intervals of E_k in I as segments of the form

$$\delta_k\Big(\frac{3\alpha_k}{\alpha_k - 1}(l + \tfrac{1}{2}) + [0, 1]\Big) \quad (l \in \mathbb{Z}).$$

In each of the rest of the intervals of E_{k-1} we define the m_k intervals of length δ_k of E_k arbitrarily so that they are separated by gaps of at least length δ_k.

This way we defined E_k so that $(*)$ holds. Let

$$E = \bigcap_{k=1}^{\infty} E_k.$$

Then E is clearly a compact subset of $\mathbb{R}$. Condition $(*)$ implies that the Hausdorff dimension of E is at least

$$\liminf_{k\to\infty} \frac{\log(m_1 \cdots m_{k-1})}{-\log(m_k \delta_k)}$$

(see [Falconer 1990, Example 4.6]). On the other hand, using (3) and (2) we get that

$$\liminf_{k\to\infty} \frac{\log(m_1 \cdots m_{k-1})}{-\log(m_k \delta_k)} = \liminf_{k\to\infty} \frac{\log(m_1 \cdots m_{k-1})}{\log(\beta_1 \cdots \beta_k) + \log(m_1 \cdots m_{k-1})} = 1,$$

and therefore the Hausdorff dimension of E is 1.

Finally, to get a contradiction, suppose that

$$x, y, z \in E, \quad x < y < z, \quad \text{and} \quad \frac{z-x}{z-y} \in A.$$

Since $\delta_k \to 0$, there exists a $k \in \mathbb{N}$ such that x, y and z are in distinct intervals of E_k. Then, by (4) there exists an $n \in \mathbb{N}$ so that

$$x \in J_n, \quad y \in K_n, \quad z \in L_n \quad \text{and} \quad \frac{z-x}{z-y} = \alpha_n.$$

By the construction of E_n, there exists $i, j, l \in \mathbb{Z}$ such that

$$x \in \delta_n(3i\alpha_n + [0,1]), \ \ y \in \delta_n(3j + [0,1]), \ \text{and } z \in \delta_n\Big(\frac{3\alpha_n}{\alpha_n - 1}(l + \tfrac{1}{2}) + [0,1]\Big).$$

Let

$$X = 3i\alpha_n + [0,1], \quad Y = 3j + [0,1], \quad \text{and} \quad Z = \frac{3\alpha_n}{\alpha_n - 1}(l + \tfrac{1}{2}) + [0,1].$$

Then $\frac{x}{\delta_n} \in X$, $\frac{y}{\delta_n} \in Y$ and $\frac{z}{\delta_n} \in Z$. On the other hand, $\frac{z-x}{z-y} = \alpha_n$ implies that $\alpha_n y = x + (\alpha_n - 1)z$, so (by using the notation $A + B = \{a + b : a \in A, b \in B\}$) we must have

$$\alpha_n Y \cap (X + (\alpha_n - 1)Z) \neq \varnothing. \tag{5}$$

By definition (and using that $\alpha_n > 1$),

$$\alpha_n Y = \alpha_n(3j + [0,1]) \tag{6}$$

and

$$\begin{aligned} X + (\alpha_n - 1)Z &= 3i\alpha_n + [0,1] + 3\alpha_n(l + \tfrac{1}{2}) + (\alpha_n - 1)[0,1] \\ &= 3(i + l)\alpha_n + \big[\tfrac{3}{2}\alpha_n, \tfrac{5}{2}\alpha_n\big] \\ &= \alpha_n(3(i + l) + \big[\tfrac{3}{2}, \tfrac{5}{2}\big]). \end{aligned} \tag{7}$$

Since $i, j, l \in \mathbb{Z}$, (6) and (7) contradict (5).

Acknowledgement

The author is grateful to Mihalis Kolountzakis for suggesting this problem and for helpful comments and suggestions.

References

[Bisbas and Kolountzakis 2006] A. Bisbas and M. N. Kolountzakis, "Avoiding affine copies of infinite sequences", unpublished manuscript, 2006.

[Bourgain 1987] J. Bourgain, "Construction of sets of positive measure not containing an affine image of a given infinite structures", *Israel J. Math.* **60**:3 (1987), 333–344. MR 89g:28004 Zbl 0647.28001

[Darji and Keleti 2003] U. B. Darji and T. Keleti, "Covering $\mathbb{R}$ with translates of a compact set", *Proc. Amer. Math. Soc.* **131**:8 (2003), 2593–2596. MR 2004d:03100 Zbl 1017.03023

[Elekes and Steprāns 2004] M. Elekes and J. Steprāns, "Less than 2^ω many translates of a compact nullset may cover the real line", *Fund. Math.* **181**:1 (2004), 89–96. MR 2005h:28034 Zbl 1095.28005

[Erdős 1974] P. Erdős, "Remarks on some problems in number theory", *Math. Balkanica* **4** (1974), 197–202. MR 55 #2715 Zbl 0313.10045

[Erdős and Kakutani 1957] P. Erdős and S. Kakutani, "On a perfect set", *Colloq. Math.* **4** (1957), 195–196. MR 19,734e Zbl 0077.27103

[Falconer 1984] K. J. Falconer, "On a problem of Erdős on sequences and measurable sets", *Proc. Amer. Math. Soc.* **90**:1 (1984), 77–78. MR 85e:28008 Zbl 0528.28005

[Falconer 1990] K. Falconer, *Fractal geometry*, John Wiley & Sons Ltd., Chichester, 1990. MR 92j:28008 Zbl 0689.28003

[Humke and Laczkovich 1998] P. D. Humke and M. Laczkovich, "A visit to the Erdős problem", *Proc. Amer. Math. Soc.* **126**:3 (1998), 819–822. MR 98e:28003 Zbl 0899.28001

[Keleti 1998] T. Keleti, "A 1-dimensional subset of the reals that intersects each of its translates in at most a single point", *Real Anal. Exchange* **24**:2 (1998), 843–844. MR 1704757 Zbl 0971.28001

[Kolountzakis 1997] M. N. Kolountzakis, "Infinite patterns that can be avoided by measure", *Bull. London Math. Soc.* **29**:4 (1997), 415–424. MR 98i:28003 Zbl 0879.28003

[Komjáth 1983] P. Komjáth, "Large sets not containing images of a given sequence", *Canad. Math. Bull.* **26**:1 (1983), 41–43. MR 85d:28003 Zbl 0464.28001

[Laba and Pramanik 2007] I. Laba and M. Pramanik, "Arithmetic progressions in sets of fractional dimension", preprint, 2007. To appear in *Geom. Funct. Anal.* arXiv 0712.3882

[Máthé $\geq$ 2008] A. Máthé, "Covering the real line with translates of a zero-dimensional set", In preparation.

[Svetic 2000] R. E. Svetic, "The Erdős similarity problem: a survey", *Real Anal. Exchange* **26**:2 (2000), 525–539. MR 2002g:28004 Zbl 1014.28502

Received 25 Jan 2008. Revised 30 Apr 2008. Accepted 30 May 2008.

TAMÁS KELETI: elek@cs.elte.hu
Department of Analysis, Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary
www.cs.elte.hu/analysis/keleti

PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 131, Number 8, Pages 2593–2596
S 0002-9939(02)06773-4
Article electronically published on November 14, 2002

COVERING $\mathbb{R}$ WITH TRANSLATES OF A COMPACT SET

UDAYAN B. DARJI AND TAMÁS KELETI

(Communicated by Alan Dow)

ABSTRACT. Motivated by a question of Gruenhage, we investigate when $\mathbb{R}$ is the union of less than continuum many translates of a compact set $C \subseteq \mathbb{R}$. It will follow from one of our general results that if a compact set C has packing dimension less than 1, then $\mathbb{R}$ is not the union of less than continuum many translates of C.

1. INTRODUCTION

When is $\mathbb{R}$ the union of less than continuum many translates of a given compact subset of $\mathbb{R}$? Of course, if the compact set has non-empty interior, then $\mathbb{R}$ is easily seen to be the union of countably many translates of the compact set. On the other hand, if we assume the continuum hypothesis, then it follows from the Baire category theorem that there is no such nowhere dense compact set.

Gary Gruenhage observed that it is consistent with ZFC that given a compact set of positive Lebesgue measure, one can find less than continuum many translates of it whose union is $\mathbb{R}$. Hence, for nowhere dense compact sets of positive Lebesgue measure the question whether $\mathbb{R}$ can be written as less than continuum many translates of the given set is independent of ZFC.

Gruenhage also showed that $\mathbb{R}$ is not the union of less than continuum many translates of the standard “middle 1/3 Cantor set”. Motivated by these results, he asked the following natural question:

Problem 1.1. *Is there a compact set of Lebesgue measure zero and less than continuum many translates of it whose union is* $\mathbb{R}$*?*

Of course, a positive answer to this problem would require some extra set-theoretic assumption.

For the sake of notational convenience, let us call a compact set $C \subseteq \mathbb{R}$ *thin* if it is true in ZFC that $\mathbb{R}$ is not the union of less than continuum many translates of C. (We remark here that our definition of *thin* has nothing do with the notion of *thin* in harmonic analysis.) Hence, Gruenhage's question is whether every compact set of Lebesgue measure zero is thin.

Received by the editors January 24, 2002 and, in revised form, March 14, 2002.

1991 *Mathematics Subject Classification.* Primary 03E15; Secondary 28A78.

The first author thanks the Fulbright Foundation and the Department of Analysis of Eötvös Loránd University for their hospitality.

The second author was supported by OTKA grant F 029768.

Daniel Mauldin also asked a similar question. Namely,

Problem 1.2. *Is every compact set of Hausdorff dimension less than 1 thin?*

In this note, we show that if we consider packing dimension instead of Hausdorff dimension, then the answer is affirmative.

Ronnie Levy asked whether it is true that $\mathbb{R}$ is not the union of less than continuum many similar copies of the standard middle 1/3 Cantor set. We show that the answer is affirmative. We call a set *similarity thin* if it satisfies the definition of thin with the word "translates" replaced by "similar copies". We show that compact sets with packing dimension less than 1 are similarity thin. A more general result will be obtained, too. Finally, we shall see that Problem 1.1 would be independent of ZFC if we wanted a G_δ-set instead of a compact set.

2. Results

For the sake of completeness, we first prove the following theorem due to Gruenhage. The proof given here is due to Márton Elekes.

Theorem 2.1 (Gruenhage)**.** *It is consistent with ZFC that given a compact set of positive measure, one can find less than continuum many translates of it whose union is* $\mathbb{R}$.

Proof (M. Elekes). It is consistent with ZFC that there is a set $A \subseteq \mathbb{R}$ of cardinality less than the continuum which has positive Lebesgue outer measure. (See e.g. [1].) Let C be a compact set of positive measure. By a variant of the well-known theorem of Steinhaus the sum of a measurable set with positive measure and a set with positive outer measure contains an interval. Hence, $A + C$ contains an interval. Now, let $T = \mathbb{Q} + A$. Then, T is a set with cardinality less than that of the continuum and $T + C = \mathbb{R}$. □

The basic idea behind our main result is the following simple fact. Recall that set $A \subseteq \mathbb{R}$ is similar to set B if there are numbers s, t, $s \neq 0$ such that $B = t + s \cdot A$.

Lemma 2.2. *Let C be a compact set. If there is a perfect set P such that $(t + s \cdot C) \cap P$ is countable for every t and every $s \neq 0$, then C is similarity thin.*

If $C \subseteq \mathbb{R}$, then $C^n = \{(p_1, p_2, \dots, p_n) : p_i \in C \text{ for all } 1 \leq i \leq n\}$. If $A \subseteq \mathbb{R}^n$, then $A_* = \{(x_1, x_2, \dots, x_n) \in A : (i \neq j) \implies (x_i \neq x_j)\}$. We define $F_n : \mathbb{R}^{n+2} \to \mathbb{R}^n$ by $F_n(x_1, x_2, \dots, x_n, s, t) = (t + sx_1, \dots, t + sx_n)$.

Lemma 2.3. *Suppose that C and P are compact sets. If a similar copy of C intersects P in at least n points, then $F_n(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) \cap P^n_* \neq \emptyset$.*

Proof. Let $y_1, y_2, \dots, y_n$ be n distinct points of P which are contained in some similar copy of C. Then $(y_1, \dots, y_n) \in F_n(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) \cap P^n_*$. □

The following lemma is rather well-known. A far more general result can be found in [4]. (Similar ideas were later used in [3].) We thank the referee for pointing this out.

Lemma 2.4. *Fix a positive integer n. Suppose that $M \subseteq \mathbb{R}^n$ is an F_σ, first category set. Then, there is a perfect set $P \subseteq \mathbb{R}$ such that $P^n_* \cap M = \emptyset$.*

Let us recall some terminology and facts from basic geometric measure theory. If $A \subseteq \mathbb{R}^n$, then $dim_P(A)$ and $dim_H(A)$ denote the packing dimension and the Hausdorff dimension of A, respectively. (See the definitions and basic properties e.g. [2].)

Theorem 2.5. *Every compact subset of $\mathbb{R}$ with packing dimension less than 1 is similarity thin. That is, less than continuum many similar copies of a compact set with packing dimension less than 1 cannot cover the real line.*

Proof. Let $C \subseteq \mathbb{R}$ be a compact set with packing dimension less than 1. By Lemma 2.2, it will suffice to show that there is a perfect set P such that $(t+s\cdot C)\cap P$ is finite for all real t, s with $s \neq 0$.

Recall (see e.g. [2]) that for packing dimension, we have for Borel sets A, B,

$$dim_H(A \times B) \leq dim_P(A \times B) \leq dim_P(A) + dim_P(B).$$

Hence, we may choose n sufficiently large so that $dim_P(C^n) < n-2$, which, in turn, implies that $dim_P(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) < n$ and hence

$$dim_H(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) < n.$$

We note that F_n is countably Lipschitz, i.e. we can decompose the domain of F_n into countably many compact sets $\{A_i\}$ so that $F_n|A_i$ is Lipschitz. Let $B_i = (C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) \cap A_i$. Then, since Lipschitz maps clearly cannot increase Hausdorff dimension, we have that $F_n(B_i)$ is a compact set with $(n-2)$-dimensional Lebesgue measure zero and hence is of first category. However, $F_n(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) = \bigcup_{i=1}^{\infty} F_n(B_i)$. Therefore, $F_n(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R})$ is an F_σ first category set. By Lemma 2.4, we have that there is a perfect set $P \subseteq \mathbb{R}$ such that $P_*^n \cap F_n(C^n \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}) = \emptyset$. By Lemma 2.3, we have that no similar copy of C intersects P in more than n points. By Lemma 2.2, we have that C is similarity thin. $\square$

Theorem 2.5 can be easily generalized to any countably Lipschitz, finite (say, k) parameter images instead of similar copies. One can easily check that by replacing $(x, s, t) \to t + sx$ by any other countably Lipschitz function $f : \mathbb{R} \times H \to \mathbb{R}$, where $H \subseteq \mathbb{R}^k$ replaces $(\mathbb{R} \setminus \{0\}) \times \mathbb{R}$ as the set of possible parameters and replacing 2 by k at some points, the same way as above, we can get the following generalization.

Theorem 2.6. *Let $H \subseteq \mathbb{R}^k$ be a set of possible parameters, and let $f : \mathbb{R} \times H \to \mathbb{R}$ be a countably Lipschitz function. If C is a compact set with packing dimension less than 1, then less than continuum many sets of the form $f(C, h)$ $(h \in H)$ cannot cover the real line.*

3. Remarks

As in Lemma 2.2, a negative answer to the original Problem 1.1 would follow from a negative answer to the following question:

Problem 3.1. *Is there a compact set C of Lebesgue measure zero such that every perfect set intersects at least one of the translates of C in uncountably many points?*

A positive answer to this question would only show that this method cannot solve Problem 1.1. But to this problem one can imagine (contrasted to Problem 1.1) a positive answer in ZFC.

Both problems seem just as hard if "compact" is replaced by "F_σ". However, it is consistent with ZFC that there is a set of Lebesgue measure zero and less than continuum many translates of it whose union is $\mathbb{R}$ (see e.g. [1]). In fact, if there exists a set of second category with cardinality less than the continuum (which is consistent with ZFC, see e.g. [1]), then any residual set of Lebesgue measure zero has this property. (Since, as one can easily check, the sum of a set of second category and a residual set is $\mathbb{R}$.) Since there exist residual G_δ sets with Lebesgue measure zero, this means that if we replaced "compact" in Problems 1.1 and 3.1 by "G_δ", then the positive answers would be consistent with ZFC.

References

[1] T. Bartoszyński, H. Judah, *Set theory: on the structure of the real line*, A K Peters, 1995. MR **96k:**03002

[2] P. Mattila, *Geometry of Sets and measures in Euclidean Spaces*, Cambridge University Press, 1995. MR **96h:**28006

[3] R. D. Mauldin, On the Borel Subspaces of algebraic structures, *Indiana University Math. J.* **29** (1980), 261-265. MR **81i:**54027

[4] J. Mycielski, Independent sets in topological algebras, *Fund. Math.* **55** (1964), 139-147. MR **30:**3855

Department of Mathematics, University of Louisville, Louisville, Kentucky 40292
E-mail address: ubdarj01@athena.louisville.edu

Department of Analysis, Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary
E-mail address: elek@cs.elte.hu

PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 130, Number 2, Pages 549–553
S 0002-9939(01)06344-4
Article electronically published on September 19, 2001

SHUFFLE THE PLANE

MIKLÓS ABÉRT AND TAMÁS KELETI

(Communicated by David Preiss)

Abstract. We prove that any permutation p of the plane can be obtained as a composition of a fixed number (209) of simple transformations of the form $(x, y) \to (x, y + f(x))$ and $(x, y) \to (x + g(y), y)$, where f and g are arbitrary $\mathbb{R} \to \mathbb{R}$ functions.

As a corollary we get that the full symmetric group acting on a set of continuum cardinal is a product of finitely many (209) copies of two isomorphic Abelian subgroups.

We investigate what transformations of the plane we can get by (finitely many) vertical and horizontal "slides", which we define as follows.

Definition 1. By a *vertical* (resp. *horizontal*) *slide* we mean an $\mathbb{R}^2 \to \mathbb{R}^2$ map of the form $(x, y) \to (x, y + f(x))$ (resp. $(x, y) \to (x + g(y), y)$), where f (resp. g) is an arbitrary $\mathbb{R} \to \mathbb{R}$ function.

By a *slide* we shall always mean a vertical or horizontal slide.

Note that geometrically a vertical (resp. horizontal) slide means a transformation of the plane in which we translate vertical (resp. horizontal) lines vertically (resp. horizontally).

Clearly any slide is a permutation of the plane, so the question is which permutations we can get by using (finitely many) slides. One can also ask the following (weaker) question: When can a subset of the plane be transformed to another subset using (finitely many) slides? Clearly, the sets must have the same cardinality and their complements must have the same cardinality, too — so the question is whether these conditions are sufficient or there exist other invariants of these maps.

In this paper we answer these questions by proving the following result:

Theorem 2. *Any permutation of the plane can be obtained by a fixed number* (209) *of slides. That is, for any permutation p of the plane there exist $\mathbb{R} \to \mathbb{R}$ functions $f_1, \dots, f_{105}$ and $g_1, \dots, g_{104}$ such that we have $p = F_1 G_1 \dots F_{104} G_{104} F_{105}$, where $F_i(x, y) = (x, y + f_i(x))$ and $G_i(x, y) = (x + g_i(y), y)$.*

Therefore the only invariants are the cardinality and the cardinality of the complement; a set can be mapped to another set by finitely many slides if and only if they have the same cardinality and their complements have the same cardinality,

Received by the editors July 11, 2000.

2000 *Mathematics Subject Classification.* Primary 26B40; Secondary 03E05, 20B30, 20D40.

The research of the first author was supported by the Hungarian National Foundation for Scientific Research Grant 32325.

The research of the second author was supported by the Hungarian National Foundation for Scientific Research Grant T26176.

too. In particular, there is no finitely additive non-negative function from the set of all subsets of the plane that agrees with ordinary area on squares and is invariant under both vertical and horizontal slides.

Since both the vertical and the horizontal slides form (isomorphic) Abelian subgroups of the group of all permutations of $\mathbb{R}^2$, we also get the following (purely group-theoretic) result:

Corollary 3. *The full symmetric group acting on a set of continuum cardinal is a product of finitely many* (209) *copies of two isomorphic Abelian subgroups.*

This is where the original motivation of this investigation comes from. In [1] the same result (excluding the constant 209) is proved for the full symmetric group acting on a countable set via the analogous result about slides on $\mathbb{Z} \times \mathbb{Z}$. Some ideas of the proof of Theorem 2 also come from the $\mathbb{Z} \times \mathbb{Z}$ proof of [1].

It is also proved in [1] that the full symmetric group acting on any set is a product of finitely many Abelian subgroups. There — in the non-trivial infinite case — three Abelian subgroups were used and one of them was non-isomorphic to the other two.

In order to make our proof more transparent we shall prove Theorem 2 via a lemma and several claims — the last one claims Theorem 2. At each claim we also state the number of slides we use. We never show the calculation of this number since it is straightforward using the obvious fact that the composition of two vertical (resp. horizontal) slides is just one vertical (resp. horizontal) slide.

Notation 4. In the sequel $\mathbb{N}$ will denote the set of positive integers, $\mathrm{card}(A)$ the cardinality of the set A, $\mathfrak{c}$ the continuum cardinal and $\phi|_A$ the restriction of the map ϕ to the set A.

Lemma 5. *One can give a countable partition $\bigcup_{n=1}^{\infty} A_n = \mathbb{R}$ and continuum many translated copies of every A_n such that the collection $\{A_n + t_{n,\alpha} : n \in \mathbb{N}, \alpha \in [0,1)\}$ of all translated copies are pairwise disjoint.*

Proof. Let $H = \{h_\alpha : \alpha \in [0,1)\}$ be a Hamel basis (of $\mathbb{R}$ over $\mathbb{Q}$) such that $H \subset [0,1)$. For each $k \in \mathbb{Z}, l \in \mathbb{N} \cup \{0\}$ and $q_1, \dots, q_l \in \mathbb{Q}$ let $A^{k,q_1,\dots,q_l}$ contain the reals of $[k, k+1)$ of the form $q_1 h_{\alpha_1} + \ldots + q_l h_{\alpha_l}$, where $h_{\alpha_i} \in H$ and $h_{\alpha_1} < \ldots < h_{\alpha_l}$. Clearly this is a countable partition of $\mathbb{R}$.

We shall prove that each $A^{k,q_1,\dots,q_l}$ has continuum many pairwise disjoint translated copies in $[k, k+2]$. This will complete the proof since we can easily translate the countably many intervals of length 2 into disjoint intervals, which makes the collection of all translated copies pairwise disjoint.

For fixed $l, k, q_1, \dots, q_l$, let $r \in \mathbb{Q} \cap [0,1)$ be distinct from $q_1, \dots, q_l$ such that $q_i + r \neq q_j$ for any i, j. (The last condition is not necessary but it makes the following argument simpler.) We claim that the translated sets $A^{k,q_1,\dots,q_l} + rh_\alpha$ ($\alpha \in [0,1)$) are pairwise disjoint. Indeed, expressing any $x \in A^{k,q_1,\dots,q_l} + rh_\alpha$ according to the Hamel base H there is exactly one term qh_β with coefficient q distinct from $q_1, \dots, q_l$, which implies that $h_\alpha = h_\beta$, so x cannot be in another translated set $A^{k,q_1,\dots,q_l} + rh_{\alpha'}$.

Claim 6. The horizontal line $e = \mathbb{R} \times \{0\}$ can be mapped to a set that contains the horizontal strip $S = \mathbb{R} \times [0,1)$ (by 3 slides).

Proof. We will map S into e by 3 slides; clearly the inverse of such a map has the required properties.

Let $\bigcup_{n=1}^{\infty} A_n = \mathbb{R}$ be the partition and $\{t_{n,\alpha} : n \in \mathbb{N}, \alpha \in [0,1)\}$ the collection of translations given in Lemma 5.

First, for every $n \in \mathbb{N}$ and $x \in A_n$ translate the line $\{x\} \times \mathbb{R}$ by n (vertically). The image of S is $S_1 = \bigcup_{n=1}^{\infty} A_n \times [n, n+1)$.

Then for every $n \in \mathbb{N}$ and $\alpha \in [0,1)$ translate $\mathbb{R} \times \{n+\alpha\}$ by $t_{n,\alpha}$ (horizontally). The image of S_1 is $S_2 = \bigcup_{n=1}^{\infty} \bigcup_{\alpha\in[0,1)} (A_n + t_{n,\alpha}) \times \{n+\alpha\}$. Since the sets $\{A_n + t_{n,\alpha} : n \in \mathbb{N}, \alpha \in [0,1)\}$ are pairwise disjoint we can map (in fact, project) S_2 into $e = \mathbb{R} \times \{0\}$ by a vertical slide.

Claim 7. The half-strip $[0,\infty) \times [0,1)$ can be mapped to the half-plane $[0,\infty) \times \mathbb{R}$ (by 4 slides).

Proof. First, for each $n \in \mathbb{N}$ we translate the vertical strip $[\frac{n(n-1)}{2}, \frac{n(n+1)}{2}) \times \mathbb{R}$ by $n-1$ vertically, then for each $n \in \mathbb{N}$ the horizontal strip $\mathbb{R} \times [n-1, n)$ by $-\frac{n(n-1)}{2}$ horizontally, then (for each $m = 0, 1, 2, \dots$) the strip $[m, m+1) \times \mathbb{R}$ by $-2m$ vertically and finally (for each $n \in \mathbb{N}$) the strip $\mathbb{R} \times [-n, -n+1)$ by $-n$ horizontally.

This way the half-strip $[0,\infty)\times[0,1)$ first goes to $\bigcup_{n=1}^{\infty}[\frac{n(n-1)}{2}, \frac{n(n+1)}{2})\times[n-1, n)$, then this set goes to $\bigcup_{n=1}^{\infty}[0, n) \times [n-1, n) = \bigcup_{m=0}^{\infty}[m, m+1) \times [m, \infty)$, which goes to $\bigcup_{m=0}^{\infty}[m, m+1) \times [-m, \infty) = [0,\infty] \times [0,\infty] \cup \bigcup_{n=1}^{\infty}[n,\infty) \times [-n, -n+1)$, so finally we get the half-plane $[0,\infty) \times \mathbb{R}$.

Claim 8. The set $\mathbb{R}^2 \setminus e$ can be mapped into $e = \mathbb{R} \times \{0\}$ (by 11 slides).

Proof. Gluing together Claims 6 and 7 we can map e onto a set that contains the half-plane $[0,\infty) \times \mathbb{R}$. Clearly, the same map maps $\mathbb{R}^2 \setminus e$ into the half-plane $(-\infty, 0) \times \mathbb{R}$. The inverse of this map maps the half-plane $[0,\infty) \times \mathbb{R}$ into e, so (by symmetry) we can map $(-\infty, 0) \times \mathbb{R}$ into e, too. Doing this after the first map, $\mathbb{R}^2 \setminus e$ goes into $e = \mathbb{R} \times \{0\}$.

Claim 9. Any subset A of the plane with $\mathrm{card}(A) = \mathfrak{c}$ can be mapped to a set that contains $e = \mathbb{R} \times \{0\}$ (by 3 slides).

Proof. By a vertical slide we can guarantee continuum many non-empty horizontal lines. Then we can move them horizontally so that each vertical line contains at least one point, which implies that by vertically translating them we can cover e.

Claim 10. If $\mathrm{card}(\mathbb{R}^2 \setminus H) = \mathfrak{c}$, then H can be mapped into $e = \mathbb{R} \times \{0\}$ (by 13 slides).

Proof. By Claim 9 we can map $\mathbb{R}^2 \setminus H$ to a set that contains e. Clearly, the same map maps H into $\mathbb{R}^2 \setminus e$. Then, by Claim 8, we can map the image into e.

Claim 11. If $\mathrm{card}(\mathbb{R}^2 \setminus H) = \mathfrak{c}$, then any permutation p of H can be realized by a fixed number (27) of slides. (We say that a permutation q of $\mathbb{R}^2$ "realizes" the permutation p of $H \subset \mathbb{R}^2$ if $q|_H = p$.)

Proof. Let ϕ be the map of Claim 10, which maps H into e. By a vertical slide V we can move the points of $\phi(H)$ to different horizontal lines. Then we can translate these lines horizontally, putting (for each $x \in H$) $V(\phi(x))$ above (or below) $\phi(p(x))$, then by a vertical slide to $\phi(p(x))$ and finally by ϕ^{-1} to $p(x)$.

Claim 12. If a permutation p of $\mathbb{R}^2$ has continuum many fixed points, then it can be realized by a fixed number (105) of slides.

Proof. Let $C = A \cup B$ be a partition of the set C of fixed points such that A and B have cardinality $\mathfrak{c}$. By a theorem of Ore [2] any permutation q of an infinite set can be written as $q = a^{-1}b^{-1}ab$, where a and b are permutations of the same set. Thus we can find permutations a and b of $\mathbb{R}^2 \setminus C$ such that $p|_{\mathbb{R}^2\setminus C} = a^{-1}b^{-1}ab$. Using Claim 11 (for $\mathbb{R}^2 \setminus A$ and for $\mathbb{R}^2 \setminus B$), there exists a map ϕ_a that realizes a on $\mathbb{R}^2 \setminus C$ and the identity on A, and a map ϕ_b that realizes b on $\mathbb{R}^2 \setminus C$ and the identity on B. Then $\psi = \phi_a^{-1}\phi_b^{-1}\phi_a\phi_b$ realizes p.

Claim 13. Any permutation of the plane can be realized by a fixed number (209) of slides.

Proof. Since every permutation consists of disjoint (finite or countable) cycles, any permutation of the plane can be written as a product of two permutations that have continuum many fixed points. Therefore Claim 12 completes the proof.

Remark 14. With a worse constant (417 instead of 209) one can prove Theorem 2 without using the theorem of Ore. Indeed, one can easily check that any permutation of the plane is the product of 4 permutations of order 2 such that each has an infinite support and continuum many fixed points. Then, using $(1,2)(3,4) = (1,2,3)^{-1}(1,2,4)^{-1}(1,2,3)(1,2,4)$ one can easily write these 4 permutations as commutators of the same support, so the argument of Claim 12 can be used.

Remark 15. By a different (set-theoretic) method Péter Komjáth [3] improved our result showing that if a subset of the plane and its complement have continuum cardinality, then it can be mapped to $e = \mathbb{R} \times \{0\}$ by 5 slides. This means that in Claim 10 it is enough to use 5 slides, which implies that by 81 slides any permutation of the plane can be realized.

In fact, Komjáth's method works in any infinite Abelian group: he proved that for any infinite Abelian group A, if $B \subset A \times A$ and $\mathrm{card}(B) = \mathrm{card}((A \times A) \setminus B)$, then B can be mapped to $A \times \{0\}$ by 5 slides. Note that after Claim 10 our method also works for any infinite Abelian group A (if A is countable, then in Claim 13 we also have to check that an infinite cycle can be written as the product of two permutations that have infinitely many fixed points). Therefore we get that for any infinite Abelian group A, any permutation of $A \times A$ is the composition of 81 slides. This also means that the full symmetric group acting on any infinite set is a product of 81 copies of two isomorphic Abelian subgroups.

The above results are not true for finite Abelian groups. First, for some Abelian groups (e.g. cyclic groups of odd order) any slide is an even permutation. However, one can show that for any Abelian group A any even permutation of A can be obtained as a composition of slides. But the number of slides one has to use is unbounded (see [1]).

Questions 16. It seems interesting to ask what happens if we make some natural restriction (e.g. continuous, measurable, Lipschitz, polynomial) about the functions in the definition of slides. For example what permutations of the plane can we get by finitely many "continuous" slides? Clearly we can get only orientation and measure preserving homeomorphisms but one can show that not all of them can be transformed: a continuous slide can make only a "bounded twist", so homeomorphisms with "unbounded twists" cannot be realized by finitely many continuous slides.

Clearly, "measurable" slides are measure preserving permutations of the plane. Are there other invariants besides the measure of the set and the measure of its complements? Can all measure preserving permutations of the plane be obtained by finitely many measurable slides?

References

[1] M. Abért, Symmetric groups as products of Abelian subgroups, to appear in Bull. London Math. Soc.

[2] O. Ore, Some remarks on commutators, *Proc. Amer. Math. Soc.* **2** (1951), 307-314. MR **12:**671e

[3] P. Komjáth, Five degrees of separation, to appear in Proc. Amer. Math. Soc.

Department of Algebra, Eötvös Loránd University, Kecskeméti u. 10-12, 1053 Budapest, Hungary

E-mail address: `abert@cs.elte.hu`

Department of Analysis, Eötvös Loránd University, Kecskeméti u. 10-12, 1053 Budapest, Hungary

E-mail address: `elek@cs.elte.hu`

Density and covering properties of intervals of $\mathbf{R}^n$

Tamás Keleti

(appeared in *Mathematika* **47** (2000), 229-242.)

Abstract

The key result of this paper is the existence of functions $\rho_n(h)$ for which whenever H is a (Lebesgue) measurable subset of the n-dimensional unit cube I^n with measure $|H| > h$ and $\mathcal{R}$ is a class of subintervals (n-dimensional axis-parallel rectangles) of I^n that covers H, then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $\rho_n(h)$; that is, $\frac{|H \cap R|}{|R|} > \rho_n(h)$ ($= \left(\frac{h}{2n}\right)^n$). We show how we can use this result for finding 4 points of a measurable subset of the unit square such that they are the vertices of an axis-parallel rectangle that has quite large intersection with the original set. We introduce and investigate density and covering properties of classes of subsets of $\mathbf{R}^n$. As a consequence we get a covering property of the class of intervals of $\mathbf{R}^n$: if $\mathcal{R}$ is a family of n-dimensional intervals with $|\cup \mathcal{R}| < \infty$ then there is a finite sequence $R_1, \ldots, R_m \in \mathcal{R}$ such that $|\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup \mathcal{R}|$ and $\|\sum_{k=1}^m \chi_{R_k}\|_q \leq C(n,q,\varepsilon)|\cup \mathcal{R}|^{1/q}$.

1 Introduction

While the author was working on a modified problem of A. Carbery, the following question arose:

If a measurable subset of the unit square is covered by axis-parallel rectangles (contained in the unit square) such that its density is small in each rectangle, can we conclude that the set itself must have small measure?

First note that if we allow any (not necessary axis-parallel) rectangles then the answer is negative. Indeed, a closed subset of a Nikodym set (a set in the unit square with measure one such that for each point of the set there is a straight line intersecting the set only in that single point, see e. g. [6]) with measure $1-\varepsilon$ can be easily covered by rectangles such that the density of the subset is less then ε in each rectangle.

1991 *Mathematical Reviews Classification.* Primary 28A75; Secondary 42B25.

However, as we shall prove (Theorem 2.1), for axis-parallel rectangles (even in n-dimension) the answer is affirmative. Then we can easily get a similar result for sets not necessarily in the unit cube (Theorem 2.5). Using this property of the intervals (the n-dimensional axis-parallel rectangles) of $\mathbf{R}^n$ - that we shall call the "minimal density property" or shortly MDP - we can also prove a covering property of the intervals of $\mathbf{R}^n$ (Theorem 2.6).

In Section 3 we present a result about the modified problem of A. Carbery (asked by I. Gyöngy) that motivated our investigation. Namely, we use the minimal density property of the axis-parallel rectangles to find an (axis-parallel) rectangle with vertices in a given set with large intersection with this set. For this we will also need a result for the original problem of A. Carbery.

In Section 4 we investigate the classes of subsets of $\mathbf{R}^n$ that have the minimal density property. We investigate how this property relates to some (old and new) covering properties. We show how we can improve some covering properties (the V_q property) for classes satisfying the MDP. Thus we can prove a strong covering property of the intervals of R^n.

Most of the measure theoretic results can be equivalently formulated as combinatorial ones, in the sense that the measurable sets and the intervals may be assumed to be finite unions of dyadic cubes and the coverings may be assumed to be finite. Nevertheless, the proof of our key result (Theorem 2.1) uses methods of analysis. We investigate a minimal operator analogue to the well known Hardy-Littlewood maximal operator (see e.g. [6] or [7]).

Notation 1.1 We denote by I^n the n-dimensional open unit cube; that is, $I^n = (0,1)\times\ldots\times(0,1)$. By an *interval* of R^n we mean an n-dimensional axis-parallel open rectangle: the Cartesian product of n open (1-dimensional) intervals. We denote by $\mathcal{I}^n$ the class of intervals of $\mathbf{R}^n$ and by $\mathcal{I}_0^n$ the class of (n-dimensional) subintervals of I^n.

We denote the (Lebesgue) measure and the closure of a set $A \subset \mathbf{R}^n$ by $|A|$ and $\overline{A}$, resp. By the density of A in B (with $|B| > 0$) we mean $\frac{|A\cap B|}{|B|}$.

2 Intervals of $\mathbf{R}^n$

Theorem 2.1 *If H is a measurable subset of I^n with $|H| > h$ and $\mathcal{R}$ is a class of intervals in I^n that covers H, then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $(\frac{h}{2n})^n$; that is,*

$$\frac{|H\cap R|}{|R|} > \left(\frac{h}{2n}\right)^n.$$

Proof. We prove the statement by induction on n. Let $n = 1$. Take a finite subclass of $\mathcal{R}$ that intersects H in a set of measure greater than h. It is well known that from a finite class of intervals one can always select two subclasses of disjoint intervals such that the union of the selected intervals is the same as the union of the whole class. Then, in our case, at least one of the selected classes of disjoint intervals intersects H in a set of measure greater than $h/2$. Thus in at least one of these intervals the density of H must be greater than $h/2$.

Assume that the statement is true for $n-1$. Since we can find a closed set $H' \subset H$ with $|H'| > h$ we can assume that H is closed. Then we can cover every point of $I^n \setminus H$ by an interval disjoint to H, thus we can assume that $\mathcal{R}$ covers the whole I^n.

Let

$$m(x_1, \ldots, x_n) = \inf \left\{ \frac{|H \cap (x_1 \times T)|}{|T|} \ : \ T \in \mathcal{I}_0^{n-1}, \ (x_2, \ldots, x_n) \in T \right\}.$$

Standard arguments show (see e.g. [5]) the measurability of the function $m : I^n \to [0,1]$.

Suppose that the density of H is at most b in every $R \in \mathcal{R}$. Then we prove that

$$2b \geq \int_{I^n} m > 2 \left(\frac{h}{2n} \right)^n,$$

which clearly implies our statement.

- $2b \geq \int_{I^n} m$

  Fix $x_2, \ldots, x_n \in I$. For a $t \in I$ let $K_t \times T_t$ ($K_t \in \mathcal{I}_0^1, T_t \in \mathcal{I}_0^{n-1}$) be an interval in $\mathcal{R}$ that covers $(t, x_2, \ldots, x_n)$. By definition,

  $$m(s, x_2, \ldots, x_n) \leq \frac{|H \cap (s \times T_t)|}{|T_t|} \qquad \text{(for any } s \in I\text{)}.$$

  Thus, integrating and using that the density of H in $K_t \times T_t$ is at most b, we get

  $$\int_{K_t} m(s, x_2, \ldots, x_n) ds \leq |K_t| \frac{|H \cap (K_t \times T_t)|}{|K_t \times T_t|} \leq |K_t| b. \tag{1}$$

  The intervals K_t ($t \in I$) cover I, so, taking a finite class of intervals K_t that covers I except a set of measure at most ε and selecting two subclasses of disjoint intervals with the same union, we get intervals

$K_{t_1}, \ldots, K_{t_m}$ that covers I except a set of measure at most ε such that every point is covered at most twice. From (1) we get

$$\sum_{i=1}^{m} \int_{K_{t_i}} m(s, x_2, \ldots, x_n) ds \leq \sum_{i=1}^{m} |K_{t_i}| b. \tag{2}$$

Since the intervals $K_{t_1}, \ldots, K_{t_m}$ cover I except a set of measure at most ε and $0 \leq m \leq 1$ we get that the left-hand side of (2) is at least $\int_I m(s, x_2, \ldots, x_n) ds - \varepsilon$. On the other hand, every point is covered at most twice, so the right-hand side is at most $2b$. Therefore we have $\int_I m(s, x_2, \ldots, x_n) ds \leq 2b$, which, integrated with respect to $x_2, \ldots, x_n$, gives the inequality we wanted to prove.

- $\int_{I^n} m > 2 \left(\frac{h}{2n}\right)^n$

  Let

  $$A = \{x \in H : m(x) < a\} \quad \text{where} \quad a = \left(\frac{|H|}{2n}\right)^{n-1} = \left(\frac{\frac{n-1}{n}|H|}{2(n-1)}\right)^{n-1}.$$

  Fix $x_1 \in I$. Let $A^{x_1} = \{(x_2, \ldots, x_n) : (x_1, \ldots, x_n) \in A\}$. By definition, any $(x_2, \ldots, x_n) \in A^{x_1}$ is covered by a $T \in \mathcal{I}_0^{n-1}$ with

  $$\frac{|H \cap (x_1 \times T)|}{|T|} < a.$$

  Since $A \subset H$ it implies that A^{x_1} is covered by ($n-1$-dimensional) intervals in which its density is less than a. By our induction assumption, this implies that $|A^{x_1}| \leq \frac{n-1}{n}|H|$. Thus $|A| \leq \frac{n-1}{n}|H|$, which implies that $|H \setminus A| \geq |H|/n$.

  Using this, we get that

  $$\int_{I^n} m \geq \int_{H \setminus A} m \geq \int_{H \setminus A} a = |H \setminus A| a \geq \frac{|H|}{n} \left(\frac{|H|}{2n}\right)^{n-1} > 2 \left(\frac{h}{2n}\right)^n,$$

  which completes the proof. □

Remark 2.2 Similar covering properties of intervals of the real line have been studied for very long time. The $n = 1$ case of Theorem 2.1 also follows from Youngs' First Covering Lemma ([10], 2. Lemma) from 1910, which says that if each point of a compact subset of the real line is the left-hand end-point of at least one interval then we can find a finite number of these intervals, non-overlapping, such that the measure of the non-covered part of the closed set is smaller than any fixed positive number.

Remark 2.3 The method we used in this proof is similar to the method used in [5] but, instead of Hardy-Littlewood maximal operator, we used the corresponding minimal operator. In fact, in the same way as in the proof of Theorem 2.1, we can also get a (very weak-type) inequality for the minimal operator. Namely, denoting the minimal operator associated to $\mathcal{I}_0^n$ by m_n (that is, $m_n f(x) = \inf\left\{\frac{1}{|R|}\int_R |f| : x \in R \in \mathcal{I}_0^n\right\}$ for any $f \in L_1(I^n)$), we can prove that $\rho_n\left(\int_{\{m_n f<b\}} f\right) \leq b$ for any $f : I^n \to I$ measurable function and $b > 0$. From this we can easily obtain that

$$\int_{\{m_n f<b\}} |f| \leq 2n\|f\|_\infty^{1-\frac{1}{n}} b^{\frac{1}{n}}$$

for any $f \in L_\infty(I^n)$ and $b > 0$.

A similar notion of minimal operator was introduced in [3].

Notation 2.4 We shall call the function $(\frac{h}{2n})^n$ (which appeared in Theorem 2.1) $\rho_n(h)$.

Theorem 2.5 *Suppose that H is a measurable subset of $\mathbf{R}^n$ with finite measure, $\mathcal{R}$ is a class of intervals of $\mathbf{R}^n$ that covers H and the density of H in $\cup\mathcal{R}$ is greater than $h > 0$. Then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $\rho_n(h)$; that is,*

$$\frac{|H \cap R|}{|R|} > \rho_n(h) = \left(\frac{h}{2n}\right)^n.$$

Proof. We can assume that $\mathcal{R}$ is finite since one can select a finite subclass $\mathcal{R}' \subset \mathcal{R}$ such that the density of $H \cap (\cup\mathcal{R}')$ in $\cup\mathcal{R}'$ is still greater than h.

It is known (see e.g. [6] p. 70) that if G is an open bounded subset of $\mathbf{R}^n$ and K is a compact set with positive measure then there is a disjoint sequence $\{K_k\}$ of sets homothetic to K contained in G such that $|G \setminus \cup_{k=1}^\infty K_k| = 0$.

Applying this for $G = I^n$ and $K = \overline{\cup\mathcal{R}}$, we get the sequence $\{K_k\}$ and homothecies $\phi_k : K \to K_k$. Let $H_k = \phi_k(H)$, $\mathcal{R}_k = \phi_k(\mathcal{R})$. Then $\mathcal{R}^* = \cup_{k=1}^\infty \mathcal{R}_k$ covers $H^* = \cup_{k=1}^\infty H_k$. Clearly $|H^*| = \sum |H_k|$, $1 = \sum |K_k|$ and $|H_k|/|K_k| = |H|/|K|$, so $|H^*| = |H|/|K| > h$.

Applying Theorem 2.1, we can select an $R' \in \cup_{k=1}^\infty \mathcal{R}_k$ in which the density of H^* is greater than $\rho_n(h)$. If $R' \in \mathcal{R}_k$ then the density of H in $R = \phi_k^{-1}(R') \in \mathcal{R}$ is also greater than $\rho_n(h)$. □

Theorem 2.6 *For each $n \in \mathbf{N}$ there is a function $C_n : \mathbf{R}^+ \to \mathbf{R}^+$ such that for any $\mathcal{R} \subset \mathcal{I}^n$ with $|\cup \mathcal{R}| < \infty$ and $\varepsilon > 0$ there exist $R_1, \ldots, R_m \in \mathcal{R}$ for which*

$$(i) \qquad \frac{|\cup \mathcal{R} \setminus \cup_{k=1}^m R_k|}{|\cup \mathcal{R}|} < \varepsilon$$

and

$$(ii) \qquad \frac{\sum_{k=1}^m |R_k|}{|\cup \mathcal{R}|} < C_n(\varepsilon).$$

Proof. Let

$$t = |\cup \mathcal{R}|, \qquad \mathcal{R}_\delta = \{R \in \mathcal{R} : |R| > \delta\} \quad \text{and} \quad H^\delta = \cup \mathcal{R}_\delta.$$

Then $\cup \mathcal{R} = \cup_{\delta > 0} H^\delta$, so we can choose $\delta > 0$ such that $|\cup \mathcal{R} \setminus H^\delta| < \varepsilon t/2$.

Let $H_1 = H^\delta$. Assume that $k \geq 1$ and $H_k \subset H^\delta$ is already defined. Let $a_k = |H_k|/(2t)$.

If $a_k < \varepsilon/4$ then let $m = k - 1$ and the procedure is finished.

Otherwise, applying Theorem 2.5 for the $\mathcal{R}^\delta$ covering of H_k, we get $R_k \in \mathcal{R}^\delta$ with

$$\frac{|R_k \cap H_k|}{|R_k|} > \rho_n(a_k).$$

Let $H_{k+1} = H_k \setminus R_k$.

We claim that this procedure finishes after a finite number of steps. Indeed, if $a_k \geq \varepsilon/4$ for every k then

$$|R_k \cap H_k| > \rho_n(\varepsilon/4)|R_k| > \rho_n(\varepsilon/4)\delta,$$

which is impossible since the sets $R_k \cap H_k$ are disjoint subsets of a set with finite measure.

Since $H_{m+1} = H^\delta \setminus \cup_{k=1}^m R_k$, $a_{m+1} = |H_{m+1}|/(2t) < \varepsilon/4$ and $|\cup \mathcal{R} \setminus H^\delta| < \varepsilon t/2$ we get

$$\frac{|\cup \mathcal{R} \setminus \cup_{k=1}^m R_k|}{t} < \varepsilon,$$

which means that (i) is satisfied.

Let

$$d_k = \frac{|R_k \cap H_k|}{2t} \quad (k = 1, \ldots, m).$$

Then $d_k = a_k - a_{k+1}$ and $\frac{|R_k|}{2t} < \frac{d_k}{\rho_n(a_k)}$. Thus

$$\frac{\sum_{k=1}^m |R_k|}{2t} < \sum_{k=1}^m \frac{d_k}{\rho_n(a_k)} = \sum_{k=1}^m (a_k - a_{k+1}) \frac{1}{\rho_n(a_k)}.$$

Since $\rho_n(x)$ is increasing and $a_k \geq \varepsilon/4$ the right-hand side is a lower estimate of the integral of the function $\min(1/\rho_n(x), 1/\rho_n(\varepsilon/4))$ in the interval $[a_{m+1}, a_1]$, so we get

$$\frac{\sum_{k=1}^m |R_k|}{2t} < \int_{a_{m+1}}^{a_1} \min\left(\frac{1}{\rho_n(x)}, \frac{1}{\rho_n(\frac{\varepsilon}{4})}\right) dx < \frac{\varepsilon/4}{\rho_n(\frac{\varepsilon}{4})} + \int_{\varepsilon/4}^{1/2} \frac{1}{\rho_n(x)} dx = \frac{C_n(\varepsilon)}{2}.$$

Therefore (ii) is also satisfied. $\square$

Remark 2.7 Using that $\rho_n(x) = (x/(2n))^n$, the proof above gives $C_n(\varepsilon) < \frac{n}{4(n-1)}(8n)^n(1/\varepsilon)^{n-1}$ for $n \geq 2$. For $n = 1$, using the same argument as in the very first part of the proof of Theorem 2.1, we get $C_1 = 2$.

Example 2.8 Let $0 < \delta < 1$ and let $\mathcal{R}$ be the class of axis-parallel unit squares with lower-left vertices on the segment $\{(x,y) : x + y = 0, -1 \leq x \leq 0\}$. Let $H = \{(x,y) : 0 \leq x + y \leq \delta\} \cap \cup\mathcal{R}$. Then $\mathcal{R}$ covers H and we have $|H| > \delta$ and $|\cup\mathcal{R}| = 3$. Thus the density of H in $\cup\mathcal{R}$ is greater than $\delta/3$ but its density is $\delta^2/2 = 4.5(\delta/3)^2$ in any $R \in \mathcal{R}$. Therefore Theorem 2.5 (and consequently Theorem 2.1) cannot be true with $\rho_2'(h) = 4.5h^2$, which means that only the constant can be improved in these results (for $n = 2$). (Slightly modifying this construction, we can also prove that no function greater than $h^2/2$ can be good either.)

One can check that if $R_1, \ldots, R_m \in \mathcal{R}$ (where $\mathcal{R}$ is the same as above) then $|\cup\mathcal{R} \setminus \cup_{k=1}^m R_k| > 1/m$. Thus, whenever $|\cup\mathcal{R} \setminus \cup_{k=1}^m R_k|/|\cup\mathcal{R}| < \varepsilon$, we have $m > \frac{1}{3\varepsilon}$, hence $(\sum_{k=1}^m |R_k|)/|\cup\mathcal{R}| > \frac{1}{9\varepsilon}$. Therefore Theorem 2.6 would not be true for $C_2(\varepsilon) \leq \frac{1}{9\varepsilon}$. On the other hand, according to Remark 2.7, Theorem 2.6 holds for $C_2(\varepsilon) = \frac{128}{\varepsilon}$.

In a similar way (taking n-dimensional axis-parallel cubes with "lower-left" vertices on a not too small domain of the hyperplane $\{x_1 + \ldots + x_n = 0\}$) we can show that, in higher dimensions as well, we have the best possible exponents in the above mentioned results.

3 Application

Recently A. Carbery asked the following question:

For which functions $a : [0,1] \to [0,1]$ *is it true that*

(*) *if H is a measurable subset of I^2 then one can always find 4 points of H such that they are the vertices of a (2-dimensional) interval with area at least $a(|H|)$?*

This question led I. Gyöngy to ask the following question:
For which functions $f : [0,1] \to [0,1]$ is it true that

(**) *if H is a measurable subset of I^2 then one can always find 4 points of H such that they are the vertices of a (2-dimensional) interval R such that $|R \cap H| \geq f(|H|)$?*

Clearly it is more difficult to satisfy (**) then (*). However, we shall see that, using Theorem 2.1, it is easy to obtain a function satisfying (**) from a function that satisfies (*).

A. Carbery, M. Christ and J. Wright [1] proved that $a(h) = ch^2/\log(1/h)$ (for a suitable $c > 0$ and h small enough) satisfies (*). For the sake of completeness, we sketch a proof of this result.

Proposition 3.1 *If a measurable set $H \subset I^2$ with measure u does not contain the 4 vertices of any interval with area at least v then we have*

$$u^2 \leq 2v \log \frac{1}{v} + v^2. \tag{3}$$

Proof. Since there exists a closed subset of H with measure arbitrarily close to u we can assume that H is closed. Let H_m be the union of those closed squares of the regular $m \times m$ subdivision of I^2 that intersect H. Only finitely many H_m can contain the 4 vertices of an interval with area at least v since otherwise, taking a subsequence in which all the 4 vertices converge, we would get an interval with area at least v and with vertices belonging to H. Let K_m be the set that we get by magnifying H_m with ratio m. Thus (for a fixed large m) K_m consists of at least m^2u (unit) squares and they form no (axis-parallel) rectangle with area at least m^2v.

Let k_i be the number of squares (of K_m) in the i-th row. Let P be the number of the horizontal square pairs; that is,

$$P = \sum_{i=1}^{m} \binom{k_i}{2} \geq \frac{1}{2}\frac{\left(\sum_{i=1}^{m} k_i\right)^2}{m} - \frac{1}{2}\sum_{i=1}^{m} k_i \geq \frac{1}{2}m^3(u^2 - o(1)). \tag{4}$$

The j-th and the $j+i$-th squares cannot be both in K_m in more than m^2v/i rows since otherwise K_m contains the vertices of a rectangle with area at least m^2v. Thus

$$\begin{aligned}P &\leq \sum_{i=1}^{m-1}\sum_{j=1}^{m-i}\min(m, m^2v/i) = \sum_{i=1}^{[mv]}(m-i)m + \sum_{i=[mv]+1}^{m}(m-i)m^2v/i \\ &= o(m^3) + mv(m-(mv/2))m + m^3v\log(1/v) - (m-mv)m^2v \\ &= m^3(v\log(1/v) + v^2/2 + o(1)) \qquad (5)\end{aligned}$$

Combining (4) and (5) and letting m tend to infinity, we get the inequality (3). $\square$

Corollary 3.2 *The function $a(u) = cu^2/\log(1/u)$ (if $u \leq \delta < 1$ and $a(u) = a(\delta)$ if $u > \delta$) satisfies $(*)$, where $c > 0$ depends on δ (and $c \to 1/4$ as $\delta \to 0$).* $\square$

Example 3.3 Let H_m be the union of the diagonal squares of the regular $m \times m$ subdivision of the unit square. Then clearly $|H_m| = 1/m$ and each axis-parallel rectangle with vertices in H_m has area at most $\frac{1}{m^2}$. Thus a function that satisfies $(*)$ cannot be greater than u^2. It is unknown weather $a(u) = cu^2$ satisfies $(*)$ (for a sufficiently small $c > 0$).

Proposition 3.4 *If the function a satisfies $(*)$ then $f(h) = \rho_2(h/2)a(h/2)$ satisfies $(**)$, (where $\rho_2(h) = h^2/16$ is the function that appeared in Theorem 2.1 for $n = 2$).*

Proof. Let H be a measurable subset of I^2 with measure h and let

$$\mathcal{R} = \left\{\overline{R} : R \in \mathcal{I}_0^2 : |R| \geq a\left(\frac{h}{2}\right) \text{ and } \frac{|R\cap H|}{|R|} < \rho_2\left(\frac{h}{2}\right)\right\}.$$

Let $H' = H \cap \cup\mathcal{R}$.

The class $\mathcal{R}$ covers H' but the density of H is less than $\rho_2(h/2)$ in any $R \in \mathcal{R}$. Thus, by Theorem 2.1, we must have $|H'| \leq h/2$. Hence $|H \setminus H'| \geq h/2$, so, using that the function a satisfies $(*)$, we can find an interval R with vertices in $H \setminus H'$ such that $|R| \geq a(h/2)$. Since $\overline{R}$ is not in $\mathcal{R}$ we get that $|R\cap H|/|R| \geq \rho_2(h/2)$. Thus $|R\cap H| \geq \rho_2(h/2)|R| \geq \rho_2(h/2)a(h/2) = f(h)$. $\square$

Corollary 3.5 *The function $f(h) = c'h^4/\log(1/h)$ (if $h \leq \delta < 1$ and $f(h) = f(\delta)$ if $h > \delta$) satisfies $(**)$, where c' depends only on δ.*

Proof. It follows from Corollary 3.2 and Proposition 3.4. $\square$

Example 3.6 We use a construction of I. Reiman [9]. Let p be a prime number (or a power of a prime) and let $a_1, \ldots, a_m$ be the points and $b_1, \ldots, b_m$ be the lines of the (finite) projective plane of order p, where $m = p^2 + p + 1$. We take H_p to be the union of open squares of the regular $m \times m$ subdivision of the unit square as follows: we take the square in the i-th row and j-th column if and only if a_i is on b_j. Then, using that two lines meet only in one point, we have that whenever 4 points of H_p are the vertices of an axis-parallel rectangle R then the vertices must be in one row or in one column of the subdivision. On the other hand, each line contains $p+1$ points and each point is on $p+1$ lines, so we get that $h = |H_p| = (p+1)/m > m^{-1/2}$ and $|H \cap R| \le (p+1)/m^2 \le m^{-3/2} + m^{-2} < h^3 + h^4$. Therefore $(**)$ does not hold for the function $h^3 + h^4$ ($\sim h^3$).

Therefore the best exponent (or the infimum of the exponents) for functions satisfying $(**)$ is in the interval [3,4]. This is the best we currently know.

Remark 3.7 All positive results of this section can be easily generalized to n-dimensional spaces: Instead of Proposition 3.1, with a similar counting argument, we can prove by induction that if $H \subset I^n, |H| = u_n$ and H does not contain the 2^n vertices of any n-dimensional interval then we have $u_n \le o\left(v_n^{\frac{1}{2^{n-1}+\alpha}}\right)$ (as $v_n \to 0$) for any $\alpha > 0$. Then we get that $a_n(u) = c_n u^{2^{n-1}+\alpha}$ satisfies the n-dimensional version of $(*)$ (for proper $c_n > 0$ depending only on n and α).

The proof of Proposition 3.4 clearly works in any dimension, hence the statement holds also in n-dimensions. Thus $f_n(h) = c'_n h^{n+2^{n-1}+\alpha}$ satisfies the n-dimensional version of $(**)$.

However, it is considerably more difficult to construct examples showing that we cannot have much better results than the above mentioned. The natural n-dimensional generalization of Example 3.3 (e. g. the union of those cubes of the regular $m \times \ldots \times m$ subdivision of the unit cube for which the sum of the coordinates is divisible by m) shows only that a function satisfying the n-dimensional version of $(*)$ cannot be greater than u^n. No natural generalization of Example 3.6 seems to be known.

By standard probabilistic method, it is easy to prove the following combinatorial result:

One can select $O(m^{n-n/2^{n-1}})$ points of the regular n-dimensional $m \times \ldots \times m$ lattice such that no 2^n of them are the vertices of an n-dimensional interval. Moreover, we can assume that we chose $O(m^{n-1-n/2^{n-1}})$ points of each $n-1$-dimensional $m \times \ldots \times m$ sublattice.

Then, taking the union of the corresponding open cubes of a regular subdivision of the unit cube, we get a set H with measure $O(1/m^{n/2^{n-1}})$ such that if the vertices of an n-dimensional interval R are in H then $|R| < 1/m$ and $|R \cap H| < O(1/m^{1+n/2^{n-1}})$. Thus we get $O(u^{2^{n-1}/n})$ and $O(u^{(2^{n-1}/n)+1})$ functions that do not satisfy the n-dimensional versions of $(*)$ and $(**)$, respectively; which are still quite far from our positive results.

One possible way to obtain better examples is to show that, as Erdős [4] conjectured, one can also select $O(m^{n-1/2^{n-1}})$ points of the regular n-dimensional $m \times \ldots \times m$ lattice such that no 2^n of them are the vertices of an n-dimensional interval.

Then we would have $O(u^{2^{n-1}})$ and $O(u^{2^{n-1}+1})$ functions that do not satisfy the n-dimensional versions of $(*)$ and $(**)$, respectively, which would be quite close to our positive results.

4 The minimal density property

Notation 4.1 We denote the L_q norm of a function $f : \mathbf{R}^n \to \mathbf{R}$ by $\|f\|_q$; that is, $\|f\|_q = (\int_{\mathbf{R}^n} |f|^q)^{1/q}$. The characteristic function of a set $A \subset \mathbf{R}^n$ is denoted by χ_A.

Definition 4.2 Let $\mathcal{B}$ be a class of nonempty open bounded subsets of $\mathbf{R}^n$ and $1 \le q \le \infty$.

- We say that $\mathcal{B}$ has the *minimal density property* (MDP) if there exists a function $\rho : \mathbf{R}^+ \to \mathbf{R}^+$ such that if $H \subset \mathbf{R}^n$ is measurable with finite measure, $\mathcal{R} \subset \mathcal{B}$ covers H and the density of H in $\cup\mathcal{R}$ is $d > 0$ then one can find an $R \in \mathcal{R}$ in which the density of H is greater than $\rho(d)$; that is,
$$\frac{|R \cap H|}{|R|} > \rho\left(\frac{|H|}{|\cup \mathcal{R}|}\right).$$
- The class $\mathcal{B}$ is said to have the *covering property* V_q (see [2]) if there exist constants $C < \infty$ and $c > 0$ such that for any $\mathcal{R} \subset \mathcal{B}$ with $|\cup \mathcal{R}| < \infty$ we can find $R_1, \ldots, R_m \in \mathcal{R}$ such that
$$\text{(i')} \quad |\cup_{k=1}^m R_k| \ge c\, |\cup \mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad \|\sum_{k=1}^m \chi_{R_k}\|_q \le C\, |\cup \mathcal{R}|^{1/q}.$$
- We say that $\mathcal{B}$ has the *complete covering property* V_q (CV_q) if there exists a function $C : \mathbf{R}^+ \to \mathbf{R}^+$ such that for any $\varepsilon > 0$ and $\mathcal{R} \subset \mathcal{B}$

with $|\cup\mathcal{R}|<\infty$ we can find $R_1,\ldots,R_m\in\mathcal{R}$ such that

$$\text{(i)}\quad |\cup_{k=1}^m R_k|\geq(1-\varepsilon)|\cup\mathcal{R}|\quad\text{and}\quad\text{(ii)}\quad \|\sum_{k=1}^m\chi_{R_k}\|_q\leq C(\varepsilon)|\cup\mathcal{R}|^{1/q}.$$

Remark 4.3 If $\mathcal{R}$ is a class of open sets with $|\cup\mathcal{R}|<\infty$ then for any $\varepsilon_1>0$ there exists a finite subclass $\mathcal{R}'\subset\mathcal{R}$ such that $|\cup\mathcal{R}'|\geq(1-\varepsilon_1)|\cup\mathcal{R}|$. (Indeed, since $\mathbf{R}^n$ is hereditary Lindelöf, there exist $R_1,R_2,\ldots\in\mathcal{R}$ such that $\cup_i R_i=\cup\mathcal{R}$, hence $\lim_{N\to\infty}|\cup_{j=1}^N R_j|=|\cup_i R_i|=|\cup\mathcal{R}|$.)

Therefore, if we want to prove any of the above mentioned properties, we can assume that $\mathcal{R}$ is finite.

Remark 4.4 Note that Theorem 2.5 and Theorem 2.6 state that $\mathcal{I}^n$ (the intervals of $\mathbf{R}^n$) has the minimal density property and the CV_1 property. In the proof of Theorem 2.6 we used only the minimal density property of $\mathcal{I}^n$, so we proved that MDP implies CV_1.

Remark 4.5 If we also assume that each $x\in\mathbf{R}^n$ is contained in sets $R\in\mathcal{B}$ with arbitrarily small diameter then clearly $\mathcal{B}$ is a Busemann-Feller differentiation basis with $\mathcal{B}(x)=\{R:x\in R\in\mathcal{B}\}$.

It is a standard argument that the V_1 property (which is clearly weaker than the CV_1 property) of a B-F basis $\mathcal{B}$ implies that $\mathcal{B}$ differentiates the L_∞ functions, which clearly implies the density property of the basis $\mathcal{B}$. (In fact, as Busemann and Feller proved, differentiating L_∞ is equivalent to the density property). Therefore the minimal density property implies the density property. On the other hand, as we proved the minimal density property of $\mathcal{I}^n$, we have an alternative proof of Saks' strong maximal theorem. (For these definitions and results see e. g. [6] or [7].)

Theorem 4.6

$$\mathrm{MDP}\Leftrightarrow\mathrm{CV}_1.$$

That is, for any class $\mathcal{B}$ of nonempty open bounded subsets of $\mathbf{R}^n$ the minimal density property and the CV_1 property are equivalent.

Proof. According to Remark 4.4, it is enough to prove that $\mathrm{CV}_1\Rightarrow$ MDP.

Suppose that $\mathcal{R}\subset\mathcal{B}$ covers the measurable $H\subset\mathbf{R}^n$ such that the density of H is d in $\cup\mathcal{R}$ but at most s in any $R\in\mathcal{R}$.

Using the CV_1 property of $\mathcal{B}$ for $\varepsilon = d/2 = \frac{|H|}{2|\cup\mathcal{R}|}$ we get a sequence $R_1, \ldots, R_n \in \mathcal{R}$ such that

$$(i) \qquad |\cup\mathcal{R} \setminus \cup_{k=1}^{m} R_k| \leq \varepsilon|\cup\mathcal{R}| = |H|/2$$

and

$$(ii) \qquad \sum_{k=1}^{m} |R_k| \leq C(d/2)|\cup\mathcal{R}|.$$

Since $H \subset \cup\mathcal{R}$, (i) implies that $|H \cap (\cup_{k=1}^{m} R_k)| \geq |H|/2$. Thus, using that the density of H is at most s in each R_k, we get

$$\frac{|H|}{2} \leq |H \cap (\cup_{k=1}^{m} R_k)| \leq \sum_{k=1}^{m} |H \cap R_k| \leq s \sum_{k=1}^{m} |R_k| \leq s\ C(d/2)|\cup\mathcal{R}|.$$

Therefore

$$s \geq \frac{d/2}{C(d/2)},$$

which means that choosing $\rho(d) < \frac{d}{2C(d/2)}$ we get the minimal density property of $\mathcal{B}$. $\square$

Example 4.7 Let $\mathcal{R}$ consist of sets that are the union of an open disc and an open sector with the same centre and twice larger radius.

Then $\mathcal{R}$ is clearly a regular B-F base, so it has several standard nice properties (e.g. weak 1-1 property of the maximal operator, density property, it differentiates L_1 functions).

However $\mathcal{R}$ does not have the minimal density property. Indeed, we can cover an annulus by sets of $\mathcal{R}$ (with the same centre and radius) such that the density of the annulus is arbitrary small in each set.

Therefore

1. The minimal density property is strictly stronger than the density property.

2. The minimal density property and the CV_q properties of a class cannot be proved by using only the standard methods (e.g. properties of the maximal operator).

Remark 4.8 It would be interesting to find a weak sufficient geometrical condition that guarantees the MDP. We could find (see [8]) a quite weak sufficient condition that includes for example the regular convex sets and also the star-shaped sets that contain a ball in their hub with radius at least a fixed constant times the diameter of the set. In fact, we could prove in [8] that this condition implies a Besicovitch type property, which is much stronger than the MDP (or even the CV_∞ property), which shows that the condition is too strong.

Lemma 4.9 *Let $\mathcal{B}$ be a class of nonempty bounded open subsets of $\mathbf{R}^n$ satisfying the minimal density property with the function ρ. Then for any $\varepsilon > 0$ from any sequence $R_1, R_2, \ldots \in \mathcal{B}$ with $|\cup_{i=1}^\infty R_i| < \infty$ one can select a finite subsequence $\tilde{R}_1, \ldots, \tilde{R}_m$ with the following properties:*

$$(i) \qquad |\cup_{k=1}^m \tilde{R}_k| \geq (1-\varepsilon)|\cup_{i=1}^\infty R_i| \qquad \text{and}$$

$$P_1^{\rho(\varepsilon)}: \qquad |\tilde{R}_k \setminus \cup_{j<k}\tilde{R}_j| > \rho(\varepsilon)|\tilde{R}_k| \qquad (k = 1, \ldots, m).$$

Proof. We define the subsequence $\tilde{R}_k$ by induction. If $\tilde{R}_1, \ldots, \tilde{R}_{k-1}$ is defined then let $\tilde{R}_k$ be the first element of the sequence (R_i) for which the disjoint part property $\mathrm{P}_1^{\rho(\varepsilon)}$ is satisfied for k. If there is no such R_i then the procedure is finished and $m = k - 1$. Thus we get a finite or infinite subsequence $(\tilde{R}_k)$.

The disjoint part property $\mathrm{P}_1^{\rho(\varepsilon)}$ is clearly satisfied for every k, so we have to prove only (i).

Let
$$H = \cup_{i=1}^\infty R_i \setminus \cup_k \tilde{R}_k.$$
Suppose that $x \in H$. Then there exists an index l_x for which $x \in R_{l_x}$. Since R_{l_x} were not chosen in the subsequence $(\tilde{R}_k)$ there exists a k_x for which

$$|R_{l_x} \cap H| \leq |R_{l_x} \setminus \cup_{j<k_x}\tilde{R}_j| \leq \rho(\varepsilon)|R_{l_x}|.$$

Therefore H is covered by $\mathcal{R} = \{R_{l_x} : x \in H\} \subset \mathcal{B}$ such that the density of H is at most $\rho(\varepsilon)$ in each $R \in \mathcal{R}$. Thus, by the minimal density property of $\mathcal{B}$, the density of H in $\cup\mathcal{R}$ is less than ε. Therefore

$$\varepsilon > \frac{|H|}{|\cup_{x\in H} R_{l_x}|} \geq \frac{|H|}{|\cup_{i=1}^\infty R_i|} = \frac{|\cup_{i=1}^\infty R_i \setminus \cup_k \tilde{R}_k|}{|\cup_{i=1}^\infty R_i|}.$$

Hence $|\cup_k \tilde{R}_k| > (1-\varepsilon)|\cup_{i=1}^\infty R_i|$, so, taking m large enough in the case when $(\tilde{R}_k)$ is infinite, we can satisfy (i). $\square$

Notation 4.10 Let $M_{\mathcal{B}}$ denotes the maximal operator corresponding to $\mathcal{B}$, that is

$$M_{\mathcal{B}}(f)(x) = \sup_{x \in R \in \mathcal{B}} \frac{1}{|R|} \int_R |f| \quad \text{if } x \in \cup \mathcal{B}$$

and $M_{\mathcal{B}}(f)(x) = 0$ otherwise.

Theorem 4.11 *Let $1 < p \leq \infty$ and $1/p + 1/q = 1$. If $\mathcal{B}$ has the minimal density property (or the equivalent CV_1 property) and the maximal operator $M_{\mathcal{B}}$ is weak-(p,p) than $\mathcal{B}$ has the CV_q property as well.*

Proof. Let $\mathcal{R} \subset \mathcal{B}$, $|\cup \mathcal{R}| < \infty$ and $\varepsilon > 0$. We can assume that $\mathcal{R} = \{R_1, R_2, \ldots\}$. Then, applying Lemma 4.9, we get a finite subsequence $\tilde{R}_1, \ldots, \tilde{R}_m$ satisfying (i) and $\mathrm{P}_1^{\rho(\varepsilon)}$.

Therefore we only have to prove that the disjoint part property $\mathrm{P}_1^{\rho(\varepsilon)}$ and the weak type (p,p) property of the maximal operator $M_{\mathcal{B}}$ implies that

$$\|\sum_{k=1}^{m} \chi_{\tilde{R}_k}\|_q \leq C(\varepsilon)|\cup \mathcal{R}|^{1/q}.$$

This is essentially proved in [2] in the proof of Proposition 1. (One should only replace $1/2$ by $\rho(\varepsilon)$ in that proof). $\square$

Corollary 4.12 *If $\mathcal{B}$ has the MDP (or the equivalent CV_1) then*

$$V_q \Leftrightarrow CV_q \quad (1 \leq q < \infty).$$

Proof. It is proved in [2] that the V_q property of $\mathcal{B}$ and the weak type (p,p) property of the maximal operator $M_{\mathcal{B}}$ are equivalent (if $\frac{1}{p} + \frac{1}{q} = 1$), (in fact, we need only the easy $V_q \Rightarrow$ weak-(p,p) part of this result), hence the non-trivial $V_q \Rightarrow CV_q$ implication follows from Theorem 4.11. $\square$

Corollary 4.13 *The class $\mathcal{I}^n$ has the CV_q property for any $1 \leq q < \infty$; that is, for any $n \in \mathbf{N}$, $1 \leq q < \infty$ and $\varepsilon > 0$ there exists a constant $C(n,q,\varepsilon)$ such that if $\mathcal{R}$ is a family of n-dimensional intervals and $|\cup \mathcal{R}| < \infty$ then there is a finite sequence $R_1, \ldots, R_m \in \mathcal{R}$ such that*

(i) $|\cup_{k=1}^{m} R_k| \geq (1-\varepsilon)|\cup \mathcal{R}|$ *and* (ii) $\|\sum_{k=1}^{m} \chi_{R_k}\|_q \leq C(n,q,\varepsilon)|\cup \mathcal{R}|^{1/q}$. $\square$

Remark 4.14 Taking a general Orlicz norm $\|.\|_\Phi$ we can also define the V_Φ and CV_Φ properties by replacing in Definition 4.2 (ii) by $\|\sum_{k=1}^{m}\chi_{R_k}\|_\Phi \leq C\|\chi_{\cup\mathcal{R}}\|_\Phi$. We do not know weather it is always true that if $\mathcal{B}$ has the MDP then $V_\Phi \Leftrightarrow CV_\Phi$.

Acknowledgment. I would like to thank Professors M. Laczkovich and D. Preiss for helpful discussions.

This research was, in part, supported by OTKA grant F 019468. Part of this research was done while the author was at the Mathematical Institute of the Hungarian Academy of Sciences. This note was completed while the author was visiting the University College London having a Royal Society/NATO Postdoctoral Fellowship award.

References

[1] A. Carbery, M. Christ and J. Wright, Multidimensional van der Corput and sublevel set estimates, *preprint*.

[2] A. Córdoba, R. Fefferman, A geometric proof for the strong maximal theorem, *Ann. of Math.* **102** (1975), 95-100.

[3] D. Cruz-Uribe, SFO and C. J. Neugebauer, The structure of the reverse Hölder classes, *Trans. Amer. Math. Soc.* **347** (1995), 2941-2960.

[4] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183-190.

[5] M. de Guzmán, An inequality for the Hardy-Littlewood maximal operator with respect to a product of differentiation bases, *Studia Math.* **49** (1974), 185-194.

[6] M. de Guzmán, *Differentiation of Integrals in* $\mathbf{R}^n$, Springer, Lecture Notes in Mathematics Vol. 481, Berlin, 1975.

[7] M. de Guzmán, *Real Variable Methods in Fourier Analysis*, North-Holland, Mathematics Studies Vol. 46, Amsterdam, 1981.

[8] T. Keleti, A covering property of some classes of sets in $\mathbf{R}^n$, *Acta Univ. Carol. Math. Phys.* **39** (1998), 111-118.

[9] I. Reiman, Über ein problem von K. Zarankiewicz, *Acta Math. Acad. Sci. Hung.* **9** (1958), 269-273.

[10] W. H. Young and G. C. Young, On the existence of a differential coefficient, *Proc. London Math. Soc.*, Ser. 2, **9** (1910), 325-335.

Department of Analysis,
Eötvös Loránd University,
Múzeum krt. 6-8,
H-1088 Budapest,
Hungary
e-mail: elek@cs.elte.hu

A covering property of some classes of sets in $\mathbf{R}^n$

Tamás Keleti

(appeared in *Acta Univ. Carol., Math. Phys.* **39** (1998), 111-118.)

Abstract

We prove that if $\mathcal{B}$ is a class of open bounded subsets of $\mathbf{R}^n$ satisfying a simple geometric condition then the following Besicovitch-type covering property is true. For any ε there exists an M such that from any subclass $\mathcal{R} \subset \mathcal{B}$ one can select M subclasses of disjoint sets such that the selected sets cover at least the $1-\varepsilon$ part of $\cup\mathcal{R}$.

Thus we get sufficient geometric condition for the minimal density property and for the CV_q covering properties introduced in [2].

During the proof we also get a reverse isoperimetric inequality for the union of star-shaped sets.

1 The result

In this note we prove a covering result (Theorem 3) that can be interesting in itself but also has connection with the following recently defined notions [2]. (Throughout the paper $|A|$ denotes the (Lebesgue) measure of A.)

Definition 1 Let $\mathcal{B}$ be a class of nonempty open bounded subsets of $\mathbf{R}^n$.

The class $\mathcal{B}$ is said to have the *minimal density property* (MDP) if there exists a function $\rho : \mathbf{R}^+ \to \mathbf{R}^+$ such that if $H \subset \mathbf{R}^n$ is measurable with finite measure, $\mathcal{R} \subset \mathcal{B}$ covers H and the density of H in $\cup\mathcal{R}$ is $d > 0$ then one can find an $R \in \mathcal{R}$ in which the density of H is greater than $\rho(d)$; that is,

$$\frac{|R \cap H|}{|R|} > \rho\left(\frac{|H|}{|\cup\mathcal{R}|}\right).$$

The class $\mathcal{B}$ is said to have the *complete covering property* V_q (CV_q) for a fixed $1 \le q \le \infty$ if there exists a function $C : \mathbf{R}^+ \to \mathbf{R}^+$ such that for any $\varepsilon > 0$ and $\mathcal{R} \subset \mathcal{B}$ with $|\cup\mathcal{R}| < \infty$ we can find $R_1, \ldots, R_m \in \mathcal{R}$ such that

$$\text{(i)} \quad |\cup_{k=1}^m R_k| \ge (1-\varepsilon)|\cup\mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad \|\sum_{k=1}^m \chi_{R_k}\|_q \le C(\varepsilon)|\cup\mathcal{R}|^{1/q},$$

where $\|.\|_q$ denotes the L_q norm and χ_R is the characteristic function of R.

Note that CV_q implies $\mathrm{CV}_{q'}$ if $q > q'$. It is proved in [2] that MDP and CV_1 are equivalent and also that MDP implies that CV_q is equivalent with the classical (and weaker) covering property V_q for any $1 \leq q < \infty$. (The covering property V_q is defined in [1], where - among others - the authors proved that V_q is equivalent with the weak type (p,p) property of the maximal operator associated to $\mathcal{B}$ if $1 < q < \infty$ and $1/p + 1/q = 1$.)

Unfortunately, it is not easy to prove the minimal density property (even for the simplest classes, like the class of balls), which makes the applicability of this notion harder. (In [2] the MDP is proved only for the class of intervals of $\mathbf{R}^n$ (that is, n-dimensional axis-parallel rectangles)). It would be useful and interesting to have a weak sufficient geometric condition that guarantees the MDP. One can check (see [2] Example 4.7) that the class of sets in the plane that are the union of an open disc and an open sector with the same center and twice larger radius does not have the minimal density property . However, this is a regular class of sets (see Definition 7), which shows that the standard properties (regularity, V_q property (even for $q = \infty$), weak (1,1) property of the maximal operator, density property, differentiating properties) cannot guarantee the MDP. In this example the too "sharp" "thorn" is the obstacle of the MDP.

Our main result is the theorem below that shows that if the sets of $\mathcal{B}$ are "non-thorny" in the below defined sense then $\mathcal{B}$ has a much stronger property than the MDP or the CV_q properties: instead of (ii) of Definition 1, in this case, we have a better (Besicovitch type) control for the overlapping.

Definition 2 By a *drop* we mean the interior of the convex hull of a ball and a point (not contained in the ball). The angle of the drop is the angle between the line through the point and the center of the ball and any tangent line.

Let $0 < d < 1$ and $0 < \alpha < \pi/2$. We say that a bounded open set $H \subset \mathbf{R}^n$ is (d,α)-*non-thorny* if H is the union of drops with angle at least α and diameter at least $d \cdot \mathrm{diam} H$.

Theorem 3 *Let $\mathcal{R}$ be a family of (d,α)-non-thorny sets in $\mathbf{R}^n$ with bounded diameter. Then for any $\varepsilon > 0$ one can choose sets $R_1, \ldots, R_m \in \mathcal{R}$ such that*

(i)

$$|\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup \mathcal{R}| \qquad \textit{and}$$

(ii) the sequence $R_1, \ldots, R_m$ can be distributed in M families of disjoint sets, where M depends only on n, d, α and ε.

Remark 4 This covering property is similar to the Besicovitch property, the only difference is that, instead of all the centers, we cover a big part of the union. But, as the earlier mentioned example showed, in our case the Besicovitch property itself is not enough. However, we shall use the classical Besicovitch covering theorem (for balls) in the proof but we will also need estimate for the "edge" of the union of drops. This estimate will give us a reverse isoperimetric inequality for the union of star-shaped sets (Corollary 12), which can be interesting in itself.

Corollary 5 *For any* $0 < d < 1$ and $0 < \alpha < \pi/2$, *any class of (d, α)-non-thorny sets in $\mathbf{R}^n$ has the CV_∞ property and consequently the CV_q property for any $1 \leq q < \infty$ and the minimal density property as well.*

Therefore this non-thornity is a sufficient condition for the MDP but it is in fact too strong. However, as we shall see below, quite large and important classes satisfy it.

Definition 6 A set $H \subset \mathbf{R}^n$ is said to be *star-shaped at x* if $\overline{xy} \subset H$ for every $y \in H$, where $\overline{xy}$ denotes the closed segment between x and y.

The *hub of H* ($\text{hub}(H)$) is the set of all points at which H is star-shaped.

Let $r > 0$. We say that H is *r-star-shaped* if $\text{hub}(H)$ contains an open ball with radius $r \cdot \text{diam} H$.

Definition 7 A set $H \subset \mathbf{R}^n$ is *r-regular* if there exists a cube Q that contains H such that $|H|/|Q| > r$.

It is not hard to see (and probably well-known) that if H is a convex open r-regular set in R^n then H is r'-star-shaped, where r' depends only on n and r. It is easy to see that any r-star-shaped set is (d, α)-non-thorny, where d and α depend only on r. Thus Theorem 3 has the following consequences:

Corollary 8 *If $\mathcal{R}$ is a class of convex open r-regular sets or a class of r-star-shaped sets then for any $\varepsilon > 0$ one can select M subclasses of disjoint sets such that the selected sets cover the $1 - \varepsilon$ part of $\cup \mathcal{R}$, where M depends only on n, r and ε.*

Corollary 9 *Any class of convex open r-regular sets or of r-star-shaped sets in $\mathbf{R}^n$ has the CV_∞ property and consequently the CV_q property for any $1 \leq q < \infty$ and the minimal density property as well.*

2 The proof of the result

Notation 10 Let $x \in \mathbf{R}^n$, $H \subset \mathbf{R}^n$ and $\delta > 0$. Let $S(x, \delta)$ denote the open ball with center x and radius δ. We denote the open neighborhood of H with radius δ by $S(H, \delta)$; that is,

$$S(H, \delta) = \cup_{x \in H} S(x, \delta).$$

We also introduce the δ-interior by the following definition:

$$\mathrm{int}(H, \delta) = \{x : S(x, \delta) \subset H\}.$$

We denote the diameter of a set H by $\mathrm{diam}(H)$.

Lemma 11 *Let $\mathcal{H}$ be a family of r-star-shaped sets in $\mathbf{R}^n$ with diameter D and let $A = \cup \mathcal{H}$. Then for any $\delta \leq D$ we have*

$$|S(A, \delta) \setminus A| \leq C(n, r) \frac{\delta}{D} |A|, \tag{1}$$

where $C(n, r)$ depends only on n and r. (In fact, we can choose $C(n, r) = ((1 + \frac{\sqrt{n}+1}{r})^n - 1)(\frac{\sqrt{n}+1}{r})^n$.)

Proof. By homogeneity we can assume that $D = 1$.

For any $H \in \mathcal{H}$ there exists a ball $S(O_H, r) \subset \mathrm{hub}(H)$. Consider a cubic lattice with side $\frac{2r}{\sqrt{n}+1}$ and for a lattice point P let $S_P = S(P, \frac{r}{\sqrt{n}+1})$. Let P_H be the nearest lattice point to O_H. Clearly, $P_H O_H \leq \sqrt{n} \frac{r}{\sqrt{n}+1}$, so $S_{P_H} \subset S(O_H, r) \subset \mathrm{hub}(H)$. On the other hand, the balls S_P are disjoint.

For a lattice point P let

$$K_P = \cup \{H \in \mathcal{H} : S_P \subset \mathrm{hub}(H)\}.$$

Then $A = \cup \mathcal{H} = \cup_P K_P$ and for every lattice point P we have $S_P \subset \mathrm{hub}(K_P)$ and $K_P \subset S(P, 1)$.

One can show (see e. g. [3] p. 286) that if $K \subset S(P, 1)$ and $S(P, a) \subset \mathrm{hub}(K)$ then the magnification of K with center P and ratio $1 + \frac{\delta}{a}$ contains $S(K, \delta)$. Then clearly

$$|S(K, \delta) \setminus K| \leq \left(\left(1 + \frac{\delta}{a} \right)^n - 1 \right) |K| \leq \left(\left(1 + \frac{\delta}{a} \right)^n - 1 \right) |S(0, 1)|.$$

Therefore in our case we have

$$|S(K_p, \delta) \setminus K_p| \leq \left(\left(1 + \frac{\delta}{r/(\sqrt{n}+1)} \right)^n - 1 \right) |S(0, 1)|.$$

Thus, denoting by N the number of those lattice points P for which K_P is nonempty, we have

$$|S(A,\delta)\setminus A| \le \sum_P |S(K_p,\delta)\setminus K_p| \le N\left(\left(1+\frac{\delta(\sqrt{n}+1)}{r}\right)^n - 1\right)|S(0,1)|.$$

On the other hand the balls S_P are disjoint subsets of A, hence

$$|A| \ge \sum_P |S_P| = N\left(\frac{r}{\sqrt{n}+1}\right)^n |S(0,1)|.$$

Therefore, using that $\delta \le D = 1$, we get

$$\begin{aligned}\frac{|S(A,\delta)\setminus A|}{\delta|A|} &\le \left(\frac{\sqrt{n}+1}{r}\right)^n \frac{\left((1+\frac{\delta(\sqrt{n}+1)}{r})^n - 1\right)}{\delta} \\ &\le \left(\frac{\sqrt{n}+1}{r}\right)^n \left(\left(1+\frac{\sqrt{n}+1}{r}\right)^n - 1\right) = C(n,r). \quad \square\end{aligned}$$

Corollary 12 *If E is the union of r-star-shaped sets in $\mathbf{R}^n$ with diameter D then we have*

$$\frac{\tilde{A}_+(E)}{|E|} \le \frac{C(n,r)}{D},$$

where $\tilde{A}_+(E)$ denotes the upper outer surface area in the sense of Minkowski, that is

$$\tilde{A}_+(E) = \limsup_{\delta\to 0+} \frac{|S(E,\delta)| - |E|}{\delta}. \ \square$$

Remark 13 If the diameters are not the same but between D_1 and D_2 then the same proof gives $\tilde{A}_+(E)/|E| \le C(n, rD_1/D_2)/D_2$.

Remark 14 As a special case of Corollary 12, for example, we have that the ratio of the perimeter and the area of any finite union of (not necessary axis-parallel) unit squares is at most an absolute constant.

The author does not know the best constant. Is it 4?

Facts 15 *Let D be a drop (see Definition 2) with angle $0 < \alpha < \pi/2$ and with* $\mathrm{diam} D = d$*. Let $E_\alpha = \frac{1}{\sin\alpha} + 1$ and $\delta < d/E_\alpha$. Then*

1. *the radius of the "ball part" of D is d/E_α,*

2. *the set D is $1/E_\alpha$-star-shaped,*

3. *the set* $\mathrm{int}(D,\delta)$ *(see Notation 10) is a drop with angle* α *and with diameter* $d - E_\alpha\delta$*,*

4. *we have* $S(\mathrm{int}(D,\delta), E_\alpha\delta) \supset S(D,\delta)$ *and*

5. *for any* $0 < d' < d$ *and* $0 < \alpha' < \alpha$*,* D *can be written as the union of drops with angle* α' *and diameter* d'*.* $\square$

Lemma 16 *Let* $\mathcal{K}$ *be a family of* (d,α)*-non-thorny (see Definition 2) sets in* $\mathbf{R}^n$ *with diameter between* Δ *and* 2Δ*, let* $K = \cup\mathcal{K}$ *and let* $\delta \le d/2E_\alpha$*. Then one can choose sets* $K_1, \dots, K_m \in \mathcal{K}$ *such that*

$$|S(K,\delta\Delta) \setminus \cup_{k=1}^m K_k| \le C\delta|K|, \tag{2}$$

and the sequence $K_1, \dots, K_m$ *can be distributed in* $M(\delta)$ *families of disjoint sets, where* C *depends only on* n, d *and* α *and* $M(\delta)$ *depends only on* n *and* δ*.*

Proof. By homogeneity, we can assume that $\Delta = 1$.

Let $\mathcal{D}$ be the family of those drops with diameter d and angle α that are contained in at least one of the sets of $\mathcal{K}$. Let $\mathcal{B}$ consist of the balls with radius δ contained in any drop of $\mathcal{D}$. Put

$$\mathcal{D}^* = \{\mathrm{int}(D,\delta) : D \in \mathcal{D}\} \quad \text{and} \quad K^* = \cup\mathcal{D}^*.$$

Note that, by definition and Fact 15.5, $K = \cup\mathcal{K} = \cup\mathcal{D}$ and that K^* is covered by the centers of the balls of $\mathcal{B}$. Thus, applying the classical covering theorem of Besicovitch, we get balls $B_1, \dots, B_m \in \mathcal{B}$ that cover K^* but no point of $\mathbf{R}^n$ is covered more than C_n times. For $k = 1, \dots, m$ let K_k be one of the sets of $\mathcal{K}$ that contain B_k. Then we have $\cup K_m \supset K^*$.

We claim that every set K_k intersects at most $C_n(4/\delta)^n$ sets of the sequence $K_1, \dots, K_m$ (including itself). Indeed, for a fixed k the sets K_i, that intersect K_k, are contained in a ball with radius 4 (since each set has diameter at most 2), but on the other hand, they contain balls with radius δ that cover each point at most C_n times, hence the number of sets that intersect K_k is at most $C_n|S(0,4)|/|S(0,\delta)| = C_n(4/\delta)^n$.

Thus the sequence $K_1, \dots, K_m$ can clearly be distributed in $M(\delta) = C_n(4/\delta)^n$ families of disjoint sets: the greedy algorithm easily gives a proper distribution.

Now we prove (2). Using Fact 15.4 we get

$$S(K^*, E_\alpha\delta) = \cup_{D\in\mathcal{D}} S(\mathrm{int}(D,\delta), E_\alpha\delta) \supset \cup_{D\in\mathcal{D}} S(D,\delta) = S(\cup\mathcal{D},\delta) = S(K,\delta).$$

Thus, using that $\cup_{k=1}^m K_k \supset K^*$, we have

$$S(K,\delta) \setminus \cup_{k=1}^m K_k \subset S(K^*, E_\alpha\delta) \setminus K^*.$$

According to Facts 15.2 and 15.3, $\mathcal{D}^*$ consists of $1/E_\alpha$-star-shaped sets (in fact, drops) with diameter $d - E_\alpha\delta$. Therefore, using Lemma 11 for $(\mathcal{D}^*, K^*, 1/E_\alpha, d - E_\alpha\delta, E_\alpha\delta)$ as $(\mathcal{H}, A, r, D, \delta)$ and that $\delta \le d/2E_\alpha$, we get

$$\begin{aligned} |S(K^*, E_\alpha\delta) \setminus K^*| &\le C(n, 1/E_\alpha)\frac{E_\alpha\delta}{d - E_\alpha\delta}|K^*| \\ &\le C(n, 1/E_\alpha)\frac{E_\alpha\delta}{d/2}|K| = C\delta|K|, \end{aligned}$$

where $C = C(n, 1/E_\alpha)\frac{E_\alpha}{d/2}$ depends only on n, d and α. This completes the proof of Lemma 16. $\square$

Proof of Theorem 3. Let N be a positive integer which will be defined later. By homogeneity, we can assume that each set of $\mathcal{R}$ has diameter at most $1/2$. Let $\mathcal{R}_k$ $(k = 1, 2, \ldots)$ denote the family of sets of $\mathcal{R}$ with diameter between $1/2^{k+1}$ and $1/2^k$, and let

$$\mathcal{H}^j = \mathcal{R}_j \cup \mathcal{R}_{N+j} \cup \mathcal{R}_{2N+j} \cup \ldots \quad (j = 1, \ldots, N).$$

Clearly $\mathcal{R} = \mathcal{H}^1 \cup \ldots \cup \mathcal{H}^N$.

Fix j. Let $\mathcal{K}_1 = \mathcal{R}_j$. If $\mathcal{K}_1, \ldots, \mathcal{K}_l$ is already defined then let $\mathcal{K}_{l+1}$ be the family of those sets of $\mathcal{R}_{Nl+j}$ which intersect no set of $\mathcal{K}_1, \ldots, \mathcal{K}_l$. Then the diameters of the sets of $\mathcal{K}_l$ are between $1/2^{N(l-1)+j+1}$ and $1/2^{N(l-1)+j}$ $(l = 1, 2, \ldots)$. Let $K_l = \cup\mathcal{K}_l$ and $\delta_l = 1/2^{Nl+j}$.

We claim that

$$\cup\mathcal{H}^j \subset \cup_{l=1}^\infty S(K_l, \delta_l). \tag{3}$$

Indeed, if $x \in \cup\mathcal{H}^j \setminus \cup_{l=1}^\infty K_l$ then for an index i we have $x \in R \in \mathcal{R}_{Ni+j}$. On the other hand R cannot be contained in K_{i+1}, so there must be an $l \le i$ for which R intersects K_l. Since $R \in \mathcal{R}_{Ni+j}$ we have $\operatorname{diam} R \le 1/2^{Ni+j} \le \delta_l$. Thus $x \in S(K_l, \delta_l)$, which completes the proof of (3).

If we choose N such that $1/2^{N-1} \le d/2E_\alpha$ then we can apply Lemma 16 for $\mathcal{K} = \mathcal{K}_l, \Delta = 1/2^{N(l-1)+j+1}, \delta = 1/2^{N-1}$ to get $K_1^l, \ldots, K_{m_l}^l$ such that this sequence can be distributed in $M(1/2^{N-1})$ families of disjoint sets and

$$|S(K_l, \delta_l) \setminus \cup_{i=1}^{m_l} K_i^{m_l}| \le C\frac{1}{2^{N-1}}|K_l|,$$

where C depends only on n, d and α.

Since the sets of $\mathcal{K}_l$ do not intersect the sets of $\mathcal{K}_{l'}$ (if $l \neq l'$), the sets $\{K_i^l : l \in \mathbf{N}, i = 1, \dots, m_l\}$ can also be distributed in $M(1/2^{N-1})$ families of disjoint sets. On the other hand, we have

$$\begin{aligned} |\cup \mathcal{H}_j \setminus \cup_{i,l} K_i^l| &\leq |\cup_{l=1}^{\infty} S(K_l, \delta_l) \setminus \cup_{i,l} K_i^l| \leq |\cup_{l=1}^{\infty} (S(K_l, \delta_l) \setminus \cup_{i=1}^{m_l} K_i^l)| \\ &\leq \sum_{l=1}^{\infty} |(S(K_l, \delta_l) \setminus \cup_{i=1}^{m_l} K_i^l| \leq C \frac{1}{2^{N-1}} \sum_{l=1}^{\infty} |K_l| \\ &\leq \frac{C}{2^{N-1}} |\cup \mathcal{H}^j| \leq \frac{C}{2^{N-1}} |\cup \mathcal{R}|. \end{aligned}$$

Until this moment j was fixed. Now let $R_1, R_2, \dots$ be the union of the families $\{K_i^l\}$ we get for $j = 1, \dots, N$. Then these sets can be distributed in $NM(1/2^{N-1})$ families of disjoint sets and

$$|\cup \mathcal{R} \setminus \cup_k R_k| \leq \frac{NC}{2^{N-1}} |\cup \mathcal{R}|.$$

Therefore, if N is an integer such that $\frac{NC}{2^{N-1}} < \varepsilon$ and $1/2^{N-1} \leq d/2E_\alpha$ (depending only on n, d, α and ε) and $M = NM(1/2^{N-1})$ (depending also only on n, d, α and ε), then (i) and (ii) of Theorem 3 are satisfied if m is large enough. □

Acknowledgment. The author would like to thank Professor D. Preiss for helpful discussions.

This research was, in part, supported by OTKA grant F 019468. Part of this research was done while the author was at the Mathematical Institute of the Hungarian Academy of Sciences. This note was completed while the author was visiting the University College London having a Royal Society/NATO Postdoctoral Fellowship award.

References

[1] A. Córdoba, R. Fefferman, A geometric proof for the strong maximal theorem, *Ann. of Math.* **102** (1975), 95-100.

[2] T. Keleti, Density and covering properties of intervals of $\mathbf{R}^n$, *submitted.*

[3] S. Vrećica, A note on starshaped sets, *Publ. Inst. Math.* **29 (43)** (1981), 283-288.

Department of Analysis, Eötvös Loránd University,
Múzeum krt. 6-8, H-1088 Budapest, Hungary
e-mail: elek@cs.elte.hu

Self-similar and self-affine sets; measure of the intersection of two copies

Márton Elekes††, Tamás Keleti‡† and András Máthé‡‡

† *Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, P.O. Box 127, H-1364, Budapest, Hungary*

‡ *Department of Analysis, Eötvös Loránd University, Pázmány Péter sétány 1/c, H-1117 Budapest, Hungary*

(e-mail: emarci@renyi.hu, elek@cs.elte.hu, amathe@cs.elte.hu)

(*Received 2008*)

Abstract. Let $K \subset \mathbb{R}^d$ be a self-similar or self-affine set and let μ be a self-similar or self-affine measure on it. Let $\mathcal{G}$ be the group of affine maps, similitudes, isometries or translations of $\mathbb{R}^d$. Under various assumptions (such as separation conditions or we assume that the transformations are small perturbations or that K is a so called Sierpiński sponge) we prove theorems of the following types, which are closely related to each other;

- *(Non-stability)*

  There exists a constant $c < 1$ such that for every $g \in \mathcal{G}$ we have either $\mu\big(K \cap g(K)\big) < c \cdot \mu(K)$ or $K \subset g(K)$.

- *(Measure and topology)*

  For every $g \in \mathcal{G}$ we have $\mu\big(K \cap g(K)\big) > 0 \iff \mathrm{int}_K(K \cap g(K)) \neq \emptyset$ (where int_K is interior relative to K).

- *(Extension)*

  The measure μ has a $\mathcal{G}$-invariant extension to $\mathbb{R}^d$.

Moreover, in many situations we characterize those g's for which $\mu\big(K \cap g(K)\big) > 0$. We also get results about those g's for which $g(K) \subset K$ or $g(K) \supset K$.

† Supported by Hungarian Scientific Foundation grant no. 37758 and 61600 and János Bolyai Fellowship.
† Supported by Hungarian Scientific Foundation grant no. 72655 and János Bolyai Fellowship.
‡ Supported by Hungarian Scientific Foundation grant no. 49786.

1. *Introduction*

The study of the size of the intersection of Cantor sets has been a central research area in geometric measure theory and dynamical systems lately, see e.g. the works of Igudesman [**12**], Li and Xiao [**17**], Moreira [**23**], Moreira and Yoccoz [**24**], Nekka and Li [**25**], Peres and Solomyak [**26**]. For instance J-C. Yoccoz and C. G. T. de Moreira [**24**] proved that if the sum of the Hausdorff dimensions of two regular Cantor sets exceeds one then, in the typical case, there are translations of them stably having intersection with positive Hausdorff dimension.

The main purpose of this paper is to study the measure of the intersection of two Cantor sets which are (affine, similar, isometric or translated) copies of a self-similar or self-affine set in $\mathbb{R}^d$. By measure here we mean a self-similar or self-affine measure on one of the two sets.

We get instability results stating that the measure of the intersection is separated from the measure of one copy. This strong non-continuity property is in sharp contrast with the well known fact that for any Lebesgue measurable set $H \subset \mathbb{R}^d$ with finite measure the Lebesgue measure of $H \cap (H + t)$ is continuous in t.

We get results stating that the intersection is of positive measure if and only if it contains a relative open set. This result resembles some recent deep results (e.g. in [**16**], [**24**]) stating that for certain classes of sets having positive Lebesgue measure and nonempty interior is equivalent. In the special case when the self-similar set is the classical Cantor set our above mentioned results were obtained by F. Nekka and Jun Li [**25**]. For other related results see also the work of Falconer [**5**], Feng and Wang [**8**], Furstenberg [**9**], Hutchinson [**11**], Järvenpää [**13**] and Mattila [**19**], [**20**], [**21**].

As an application we also get isometry (or at least translation) invariant measures of $\mathbb{R}^d$ such that the measure of the given self-similar or self-affine set is 1.

Feng and Wang [**8**] has proved recently "The Logarithmic Commensurability Theorem": they showed logarithmic commensurability of the similarity ratio of a homogeneous self-similar set in $\mathbb{R}$ with the open set condition and of a similarity map that maps the self-similar set into itself (see more precisely after Theorem 4.9). They also posed the problem of generalizing their result to higher dimensions. For self-similar sets with the strong separation condition we prove a higher dimensional generalization without assuming homogeneity.

1.1. *Self-affine sets.*

Let $K \subset \mathbb{R}^d$ be a *self-affine set* with the *strong separation condition*; that is, $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ is a compact set, where $r \geq 2$ and $\varphi_1, \ldots, \varphi_r$ are injective and (in some norm) contractive $\mathbb{R}^d \to \mathbb{R}^d$ affine maps and $\cup^*$ denotes disjoint union.

For any $p_1, \ldots, p_r \in (0, 1)$ such that $p_1 + \ldots + p_r = 1$ let μ be the corresponding *self-affine measure*; that is, the image of the infinite product of the discrete probability measure $p(\{i\}) = p_i$ on $\{1, \ldots, r\}$ under the representation map $\pi : \{1, \ldots, r\}^{\mathbb{N}} \to K, \quad \{\pi(i_1, i_2, \ldots)\} = \cap_{n=1}^{\infty} (\varphi_{i_1} \circ \ldots \circ \varphi_{i_n})(K)$.

In Section 3 we show (Theorem 3.2) that small affine perturbations of K cannot intersect a very large part of K; that is, there exists a $c < 1$ and a neighborhood U

of the identity map in the space of affine maps such that for any $g \in U \setminus \{\text{identity}\}$ we have $\mu\big(K \cap g(K)\big) < c$. We also prove (Theorem 3.5) that no isometric but nonidentical copy of K can intersect a very large part of K; that is, there exists a constant $c < 1$ such that for any isometry g either $\mu\big(K \cap g(K)\big) < c$ or $g(K) = K$.

1.2. *Self-similar sets.* Now let $K \subset \mathbb{R}^d$ be a *self-similar set* with the *strong separation condition* and μ a self-similar measure on it; that is, K and μ are defined as above with the extra assumption that $\varphi_1, \dots, \varphi_r$ are similitudes.

In Section 4 we prove (Theorem 4.1) that for any given self-similar set $K \subset \mathbb{R}^d$ with the strong separation condition and self-similar measure μ on K there exists a $c < 1$ such that for any similitude g either $\mu\big(K \cap g(K)\big) < c \cdot \mu(K) = c$ or $K \subset g(K)$. In other words, the intersection of a self-similar set with the strong separation condition and its similar copy cannot have a really big non-trivial intersection.

Let K, μ and g be as above. An obvious way of getting $\mu\big(K \cap g(K)\big) > 0$ is when $g(K)$ contains a nonempty (relative) open set in K. The main result (Theorem 4.5) of Section 4, which will follow from the above mentioned Theorem 4.1, shows that this is the only way. That is, for any self-similar set $K \subset \mathbb{R}^d$ with the strong separation condition and self-similar measure μ on K a similar copy of K has positive μ measure in K if and only if it has nonempty relative interior in K.

An immediate consequence (Corollary 4.6) of the above result is that for any fixed self-similar set with the strong separation condition and for any two self-similar measures μ_1 and μ_2 we have $\mu_1\big(g(K) \cap K\big) > 0 \iff \mu_2\big(g(K) \cap K\big) > 0$ for any similitude g. As another corollary (Corollary 4.7) we get that for any given self-similar set $K \subset \mathbb{R}^d$ with the strong separation condition and self-similar measure μ on K there exist only countably many (in fact exactly countably infinitely many) similitudes $g : A_K \to \mathbb{R}^d$ (where A_K is the affine span of K) such that $g(K) \cap K$ has positive μ-measure.

Let $K \subset \mathbb{R}^d$ be a self-similar set with the strong separation condition and let s be its Hausdorff dimension, which in this case equals its similarity and box-counting dimension. Then the s-dimensional Hausdorff measure is a constant multiple of a self-similar measure (one has to choose $p_i = a_i^s$, where a_i is the similarity ratio of φ_i). Therefore all the above results hold when μ is s-dimensional Hausdorff measure.

In Section 4 we also need and get results (Proposition 4.3, Lemma 4.8, Theorem 4.9 and Corollary 4.10) stating that only very special similarity maps can map a self-similar set with the strong separation condition into itself. Theorem 4.9 and Corollary 4.10 are the already mentioned generalizations of The Logarithmic Commensurability Theorem of Feng and Wang [**8**].

In Section 5 we apply the main result (Theorem 4.5) and some of the above mentioned results (Lemma 4.8 and Theorem 4.9) of Section 4 to characterize those self-similar measures on a self-similar set with the strong separation condition that can be extended to $\mathbb{R}^d$ as an isometry invariant Borel measure. It turns out that, unless there is a clear obstacle, any self-similar measure can be extended to $\mathbb{R}^d$ as an isometry invariant measure. Thus, for a given self-similar set with the strong separation condition, there are usually many distinct isometry invariant Borel measures for which the set is of measure 1.

Let us simply call a measure defined on K isometry invariant if it can be extended to an isometry invariant measure on $\mathbb{R}^d$. Many different collections of similitudes can define the same self-similar set. We call $\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ a *presentation* of K if $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ holds; in other words, K is the attractor of the iterated function system $\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ with the extra condition of disjointness.

The notion of a self-similar measure on K depends on the particular presentation. However, we show that the notion of isometry invariant self-similar measure on K is independent of the presentations (Theorem 5.5). By this theorem we can define a natural number for each self-similar set (satisfying the strong separation property), an invariant, which does not depend on the presentation (Theorem 5.7). This invariant is equal to the dimension of the space of isometry invariant self-similar measures, and is related to the algebraic dependence of the similitudes of some (any) presentation of K.

In Section 6 we show that the connection between different presentations of a self-similar set can be very complicated. This sheds some light on why results and their proofs in Section 5 are complicated. The structure of different presentations of a self-similar set in $\mathbb{R}$ has been also studied recently and independently by Feng and Wang in [**8**], where a similar example is presented.

1.3. *Self-affine sponges.* Take the unit cube $[0,1]^n$ in $\mathbb{R}^n$ and subdivide it into $m_1 \times \ldots \times m_n$ boxes of the same size ($m_1, \ldots, m_n \geq 2$) and cut out some of them. Then do the same with the remaining boxes using the same pattern as in the first step and so on. What remains after infinitely many steps is a self-affine set, which is called *self-affine Sierpiński sponge*. (A more precise definition will be given in Definition 2.14.)

For $n = 2$ these sets were studied in several papers (in which they were called self-affine carpets or self-affine carpets of Bedford and McMullen). Bedford [**2**] and McMullen [**22**] determined the Hausdorff and Minkowski dimensions of these self-affine carpets. (The Hausdorff and Minkowski dimension of self-affine Sierpiński sponges was determined by Kenyon and Peres [**15**]). Gatzouras and Lalley [**10**] proved that except in some relatively simple cases such a set has zero or infinity Hausdorff measure in its dimension (and so in any dimension). Peres extended their results by proving that (except in the same rare simple cases) for any gauge function neither the Hausdorff [**28**] nor the packing [**27**] measure of a self-affine carpet can be positive and finite (in fact, the packing measure cannot be σ-finite either), and remarked that these results extend to self-affine Sierpiński sponges of higher dimensions.

Recently the first and the second listed authors of the present paper showed [**4**] that some nice sets – among others the set of Liouville numbers – have zero or non-σ-finite Hausdorff and packing measure for any gauge function by proving that these sets have zero or non-σ-finite measure for any translation invariant Borel measure. (Much earlier Davies [**3**] constructed a compact subset of $\mathbb{R}$ with this property.) So it was natural to ask whether the self-affine carpets of Bedford and McMullen have this stronger property.

In Section 7 we prove (Corollary 7.7) that for any self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ with the natural Borel probability measure μ (see in Definition 2.15) on K and $t \in \mathbb{R}^n$, the set $K \cap (K+t)$ has positive μ measure if and only if it has non-empty interior relative to K.

For this we prove (Theorem 7.4) that for any self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ and translation vector $t \in \mathbb{R}^n$ we have $\mu\big(K \cap (K+t)\big) = 0$ unless K or t are of very special form.

We also characterize (Theorem 7.9) those Sierpiński sponges for which we do not have instability result for translations and the natural probability measure μ. In fact, we get that $\mu\big(K \cap (K+t)\big)$ can be close to 1 only for the same special sponges that appear in the above mentioned result.

In Section 8 we show (Theorem 8.1) that for any self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ the natural probability measure μ on K can be extended as a translation invariant Borel measure ν on $\mathbb{R}^n$. We also extend this result (Theorem 8.2, Corollary 8.3) to slightly larger classes of self-affine sets.

2. *Notation, basic facts and some lemmas*

In this section we collect several notions and well known or fairly easy statements that we will need in the sequel. Some of these might be interesting in their own right. Of course, only a few of them are needed for each specific section. Though some of these statements may be well known, for the sake of completeness we included the proofs.

Notation 2.1. We shall denote by $\cup^*$ the disjoint union and by dist the Euclidean distance.

2.1. *Affine maps, similitudes, isometries.*

Definition 2.2. A mapping $g : \mathbb{R}^d \to \mathbb{R}^d$ is called a *similitude* if there is a constant $r > 0$, called *similarity ratio*, such that $\operatorname{dist}(g(a), g(b)) = r \cdot \operatorname{dist}(a, b)$ for any $a, b \in \mathbb{R}^d$.

The *affine maps* of $\mathbb{R}^d$ are of the form $x \mapsto Ax + b$, where A is a $d \times d$ matrix and $b \in \mathbb{R}^d$ is a translation vector. Thus the set of all affine maps of $\mathbb{R}^d$ can be considered as $\mathbb{R}^{d^2+d}$ and so it can be considered as a metric space.

It is easy to check that a sequence (g_n) in this metric space converges to an affine map g if and only if g_n converges to g uniformly on any compact subset of $\mathbb{R}^d$.

Definition 2.3. For a given set $K \subset \mathbb{R}^d$ with affine span A_K let $\mathcal{A}_K$, $\mathcal{S}_K$ and $\mathcal{I}_K$ denote the metric space (with the above metric) of the injective affine maps, similitudes and isometries of A_K into itself, respectively.

Note also that all these three metric spaces with the composition can be also considered as topological groups.

2.2. *Self-similar and self-affine sets and measures.*

DEFINITION 2.4. A $K \subset \mathbb{R}^d$ compact set is *self-similar* if $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$, where $r \geq 2$ and $\varphi_1, \ldots, \varphi_r$ are contractive similitudes.

A $K \subset \mathbb{R}^d$ compact set is *self-affine* if $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$, where $r \geq 2$ and $\varphi_1, \ldots, \varphi_r$ are injective affine maps, and there is a norm in which they are all contractions.

By the *n-th generation elementary pieces of* K we mean the sets of the form $(\varphi_{i_1} \circ \ldots \circ \varphi_{i_n})(K)$, where $n = 0, 1, 2, \ldots$.

We shall use multi-indices. By a *multi-index* we mean a finite sequence of indices; for $I = (i_1, i_2, \ldots, i_n)$ let $\varphi_I = \varphi_{i_1} \circ \ldots \circ \varphi_{i_n}$ and $p_I = p_{i_1} p_{i_2} \ldots p_{i_n}$. We shall consider $I = \emptyset$ as a multi-index as well: $\varphi_\emptyset$ is the identity map and $p_\emptyset = 1$.

Note that the elementary pieces of K are the sets of the form $\varphi_I(K)$. These sets are also self-similar/self-affine; and if h is a similitude / injective affine map then $h(K)$ is also self-similar/self-affine and its elementary pieces are the sets of the form $h(\varphi_I(K))$. Note also that every point of K is contained in an arbitrarily small elementary piece.

DEFINITION 2.5. Let $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ be a self-similar/self-affine set, and let $p_1 + \ldots + p_r = 1$, $p_i > 0$ for all i. Consider the symbol space $\Omega = \{1, \ldots, r\}^{\mathbb{N}}$ equipped with the product topology and let ν be the Borel measure on Ω which is the countable infinite product of the discrete probability measure $p(\{i\}) = p_i$ on $\{1, \ldots, r\}$. Let

$$\pi : \Omega \to K, \quad \{\pi(i_1, i_2, \ldots)\} = \cap_{n=1}^{\infty} (\varphi_{i_1} \circ \ldots \circ \varphi_{i_n})(K)$$

be the continuous addressing map of K. Let μ be the image measure of ν under the projection π; that is,

$$\mu(H) = \nu\big(\pi^{-1}(H)\big) \quad \text{for every Borel set } H \subset K. \tag{1}$$

Such a μ is called a *self-similar/self-affine measure* on K.

One can also define (see e.g. in [**7**]) self-similar or self-affine measures as the unique probability measure μ on K such that

$$\mu(H) = \sum_{i=1}^{r} p_i \mu\big(\varphi_i^{-1}(H)\big)$$

holds for every Borel set $H \subset K$. It was already proved by Hutchinson [**11**] that the two definitions agree.

LEMMA 2.6. *Let $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ be a self-affine set, $p_1 + \ldots + p_r = 1$, $p_i > 0$ for all i, and let μ be the self-affine measure on K corresponding to the weights p_i. Then for every affine subspace A either $\mu(A \cap K) = 0$ or $A \supset K$.*

Proof. Let $\{x_1, x_2, \ldots, x_k\}$ be a maximal collection of affine independent points in K. Choose $U_1, \ldots, U_k$ convex open sets such that $x_j \in U_j$ $(j = 1, \ldots, k)$ and whenever we choose one point from each U_j they are affine independent. Since

$K \cap U_i$ is a nonempty relative open subset of K, we may choose an elementary piece $\varphi_{I_j}(K)$ in U_j for each j. Let $\varepsilon = \min_{1 \le j \le k} p_{I_j} > 0$.

We shall use the notation we introduced in Definition 2.5. For $1 \le i \le r$ and $\omega = (i_0, i_1, \ldots) \in \Omega$, let $\sigma_i(\omega) = (i, i_0, i_1, \ldots)$. Thus $\nu\big(\sigma_i(H)\big) = p_i \nu(H)$ for all Borel subset H of Ω.

Suppose that A is an affine subspace such that $\mu(A \cap K) > 0$. Thus $\nu\big(\pi^{-1}(A)\big) > 0$. It is easy to prove (see a possible argument later in the proof of Lemma 2.12) that this implies that there exists an elementary piece $\sigma_J(\Omega)$ such that

$$\nu\big(\pi^{-1}(A) \cap \sigma_J(\Omega)\big) > (1-\varepsilon)\nu\big(\sigma_J(\Omega)\big) = (1-\varepsilon)p_J.$$

Since $\nu\big((\sigma_J \circ \sigma_{I_j})(\Omega)\big) = p_J p_{I_j} \ge p_J \varepsilon$ $(j = 1, \ldots, k)$, the set $\pi^{-1}(A)$ must intersect the sets $(\sigma_J \circ \sigma_{I_j})(\Omega)$. Therefore the set A must intersect the sets $\pi((\sigma_J \circ \sigma_{I_j})(\Omega)) = (\varphi_J \circ \varphi_{I_j})(K)$ $(j = 1, \ldots, k)$.

By picking one point from each $A \cap (\varphi_J \circ \varphi_{I_j})(K)$, we get a maximal collection of affine independent points in K since φ_J is an invertible affine mapping. As this collection is contained in the affine subspace A, we get that K is also contained in A. $\square$

Remark 2.7. In this paper one of our main goals is to study $\mu\big(K \cap g(K)\big)$, where g is an affine map of $\mathbb{R}^d$. By the above lemma if the affine map g does not map the affine span A_K of K onto itself then $\mu\big(g(K) \cap K\big) = 0$ since $K \not\subset g(A_K)$. The other property of affine maps we are interested in is $K \subset g(K)$, which also implies that g maps A_K onto itself. Thus it is enough to consider those affine maps g of $\mathbb{R}^d$ that map the affine span A_K of K onto itself. Since then both K and $g(K)$ are in A_K, only the restriction of g to A_K matters. This is why in the next section we shall study $\mathcal{A}_K$, $\mathcal{S}_K$ and $\mathcal{I}_K$ (the injective affine maps, similitudes and isometries of A_K into itself) instead of all affine maps, similitudes and isometries of $\mathbb{R}^d$.

Therefore if we state something (about $\mu\big(g(K) \cap K\big)$ or about the property $K \subset g(K)$) for every affine map, similitude or isometry g, it will be enough to prove them for $g \in \mathcal{A}_K$, $g \in \mathcal{S}_K$ or $g \in \mathcal{I}_K$, respectively.

Note also that self-similar sets and measures are self-affine as well, so results about self-affine sets and measures also apply for self-similar sets and measures.

2.3. *Separation properties.*

Definition 2.8. A self-similar/self-affine set $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ (or more precisely, the collection $\varphi_1, \ldots, \varphi_r$ of the representing maps) satisfies the

- *strong separation condition* (SSC) if the union $\varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ is disjoint;
- *open set condition* (OSC) if there exists a nonempty bounded open set $U \subset \mathbb{R}^d$ such that $\varphi_1(U) \cup^* \ldots \cup^* \varphi_r(U) \subset U$;
- *strong open set condition* (SOSC) if there exists a nonempty bounded open set $U \subset \mathbb{R}^d$ such that $U \cap K \neq \emptyset$ and $\varphi_1(U) \cup^* \ldots \cup^* \varphi_r(U) \subset U$;

- *convex open set condition* (COSC) if there exists a nonempty bounded open convex set $U \subset \mathbb{R}^d$ such that $\varphi_1(U) \cup^* \ldots \cup^* \varphi_r(U) \subset U$;
- *measure separation condition* (MSC) if for any self-similar/self-affine measure μ on K we have $\mu\big(\varphi_i(K) \cap \varphi_j(K)\big) = 0$ for any $1 \le i < j \le r$.

We note that the first three definitions are standard but we have not seen any name for the last two in the literature.

It is easy to check the well known fact that we must have $K \subset \overline{U}$ (where $\overline{E}$ denotes the closure of a set E) for the open set U in the definition of OSC (and SOSC, COSC).

It is easy to see (U can be chosen as a small ε-neighborhood of K for the first implication) that for any self-affine set

$$SSC \Longrightarrow SOSC \Longrightarrow OSC.$$

Using the methods of C. Bandt and S. Graf [**1**], A. Schief proved in [**30**] that, in fact, $SOSC \Longleftrightarrow OSC$ holds for self-similar sets.

In [**30**] for self-similar sets $SOSC \Longrightarrow MSC$ is also proved. Since the proof works for self-affine sets as well we get that for any self-affine set

$$SOSC \Longrightarrow MSC.$$

It seems to be also true that $COSC \Longrightarrow SOSC$ and so $COSC \Longrightarrow MSC$ but we do not prove this, since we do not need the first implication and the following lemma is stronger than the second implication.

LEMMA 2.9. *Let $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ be a self-affine set in $\mathbb{R}^d$ with the convex open set condition and let μ be a self-affine measure on it. Then for any affine map $\Psi : \mathbb{R}^d \to \mathbb{R}^d$ we have*

$$\mu\Big(\Psi\big(\varphi_i(K) \cap \varphi_j(K)\big)\Big) = 0 \qquad (\forall\, 1 \le i < j \le r).$$

Proof. Let $1 \le i < j \le r$ and U be the convex open set given in the definition of COSC. Let A_K be the affine span of K. Since $\varphi_i(U \cap A_K)$ and $\varphi_j(U \cap A_K)$ are disjoint convex open sets in A_K, $\overline{\varphi_i(U \cap A_K)} \cap \overline{\varphi_j(U \cap A_K)}$ must be contained in a proper affine subspace A of A_K. Since $K \subset \overline{U} \cap A_K$, this implies that $\varphi_i(K) \cap \varphi_j(K) \subset A$, and so

$$\Psi\big(\varphi_i(K) \cap \varphi_j(K)\big) \subset \Psi(A). \tag{2}$$

Since $\Psi(A)$ is an affine subspace, which is smaller dimensional than the affine span A_K of K, we cannot have $K \subset \Psi(A)$, so by Lemma 2.6 we must have $\mu\big(K \cap \Psi(A)\big) = 0$. By (2) this implies that $\mu\big(\Psi(\varphi_i(K) \cap \varphi_j(K))\big) = 0$. □

We also note that one can find a self-similar set in $\mathbb{R}$ that satisfies even the SSC but does not satisfy the COSC [**8**, Example 5.1], so SSC and COSC are independent even for self-similar sets of $\mathbb{R}$.

NOTATION 2.10. Given a fixed measure μ, we shall say that two sets are *almost disjoint* if their intersection has μ-measure 0. The almost disjoint union will be denoted by $\cup^{**}$.

It is very easy to prove one by one each of the following facts.

FACTS 2.11. *Let $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ be a self-affine/self-similar set with the measure separation condition and let μ be a self-affine/self-similar measure on it, which corresponds to the weights $p_1, \ldots, p_r$. Then the following statements hold.*

1. *Any two elementary pieces of K are either almost disjoint or one contains the other.*
2. *Any union of elementary pieces can be replaced by an almost disjoint countable union.*
3. *For any multi-index I we have $\mu \circ \varphi_I = p_I \cdot \mu$; that is, $\mu \circ \varphi_I(B) = p_I \cdot \mu(B)$ for any Borel set $B \subset K$.*
4. *We have $\mu\big(\varphi_I(K)\big) = p_I$ for any multi-index I.*
5. *For any Borel set $B \subset K$ we have*

$$\mu(B) = \inf\Big\{\sum_{i=1}^{\infty} p_{I_i} : B \subset \bigcup_{i=1}^{\infty}{}^{**} \varphi_{I_i}(K)\Big\}.$$

Since SOSC and COSC are both stronger than MSC and one of them will be always assumed in this paper, the statements of this lemma will often be tacitly used. Sometimes, for example, we shall even handle the above almost disjoint sets as disjoint sets and often consider Fact 5 as the definition of self-affine/self-similar measures.

LEMMA 2.12. *Let $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ be a self-affine set with the measure separation property (or in particular with the SSC or SOSC or COSC) and let μ be a self-affine measure on it. Then for every $\varepsilon > 0$ and for every Borel set $B \subset K$ with positive μ-measure there exists an elementary piece $a(K)$ of K of arbitrarily large generation such that $\mu\big(B \cap a(K)\big) > (1-\varepsilon)\mu\big(a(K)\big)$.*

Proof. Since $\mu(B) > 0$, using Fact 5, B can be covered by countably many elementary pieces $\varphi_{I_i}(K)$ ($i \in \mathbb{N}$) such that

$$(1+\varepsilon)\,\mu(B) > \sum_i \mu\big(\varphi_{I_i}(K)\big).$$

By subdividing the elementary pieces if necessary, we can suppose that each is of large generation.

If there exists an $i \in \mathbb{N}$ such that

$$(1+\varepsilon)\,\mu\big(B \cap \varphi_{I_i}(K)\big) > \mu\big(\varphi_{I_i}(K)\big)$$

then we can choose φ_{I_i} as a.

Otherwise we have $(1+\varepsilon)\,\mu\big(B \cap \varphi_{I_i}(K)\big) \le \mu\big(\varphi_{I_i}(K)\big)$ for each $i \in \mathbb{N}$, hence

$$(1+\varepsilon)\,\mu(B) = (1+\varepsilon)\,\mu\Big(\bigcup_i B \cap \varphi_{I_i}(K)\Big) \le \sum_i (1+\varepsilon)\mu\big(B \cap \varphi_{I_i}(K)\big) \le \sum_i \mu\big(\varphi_i(K)\big),$$

contradicting the above inequality. □

Lemma 2.13. *Let $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ be a self-affine set with the measure separation property (or in particular with the SSC or SOSC or COSC) and let μ be a self-affine measure on it. Then for any Borel set $B \subset K$ and $\varepsilon > 0$ there exist countably many pairwise almost disjoint elementary pieces $a_i(K)$ such that $\mu\big(B \cap a_i(K)\big) > (1-\varepsilon)\mu\big(a_i(K)\big)$ and $\mu\big(B \setminus \cup_i^{**} a_i(K)\big) = 0$.*

Proof. The elementary pieces $a_i(K)$ will be chosen by greedy algorithm. In the n^{th} step $(n = 0, 1, 2, \ldots)$ we choose the largest elementary piece $a_n(K)$ such that $\mu\big(a_n(K) \cap a_i(K)\big) = 0 \quad (0 \le i < n)$ and $\mu\big(B \cap a_n(K)\big) > (1-\varepsilon)\mu\big(a_n(K)\big)$. If there is no such $a_n(K)$ then the procedure terminates.

We claim that $\mu\big(B \setminus \cup_i^{**} a_i(K)\big) = 0$. Suppose that $\mu\big(B \setminus \cup_i^{**} a_i(K)\big) > 0$. Then by Lemma 2.12 there exists an elementary piece $a(K)$ such that

$$\mu\big((B \setminus \cup_i^{**} a_i(K)) \cap a(K)\big) > (1-\varepsilon)\mu\big(a(K)\big).$$

Then $\mu\big(B \cap a(K)\big) > (1-\varepsilon)\mu\big(a(K)\big)$ but $a(K)$ was not chosen in the procedure. This could happen only if $a(K)$ intersects a chosen elementary piece $a_i(K)$ in a set of positive measure. But then either $a_i(K) \supset a(K)$ or $a_i(K) \subset a(K)$, which are both impossible. □

2.4. *Self-affine Sierpiński sponges.*

Definition 2.14. By *self-affine Sierpiński sponge* we mean self-affine sets of the following type. Let $n, r \in \mathbb{N}$, $m_1, m_2, \ldots, m_n \ge 2$ integers, M be the linear transformation given by the diagonal $n \times n$ matrix

$$M = \begin{pmatrix} m_1 & & 0 \\ & \ddots & \\ 0 & & m_n \end{pmatrix},$$

and let

$$D = \{d_1, \ldots, d_r\} \subset \{0, 1, \ldots, m_1 - 1\} \times \ldots \times \{0, 1, \ldots, m_n - 1\}$$

be given. Let $\varphi_j(x) = M^{-1}(x + d_j)$ $(j = 1, \ldots, r)$. Then the self-affine set $K(M, D) = K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ is a Sierpiński sponge.

We can also define the self-affine Sierpiński sponge as

$$K = K(M, D) = \left\{ \sum_{k=1}^{\infty} M^{-k} \alpha_k \ : \ \alpha_1, \alpha_2, \ldots \in D \right\},$$

or equivalently K is the unique compact set in $\mathbb{R}^n$ (in fact, in $[0,1]^n$) such that

$$M(K) = K + D = \bigcup_{j=1}^{r} K + d_j;$$

that is,

$$K = M^{-1}(K) + M^{-1}(D).$$

By iterating the last equation we get

$$\begin{aligned} K &= M^{-k}(K) + M^{-k}(D) + M^{-k+1}(D) + \ldots + M^{-1}(D) \\ &= \bigcup_{\alpha_1,\ldots,\alpha_k \in D} M^{-k}(K) + M^{-k}\alpha_k + \ldots + M^{-1}\alpha_1. \end{aligned}$$

Note that the k-th generation elementary pieces of K are the sets of the form $M^{-k}(K)+M^{-k}(\alpha_k)+\ldots+M^{-1}(\alpha_1)$ $(\alpha_1,\ldots,\alpha_k \in D)$ and the only 0-th generation elementary piece of K is K itself.

Definition 2.15. By the *natural* probability measure on a self-affine sponge $K = K(M,D)$ we shall mean the self-affine measure on K obtained by using equal weights $p_j = \frac{1}{r}$ $(j = 1,\ldots,r)$.

Since the first generation elementary pieces of K are translates of each other (in fact, so are the k-th generation elementary pieces), this is indeed the most natural self-affine measure on K. Using (5) of Facts 2.11 we get that

$$\mu(B) = \inf\left\{\sum_{i=1}^{\infty}\mu(S_i) : B \subset \cup_{i=1}^{\infty} S_i,\ S_i \text{ is an elementary piece of } K\ (i \in \mathbb{N})\right\}$$

for every Borel set $B \subset K$.

Let $\tilde{\mu}$ be the $\mathbb{Z}^n$-invariant extension of μ to $\mathbb{R}^n$; that is, for any Borel set $B \subset \mathbb{R}^n$ let

$$\tilde{\mu}(B) = \sum_{t\in\mathbb{Z}^n} \mu\big((B+t)\cap K\big).$$

One can check that

$$\tilde{\mu}\big(M^l(H)+v\big) = r^l\mu(H) \quad \text{for any } H \subset K \text{ Borel set, } v \in \mathbb{Z}^n,\ l = 0,1,2,\ldots. \tag{3}$$

Lemma 2.16. *Let $m_1,\ldots,m_n \geq 2$ and M be like in Definition 2.14. Let $t \in \mathbb{R}^n$ be such that $\|M^k t\| > 0$ for every $k = 0,1,2,\ldots$, where $\|.\|$ denotes the distance from $\mathbb{Z}^n$.*

Then there exists infinitely many $k \in \mathbb{N}$ such that $\|M^k t\| > \frac{1}{2\max(m_1,\ldots,m_n)}$.

Proof. This lemma immediately follows from the following clear fact:

$$\|u\| \leq \frac{1}{2\max(m_1,\ldots,m_n)} \Longrightarrow \|Mu\| \geq \min(m_1,\ldots,m_n)\|u\| \geq 2\|u\|.$$

□

2.5. *Invariant extension of measures to larger sets.*

LEMMA 2.17. *Suppose that the group G acts on a set X, $\mathcal{M}$ is a G-invariant σ-algebra on X, $A \in \mathcal{M}$, $\mathcal{M}_A = \{B \in \mathcal{M} : B \subset A\}$ and μ is a measure on $(A, \mathcal{M}_A)$.*

Then the following two statements are equivalent:

(i) $\mu\big(g(B)\big) = \mu(B)$ *whenever* $g \in G$ *and* $B, g(B) \in \mathcal{M}_A$.

(ii) *There exists a G-invariant measure $\tilde{\mu}$ on $(X, \mathcal{M})$ such that $\tilde{\mu}(B) = \mu(B)$ for every $B \in \mathcal{M}_A$.*

Proof. The implication $(ii) \Rightarrow (i)$ is obvious. For proving the other implication we construct $\tilde{\mu}$ as follows.

If H is a set of the form

$$H = \cup_{i=1}^{*\infty} B_i, \text{ where } g_1, g_2, \ldots \in G \text{ and } g_1(B_1), g_2(B_2), \ldots \in \mathcal{M}_A \tag{4}$$

then let

$$\tilde{\mu}(H) = \sum_{i=1}^{\infty} \mu\big(g_i(B_i)\big)$$

and let $\tilde{\mu}(H) = \infty$ if $H \in \mathcal{M}$ cannot be written in the above form.

First we check that $\tilde{\mu}$ is well defined; that is, if we have (4) and $H = \cup_{j=1}^{*\infty} C_j$, $h_1, h_2, \ldots \in G$ and $h_1(C_1), h_2(C_2), \ldots \in \mathcal{M}_A$ then

$$\sum_{i=1}^{\infty} \mu\big(g_i(B_i)\big) = \sum_{j=1}^{\infty} \mu\big(h_j(C_j)\big). \tag{5}$$

Using that $B_i \subset H = \cup_{j=1}^{*\infty} C_j$ we get that $g_i(B_i) = g_i(\cup_{j=1}^{*\infty} B_i \cap C_j) = \cup_{j=1}^{*\infty} g_i(B_i \cap C_j)$ and so

$$\sum_{i=1}^{\infty} \mu\big(g_i(B_i)\big) = \sum_{i=1}^{\infty} \mu\big(\cup_{j=1}^{*\infty} g_i(B_i \cap C_j)\big) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu\big(g_i(B_i \cap C_j)\big),$$

and similarly

$$\sum_{j=1}^{\infty} \mu\big(h_j(C_j)\big) = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \mu\big(h_j(B_i \cap C_j)\big).$$

Thus, using condition (i) for $B = g_i(B_i \cap C_j)$ and $g = h_j g_i^{-1}$, we get (5).

Using the freedom in (4) and that whenever $H \in \mathcal{M}$ can be written in the form (4) then the same is true for any $H \supset H' \in \mathcal{M}$, it is easy to check that $\tilde{\mu}$ is a G-invariant measure on $(X, \mathcal{M})$ such that $\tilde{\mu}(B) = \mu(B)$ for every $B \in \mathcal{M}_A$. □

We will need only the following special case of this lemma.

LEMMA 2.18. *Let μ be a Borel measure on a Borel set $A \subset \mathbb{R}^n$ and let G be a group of affine transformations of $\mathbb{R}^n$. Suppose that*

$$\mu\big(g(B)\big) = \mu(B) \text{ whenever } g \in G,\ B, g(B) \subset A \text{ and } B \text{ is a Borel set.} \tag{6}$$

Then there exists a G-invariant Borel measure $\tilde{\mu}$ on $\mathbb{R}^n$ such that $\tilde{\mu}(B) = \mu(B)$ for any $B \subset A$ Borel set. □

Remark 2.19. The extension we get in the above proof do not always give the measure we expect – it may be infinity for too many sets. For example, if $A \subset \mathbb{R}$ is a Borel set of first category with positive Lebesgue measure, G is the group of translations and μ is the restriction of the Lebesgue measure to A then the Lebesgue measure itself would be the natural translation invariant extension of μ, however the extension $\tilde{\mu}$ as defined in the proof is clearly infinity for every Borel set of second category.

Definition 2.20. Let μ be a Borel measure on a compact set K. We say that μ is isometry invariant if given any isometry g and a Borel set $B \subset K$ such that $g(B) \subset K$, then $\mu(B) = \mu\big(g(B)\big)$.

This definition makes sense since (by Lemma 2.18) exactly the isometry invariant measures on K can be extended to be isometry invariant measures on $\mathbb{R}^n$ in the usual sense.

As an illustration of Lemma 2.18 we mention the following special case with a peculiar consequence.

Lemma 2.21. *Let $A \subset \mathbb{R}^n$ ($n \in \mathbb{N}$) be a Borel set such that $A \cap (A + t)$ is at most countable for any $t \in \mathbb{R}^n \setminus \{0\}$. Then any continuous Borel measure μ on A (continuous here means that the measure of any singleton is zero) can be extended to a translation invariant Borel measure on $\mathbb{R}^n$.* □

Note that although the condition that $A \cap (A + t)$ is at most countable for any $t \in \mathbb{R}^n \setminus \{0\}$ seems to imply that A is very small, such a set can be still fairly large. For example there exists a compact set $C \subset \mathbb{R}$ with Hausdorff dimension 1 such that $C \cap (C + t)$ contains at most one point for any $t \in \mathbb{R} \setminus \{0\}$ [**14**]. Combining this with Lemma 2.21 we get the following.

Corollary 2.22. *There exists a compact set $C \subset \mathbb{R}$ with Hausdorff dimension 1 such that any continuous Borel measure μ on C can be extended to a translation invariant Borel measure on $\mathbb{R}$.* □

2.6. *Some more lemmas.*

The following simple lemmas might be known but for completeness (and because it is easier to prove them than to find them) we present their proof.

Recall that the support of a measure is the smallest closed set with measure zero complement.

Lemma 2.23. *Let μ be a finite Borel measure on $\mathbb{R}^n$ with compact support K. Then for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$|u| \geq \varepsilon \Longrightarrow \mu\big(K \cap (K + u)\big) \leq (1 - \delta)\mu(K).$$

Proof. We prove by contradiction. Assume that there exists an $\varepsilon > 0$ and a sequence $u_1, u_2, \ldots \in \mathbb{R}^n$ such that $|u_n| \geq \varepsilon$ (for every $n \in \mathbb{N}$) and $\mu\big(K \cap (K + u)\big) \to \mu(K) > 0$ $(n \to \infty)$. By omitting some (at most finitely many) zero terms we can guarantee that every u_n is in the compact annulus $\{x : \varepsilon \leq |x| \leq \operatorname{diam}(K)\}$ (where diam denotes the diameter). Hence, by taking a subsequence, we can suppose that (u_n) converges, say to u. Since $K \cap (K+u)$ is a proper compact subset of K (since K is compact and $u \neq 0$, $K+u \supset K$ is impossible) and K is the support of μ, we must have $\mu(K) > \mu\big(K \cap (K+u)\big) = \mu(K+u)$.

It is well known (see e.g. [**29**], 2.18. Theorem) that any finite Borel measure is outer regular in the sense that the measure of any Borel set is the infimum of the measures of the open sets that contain the Borel set. Thus $\mu(K+u) < \mu(K)$ implies that there exists an open set $G \supset K+u$ such that $\mu(G) < \mu(K)$. Then whenever $|u_n - u|$ is less than the (positive) distance between K and the complement of G, G contains $K+u_n$ and so $\mu(K) > \mu(G) \geq \mu(K+u_n)$. This is a contradiction since $u_n \to u$ and $\mu(K+u_n) = \mu\big(K \cap (K+u_n)\big) \to \mu(K)$. □

LEMMA 2.24. *Let μ be a probability Borel measure on a compact set $K \subset \mathbb{R}^d$ such that any nonempty relative open subset of K has positive μ measure. Then if the sequence (g_n) of affine maps converges to an affine map g and $\mu\big(g_n(K) \cap K\big) \to 1$ then $\mu\big(g(K) \cap K\big) = 1$. Moreover, $K \subset g(K)$.*

Proof. Suppose that $\mu\big(g(K) \cap K\big) = q < 1$. Let $g(K)_\varepsilon$ denote the ε-neighborhood of $g(K)$. Since $\bigcap_{n=1}^{\infty} (g(K)_{1/n} \cap K) = g(K) \cap K$ and μ is a finite measure we have $\mu\big(g(K)_{1/n} \cap K\big) \to \mu\big(g(K) \cap K\big) = q$. Thus there exists an $\varepsilon > 0$ for which $\mu\big(g(K)_\varepsilon \cap K\big) \leq \frac{1+q}{2} < 1$. Since g_n converges uniformly on K, for n large enough we have $g_n(K) \subset g(K)_\varepsilon$ and so $\mu\big(g_n(K) \cap K\big) \leq \frac{1+q}{2}$, contradicting $\mu\big(g_n(K) \cap K\big) \to 1$. Therefore we proved that $\mu\big(g(K) \cap K\big) = 1$.

Then $K \setminus g(K)$ is relative open in K and has μ measure zero, so it must be empty, therefore $K \subset g(K)$. □

3. *Self-affine sets with the strong separation condition*

PROPOSITION 3.1. *For any self-affine set $K \subset \mathbb{R}^d$ with the strong separation condition there exists an open neighborhood $U \subset \mathcal{A}_K$ of the identity map such that for any $g \in U$,*

$$g(K) \supset K \iff g = \textit{identity}.$$

Proof. Let n denote the dimension of the affine span of K.

We shall prove that there exists a small open neighborhood $V \subset \mathcal{A}_K$ of the identity map such that for any $g \in V$ we have $g(K) \subset K \iff g = \text{identity}$. This would be enough since then for any $g \in V$ we get $K \subset g^{-1}(K) \iff g = \text{identity}$, therefore $U = V^{-1} = \{g^{-1} : g \in V\}$ has all the required properties.

Similarly as in the proof of Lemma 2.6, choose $n+1$ elementary pieces $\varphi_{I_1}(K), \dots, \varphi_{I_{n+1}}(K)$ of K so that if we pick one point from the convex hull of each of them then we get a maximal collection of affine independent points in the affine span of K.

Let $d = \min_{1 \leq i \leq n+1} \operatorname{dist}(\varphi_{I_i}(K), K \setminus \varphi_{I_i}(K))$, then $d > 0$. Let V be a so small neighborhood of the identity map that $\operatorname{dist}(x, g(x)) < d$ for any $g \in V$ and $x \in K$.

Let $g \in V$ and $g(K) \subset K$. Then, by the definition of d and V we have $g(\varphi_{I_i}(K)) \subset \varphi_{I_i}(K)$ for every $1 \leq i \leq n+1$. Then the convex hulls of these elementary pieces are also mapped into themselves. Since each of these convex hulls is homeomorphic to a ball, by Brouwer's fixed point theorem we get a fixed point of g in each of these elementary pieces. So we obtained $n+1$ fixed points of g such that their affine span is exactly the affine span of K. Since g is an affine map, the set of its fixed points form an affine subspace, thus the set of fixed points of g contains the affine span of K. Since $g \in \mathcal{A}_K$, g is defined exactly on the affine span of K, therefore g must be the identity map. □

THEOREM 3.2. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-affine set satisfying the strong separation condition and let μ be a self-affine measure on K. Then there exists a $c < 1$ and an open neighborhood $U \subset \mathcal{A}_K$ of the identity map such that $g \in U \setminus \{identity\} \Longrightarrow \mu\big(K \cap g(K)\big) < c$.*

Proof. According to our definition of self-affine set (see Definition 2.4) there exists a norm in which every φ_i is contractive. Let $\operatorname{dist}_\varphi$ denote the metric determined by this norm.

Using Proposition 3.1 we can choose a small open neighborhood $U \subset \mathcal{A}_K$ of the identity map such that even in the closure of U the only affine map g for which $g(K)$ contains K is the identity map and so that

$$\operatorname{dist}_\varphi(x, g(x)) < 1 \text{ for any } g \in U \text{ and } x \in K. \tag{7}$$

Since $\mathcal{A}_K$ is locally compact, we may also assume that the closure of U is compact.

We claim that we can choose an even smaller open neighborhood $V \subset U$ of the identity map such that $\varphi_i^{-1} \circ V \circ \varphi_i \subset U$ for $i = 1, \ldots, r$ and that $g(\varphi_i(K)) \cap \varphi_j(K) = \emptyset$ for any $i \neq j$ and $g \in V$. Indeed, the first property can be satisfied since $\mathcal{A}_K$ is a topological group and those g's for which the second property do not hold are far from the identity map.

Now we claim that there exists a $c < 1$ such that $g \in \overline{U} \setminus V \Longrightarrow \mu\big(g(K) \cap K\big) < c$. Suppose that there exists a sequence $(g_n) \subset \overline{U} \setminus V$ such that $\mu\big(K \cap g_n(K)\big) \to 1$. Since $\overline{U} \setminus V$ is compact there exists a subsequence g_{n_i} such that $g_{n_i} \to h \in \overline{U} \setminus V$. By Lemma 2.24 this implies that $h(K) \supset K$ but in $\overline{U} \setminus V$ there is no such affine map h.

We prove that this U and this c have the required properties; that is, $g \in U \setminus \{\text{identity}\} \Longrightarrow \mu\big(K \cap g(K)\big) < c$.

If $g \in \overline{U} \setminus V$ then we are already done, so suppose that $g \in V \setminus \{\text{identity}\}$. Let F denote the set of fixed points of g.

The heuristics of the remaining part of the proof is the following. The affine map g moves K too slightly. We zoom in on small elementary pieces $a(K)$ of K so that each $g(a(K))$ intersects only $a(K)$ in K, but g moves $a(K)$ far enough (compared to its size). Technically this second requirement means that $a^{-1} \circ g \circ a \in U \setminus V$, so we can use the $g \in U \setminus V$ case for the elementary piece $a(K)$. We find such an elementary piece around each point of K that is not a fixed point of g, and so we get a partition of $K \setminus F$ into elementary pieces with the above property. Finally, by adding up the estimates for these elementary pieces we derive $\mu\big(g(K) \cap K\big) < c$.

Claim 3.3. *For any $x \in K \setminus F$ there exists a largest elementary piece $\varphi_{I_x}(K)$ of K that contains x and for which $\varphi_{I_x}^{-1} \circ g \circ \varphi_{I_x} \in U \setminus V$.*

Proof. Let $(i_1, i_2, \ldots)$ be the sequence of indices for which

$$\{x\} = \bigcap_{n=1}^{\infty} (\varphi_{i_1} \circ \varphi_{i_2} \circ \ldots \circ \varphi_{i_n})(K),$$

and let $I_n = (i_1, \ldots, i_n)$. Since $g \in V$, we have $\varphi_{i_1}^{-1} \circ g \circ \varphi_{i_1} \in U$ by the definition of V. If for some n we have $\varphi_{I_n}^{-1} \circ g \circ \varphi_{I_n} \in V$ then by the definition of V we have

$$\varphi_{I_{n+1}}^{-1} \circ g \circ \varphi_{I_{n+1}} = \varphi_{i_{n+1}}^{-1} \circ \varphi_{I_n}^{-1} \circ g \circ \varphi_{I_n} \circ \varphi_{i_{n+1}} \in U.$$

Therefore it is enough to find an n such that $\varphi_{I_n}^{-1} \circ g \circ \varphi_{I_n} \notin V$ since then taking the smallest such n, $I_x = I_n$ has the desired property. Letting $y_n = \varphi_{I_n}^{-1}(x)$ we have $y_n \in K$ (since $\{x\} = \bigcap_{n=1}^{\infty} \varphi_{I_n}(K)$) and $(\varphi_{I_n}^{-1} \circ g \circ \varphi_{I_n})(y_n) = \varphi_{I_n}^{-1}(g(x))$.

Since x is not a fixed point of g, for n large enough we have

$$\mathrm{dist}_{\varphi}\big(g(x), \varphi_{I_n}(K)\big) > \frac{\mathrm{dist}_{\varphi}(g(x), x)}{2} \stackrel{\mathrm{def}}{=} t > 0.$$

Recall that dist_{φ} was defined as a metric in which every φ_i is contractive. Hence for each i there exists an $\alpha_i < 1$ such that $\mathrm{dist}_{\varphi}(\varphi_i(a), \varphi_i(b)) \leq \alpha_i \cdot \mathrm{dist}_{\varphi}(a, b)$ for any a, b. Then, using the multi-index notation $\alpha_{I_n} = \alpha_{i_1} \cdot \ldots \cdot \alpha_{i_n}$, we clearly have $\mathrm{dist}_{\varphi}(\varphi_{I_n}(a), \varphi_{I_n}(b)) \leq \alpha_{I_n} \cdot \mathrm{dist}_{\varphi}(a, b)$ for any a, b. Then $\mathrm{dist}_{\varphi}\big(\varphi_{I_n}^{-1}(g(x)), K\big) > t/\alpha_{I_n}$, hence $\mathrm{dist}_{\varphi}\big((\varphi_{I_n}^{-1} \circ g \circ \varphi_{I_n})(y_n), K\big) > t/\alpha_{I_n}$, which is bigger than 1 if n is large enough. Thus for n large enough, $\varphi_{I_n}^{-1} \circ g \circ \varphi_{I_n}$ is not in V, since it is not even in U by (7). □

Claim 3.4. *For any $x \in K \setminus F$ we have $g(\varphi_{I_x}(K)) \cap K \subset \varphi_{I_x}(K)$, where $I_x = I_n = (i_1, \ldots, i_n)$ is the multi-index we got in Claim 3.3.*

Proof. Let $k \in \{0, 1, \ldots, n-1\}$ be arbitrary and let $I_k = (i_1, \ldots, i_k)$. Then $\varphi_{I_k}^{-1} \circ g \circ \varphi_{I_k} \in V$, hence for any $l \neq i_{k+1}$ we have $(\varphi_{I_k}^{-1} \circ g \circ \varphi_{I_k} \circ \varphi_{i_{k+1}})(K) \cap \varphi_l(K) = \emptyset$, which is the same as $(g \circ \varphi_{I_{k+1}})(K) \cap (\varphi_{I_k} \circ \varphi_l)(K) = \emptyset$ $(l \neq i_{k+1})$. Since $(g \circ \varphi_{I_n})(K) \subset (g \circ \varphi_{I_{k+1}})(K)$, this implies that

$$(g \circ \varphi_{I_n})(K) \cap (\varphi_{I_k} \circ \varphi_l)(K) = \emptyset \quad (k \in \{0, 1, \ldots, n-1\},\ l \neq i_{k+1}).$$

Since $K \setminus \varphi_{I_n}(K) = \cup_{k=0}^{n-1} \cup_{l \neq i_{k+1}} (\varphi_{I_k} \circ \varphi_l)(K)$, this implies that $g(\varphi_{I_n}(K)) \cap K \subset \varphi_{I_n}(K)$. □

The elementary pieces $\{\varphi_{I_x}(K) : x \in K \setminus F\}$ clearly cover $K \setminus F$. Since for any $x \neq y$ we have $\varphi_{I_x}(K) \cap \varphi_{I_y}(K) = \emptyset$ or $\varphi_{I_x}(K) \subset \varphi_{I_y}(K)$ or $\varphi_{I_x}(K) \supset \varphi_{I_y}(K)$, one can choose a

$$K \setminus F \subset \bigcup_{i=1}^{\infty}{}^{*} \varphi_{J_i}(K) \tag{8}$$

countable disjoint subcover. By Claim 3.4 we have

$$g(\varphi_{J_i}(K)) \cap K \subset \varphi_{J_i}(K). \tag{9}$$

Since g is not the identity map (of the affine span of K) and F is the set of fixed points of the affine map g, the dimension of the affine subspace F is smaller than the dimension of the affine span of K, and so we cannot have $g(F) \supset K$. By Lemma 2.6 this implies that $\mu\big(g(F) \cap K\big) = 0$. Using this last equation, (8), (9), and finally the definition of a self-affine measure we get that

$$\begin{aligned}
\mu\big(g(K) \cap K\big) &\le \mu\big(g(F) \cap K\big) + \mu\big(g(K \setminus F) \cap K\big) = \mu\big(g(K \setminus F) \cap K\big) \\
&\le \mu\Big(g\Big(\bigcup_{i=1}^{\infty}{}^{*} \varphi_{J_i}(K)\Big) \cap K\Big) = \sum_{i=1}^{\infty} \mu\Big(g\big(\varphi_{J_i}(K)\big) \cap K\Big) \\
&= \sum_{i=1}^{\infty} \mu\Big(g\big(\varphi_{J_i}(K)\big) \cap \varphi_{J_i}(K)\Big) = \sum_{i=1}^{\infty} \mu\Big(\varphi_{J_i}\big((\varphi_{J_i}^{-1} \circ g \circ \varphi_{J_i})(K) \cap K\big)\Big) \\
&= \sum_{i=1}^{\infty} p_{J_i}\, \mu\Big((\varphi_{J_i}^{-1} \circ g \circ \varphi_{J_i})(K) \cap K\Big).
\end{aligned}$$

Since $\varphi_{J_i}^{-1} \circ g \circ \varphi_{J_i} \in U \setminus V$, the measures in the last expression are less than c. Thus $\mu\big(g(K) \cap K\big) < \sum p_{J_i} \cdot c = \sum \mu\big(\varphi_{J_i}(K)\big) \cdot c = \mu\big(\bigcup_i^* \varphi_{J_i}(K)\big) \cdot c = c$, which completes the proof. $\square$

THEOREM 3.5. *Let $K \subset \mathbb{R}^d$ be a self-affine set with the strong separation condition and let μ be a self-affine measure on K. Then there exists a constant $c < 1$ such that for any isometry g we have $\mu\big(K \cap g(K)\big) < c$ unless $g(K) = K$.*

Proof. Suppose that $g_n \in \mathcal{I}_K$ (that is, g_n is an isometry of the affine span of K) such that $g_n(K) \neq K$ $(n \in \mathbb{N})$ and $\mu\big(K \cap g_n(K)\big) \to 1$. We can clearly assume that $K \cap g_n(K) \neq \emptyset$ for each n and so the whole sequence (g_n) is in a compact subset of $\mathcal{I}_K$. Thus, after choosing a subsequence if necessary, we can also assume that g_n converges to an $h \in \mathcal{I}_K$. By Lemma 2.24 we must have $K \subset h(K)$. It is well known and not hard to prove that no compact set in $\mathbb{R}^d$ can have an isometric proper subset, so $K \subset h(K)$ implies that $h(K) = K$.

Applying Theorem 3.2 we get a $c < 1$ and an open neighborhood $U \subset \mathcal{A}_K$ of the identity such that $g \in U \setminus \{\text{identity}\} \Longrightarrow \mu\big(K \cap g(K)\big) < c$.

Since $g_n \to h$ we get $g_n \circ h^{-1} \to$ identity. Let n be large enough to have $g_n \circ h^{-1} \in U$ and $\mu\big(K \cap g_n(K)\big) > c$. Since $g_n(K) \neq K$ but $h(K) = K$ we cannot have $g_n = h$ and so $g_n \circ h^{-1} \in U \setminus \{\text{identity}\}$. Then, by the previous paragraph, we get $\mu\big(K \cap g_n(K)\big) < c$, contradicting $\mu\big(K \cap g_n(K)\big) > c$. $\square$

4. *Self-similar sets with the strong separation property*

Our first goal in this section is to prove the following theorem.

THEOREM 4.1. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set satisfying the strong separation condition and μ be a self-similar measure on it. There exists $c < 1$ such that for every similitude g either $\mu\big(g(K) \cap K\big) < c$ or $K \subset g(K)$.*

Now, for the sake of transparency we outline the proof. At first we need a new notation.

From S_K we excluded those similarity maps which map everything to a single point. So let S_K^* be the metric space of all degenerate and all non-degenerate similarity maps in the affine span A_K of K; that is,

$$S_K^* = S_K \cup \{f \mid f : A_K \to \{y\},\, y \in A_K\}. \tag{10}$$

First we show that there exists a compact set $\mathcal{G} \subset S_K^*$ of similarity maps such that for every $g' \in S_K^*$ there exists $g \in \mathcal{G}$ for which $g'(K) \cap K = g(K) \cap K$. Then clearly it suffices to prove the theorem for $g \in \mathcal{G}$. (It is easy to see that no such compact set $\mathcal{G}$ in $\mathcal{S}_K$ exists.)

Let μ_H be a constant multiple of Hausdorff measure of appropriate dimension so that $\mu_H(K) = 1$. The restriction of this measure to K is a self-similar measure. Let us consider those $h \in \mathcal{G}$ for which $K \subset h(K)$ holds. Using Hausdorff measures and Theorem 3.2 we prove that there are only finitely many such h, and also that the theorem holds in small neighbourhoods of each such h for the measure μ_H. The maximum of the corresponding finitely many values c is still strictly smaller than 1. Let us now cut these small neighbourhoods out of $\mathcal{G}$. Using upper semicontinuity of our measure (Lemma 2.24) we produce a $c < 1$ such that for the remaining similarity maps g we have $\mu_H\big(g(K) \cap K\big) < c$. Then clearly the same holds for all elements of $\mathcal{G}$, possibly with a larger $c < 1$, finishing the proof for the measure μ_H.

Applying the theorem for μ_H, and also in a small open neighbourhood U of the identity for every self-similar measure μ, we show that if $h \in \mathcal{G}$, $K \subset h(K)$, and g is in a small neighbourhood of h then $\mu\big(g(K) \cap K\big) < c$. Then the same argument as above (using upper semicontinuity) yields the theorem, possibly with a larger constant again.

Proposition 4.2. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set satisfying the strong separation condition. Then there exists a compact set $\mathcal{G} \subset S_K^*$ such that for every similarity map $g' \in S_K^*$ there is a $g \in \mathcal{G}$ for which $g'(K) \cap K = g(K) \cap K$ holds.*

Proof. Let D denote the diameter of K, let $\delta = \min_{1 \le i < j \le r} \operatorname{dist}(\varphi_i(K), \varphi_j(K))$ and let

$$\mathcal{G} = \{g \in S_K^* : g(K) \cap K \neq \emptyset, \text{ the similarity ratio of } g \text{ is at most } D/\delta\} \cup \{g_0\},$$

where $g_0 \in S_K^*$ is an arbitrary fixed similarity map such that $g(K) \cap K = \emptyset$. It is easy to check that $\mathcal{G} \subset S_K^*$ is compact.

Let $g' \in S_K^*$. If $g' \in \mathcal{G}$ or $g'(K) \cap K = \emptyset$ then we can choose $g = g'$ or $g = g_0$, respectively. So we can suppose that $g'(K) \cap K \neq \emptyset$ and the similarity ratio of g' is greater than D/δ. Then the minimal distance between the first generation elementary pieces $g'(\varphi_j(K))$ of $g'(K)$ is larger than D. So there exists φ_i such that $g'(K) \cap K = g'(\varphi_i(K)) \cap K$. Therefore g' can be replaced by $g' \circ \varphi_i$, which has similarity ratio α_i times smaller than the similarity ratio of g', where α_i denotes the similarity ratio of φ_i. Since $\max(\alpha_1, \ldots, \alpha_r) < 1$, this way in finitely many steps we get a g with similarity ratio at most D/δ such that $g(K) \cap K = g'(K) \cap K \neq \emptyset$, which completes the proof. □

PROPOSITION 4.3. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set satisfying the strong separation condition.*

(i) Then $\{g \in \mathcal{S}_K : g(K) \supset K\}$ is discrete in $\mathcal{S}_K$, hence countable, and also closed in $\mathcal{S}_K$.

(ii) Let μ_H be a constant multiple of Hausdorff measure of appropriate dimension so that $\mu_H(K) = 1$. There exists $c < 1$ such that for every similitude g either $\mu_H\big(g(K) \cap K\big) < c$ or $K \subset g(K)$.

Proof. By Lemma 2.24 $\{g \in \mathcal{S}_K : g(K) \supset K\}$ is closed. Since every discrete subset of a subspace of $\mathbb{R}^{d^2+d}$ is countable, in order to prove (i) it is enough to prove that $\{g \in \mathcal{S}_K : g(K) \supset K\}$ is discrete.

Let ε be a positive number to be chosen later, and h be a similitude for which $K \subset h(K)$. Denote by K_δ the δ-neighbourhood of K. As $\mu_H\big(h(K)\big)$ is finite, there is a small $\delta > 0$ such that $\mu_H\big(K_\delta \cap (h(K) \setminus K)\big) < \varepsilon$. Applying Theorem 3.2 to K and μ_H we obtain an open neighbourhood $U \subset \mathcal{A}_K$ and a constant c_H. There exists an open neighbourhood $W_\varepsilon \subset \mathcal{S}_K$ of the identity such that

(a) $W_\varepsilon = W_\varepsilon^{-1} \subset U$,

(b) $\operatorname{dist}(g(x), x) < \delta$ for every $x \in K$,

(c) $\mu_H\big(g(B)\big) \le (1+\varepsilon)\mu_H(B)$ for every $g \in W_\varepsilon$ and Borel set B,

where for (c) we use that a similitude of ratio α multiplies the s-dimensional Hausdorff measure by α^s.

Let $g \in W_\varepsilon h$ and $g \neq h$. Clearly $W_\varepsilon h$ is an open neighbourhood of h and $g \circ h^{-1}$, $h \circ g^{-1} \in W_\varepsilon \setminus \{\text{identity}\}$, and $(h \circ g^{-1})(K) \subset K_\delta$. Hence

$$\begin{aligned}
\mu_H\big(K \cap g(K)\big) &\le (1+\varepsilon)\mu_H\big((h \circ g^{-1})(K \cap g(K))\big) = \\
&= (1+\varepsilon)\mu_H\big((h \circ g^{-1})(K) \cap h(K)\big) = \\
&= (1+\varepsilon)\mu_H\big((h \circ g^{-1})(K) \cap K\big) + (1+\varepsilon)\mu_H\big((h \circ g^{-1})(K) \cap (h(K) \setminus K)\big) \le \\
&\le (1+\varepsilon)c_H + (1+\varepsilon)\mu_H\big(K_\delta \cap (h(K) \setminus K)\big) \le (1+\varepsilon)c_H + (1+\varepsilon)\varepsilon. \quad (11)
\end{aligned}$$

The last expression is clearly smaller than 1 if ε is small enough, so let us fix such an ε. Therefore if $g \in W_\varepsilon h$ and $g \neq h$ then $g(K) \not\supset K$, which shows that $\{g \in \mathcal{S}_K : g(K) \supset K\}$ is discrete finishing the proof of (i).

In order to prove (ii) suppose towards a contradiction that $\sup\{\mu_H(g(K) \cap K) : g \in S_K^*,\ g(K) \not\supset K\} = 1$. Then we also have $\sup\{\mu_H(g(K) \cap K) : g \in \mathcal{G},\ g(K) \not\supset K\} = 1$. Let (g_n) be a convergent sequence in $\mathcal{G}$ so that $g_n(K) \not\supset K$, $\mu_H\big(g_n(K) \cap K\big) \to 1$, $g_n \to h$. Lemma 2.24 yields $h(K) \supset K$, hence $g_n \neq h$. If n is large enough then $g_n \in W_\varepsilon h$ and, by (11), $\mu_H\big(K \cap g_n(K)\big) \le (1+\varepsilon)c_H + (1+\varepsilon)\varepsilon$, contradicting $\mu_H\big(g_n(K) \cap K\big) \to 1$. □

Proof of Theorem 4.1. By Proposition 4.2 we can assume $g \in \mathcal{G}$. Let c_H be the constant yielded by Proposition 4.3 (ii). Fix $h \in \mathcal{G}$ with $h(K) \supset K$. There are only finitely many such h by Proposition 4.3 (i) and the compactness of $\mathcal{G}$.

Let us now apply Lemma 2.12 to the self-similar set $h(K)$, μ_H, $0 < \varepsilon \le 1 - c_H$ and $B = K \subset h(K)$. We obtain φ_I such that

$$\mu_H\big(K \cap h(\varphi_I(K))\big) \ge (1-\varepsilon)\,\mu_H\big(h(\varphi_I(K))\big).$$

Hence Proposition 4.3 (ii) applied to the self-similar set $h(\varphi_I(K))$ and the similitude $(h \circ \varphi_I)^{-1}$ gives $K \supset h(\varphi_I(K))$.

Since $h(\varphi_I(K))$ is open in $h(K)$, it is also open in K and so it can be written as a union of elementary pieces of K. Since $h(\varphi_I(K))$ is compact this implies that $h(\varphi_I(K))$ is a finite union of elementary pieces of K. Let $\varphi_J(K)$ be one of these elementary pieces. So $\varphi_J(K) \subset h(\varphi_I(K)) \subset K \subset h(K)$. As $\varphi_J(K)$ is open in K, it is also open in $h(\varphi_I(K))$, hence also in $h(K)$. Therefore $\operatorname{dist}(\varphi_J(K), h(K) \setminus \varphi_J(K)) > 0$, and so for every g that is close enough to h we have

$$\big(g \circ h^{-1}\big)\big(h(K) \setminus \varphi_J(K)\big) \cap \varphi_J(K) = \emptyset.$$

Thus, as $\varphi_J(K) \subset h(K)$, for every such g we have

$$g(K) \cap \varphi_J(K) = \big(g \circ h^{-1}\big)\big(h(K)\big) \cap \varphi_J(K) = \big(g \circ h^{-1}\big)\big(\varphi_J(K)\big) \cap \varphi_J(K).$$

On the other hand, Theorem 3.2 yields that there exists a $c < 1$ such that if g is close enough to h and $g \ne h$ then

$$\mu\big((g \circ h^{-1})(\varphi_J(K)) \cap \varphi_J(K)\big) < c \cdot \mu\big(\varphi_J(K)\big) = c \cdot p_J.$$

Therefore $\mu\big(g(K) \cap \varphi_J(K)\big) = \mu\big((g \circ h^{-1})(\varphi_J(K)) \cap \varphi_J(K)\big) < c \cdot p_J$ and

$$\begin{aligned}\mu\big(g(K) \cap K\big) &= \mu\big(g(K) \cap \varphi_J(K)\big) + \mu\big(g(K) \cap (K \setminus \varphi_J(K))\big) \\ &< c \cdot p_J + 1 - p_J = 1 - (1-c)p_J. \qquad (12)\end{aligned}$$

As we only considered finitely many h's, there exists $c' < 1$ such that if g is close to one of these h's, but distinct from it, then $\mu\big(g(K) \cap K\big) < c'$. This, together with Lemma 2.24 provides a $c'' < 1$ such that for every $g \in \mathcal{G}$ either $\mu\big(g(K) \cap K\big) < c''$ or $g(K) \supset K$. (Just like at the end of the proof of Proposition 4.3.) Finally, by Proposition 4.2 this also holds outside $\mathcal{G}$. □

We will apply this theorem to elementary pieces of K instead of K itself. It is easy to see that the same c works for every elementary piece; that is, we have the following corollary of Theorem 4.1.

COROLLARY 4.4. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set satisfying the strong separation condition and μ be a self-similar measure on it. There exists $c < 1$ such that for every similitude g and every elementary piece $a(K)$ of K either $\mu\big(g(K) \cap a(K)\big) < c \cdot \mu\big(a(K)\big)$ or $a(K) \subset g(K)$.* □

Now we are ready to prove the second main result of this section.

Theorem 4.5. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set satisfying the strong separation condition, μ be a self-similar measure on it, and g be a similitude. Then $\mu\big(g(K) \cap K\big) > 0$ if and only if the interior (in K) of $g(K) \cap K$ is nonempty. Moreover, $\mu\big(int_K(g(K) \cap K)\big) = \mu\big(g(K) \cap K\big)$.*

Proof. If the interior (in K) of $g(K) \cap K$ is nonempty then clearly it is of positive measure, since the measure of every elementary piece is positive.

Let c be the constant given by Corollary 4.4, and let g be a similitude such that $\mu\big(g(K) \cap K\big) > 0$. Applying Lemma 2.13 for $B = g(K) \cap K$ and $\varepsilon = 1 - c$ we obtain countably many disjoint elementary pieces $a_i(K)$ of K such that

$$\mu\big(g(K) \cap a_i(K)\big) = \mu\big((g(K) \cap K) \cap a_i(K)\big) > c \cdot \mu\big(a_i(K)\big) \tag{13}$$

and $\big(g(K) \cap K\big) \setminus \bigcup_i^* a_i(K)$ is of μ-measure zero. By Corollary 4.4, (13) implies that $a_i(K) \subset g(K)$. Since $a_i(K)$ is open in K, it is open in $g(K) \cap K$, so $\bigcup_i^* a_i(K) \subset \mathrm{int}_K(g(K) \cap K)$. Hence

$$\begin{aligned}\mu\big(g(K) \cap K\big) = \mu\big(g(K) \cap K \cap \bigcup_i^* a_i(K)\big) + \mu\big((g(K) \cap K) \setminus \bigcup_i^* a_i(K)\big)\\ = \mu\big(\bigcup_i^* a_i(K)\big) \le \mu\big(\mathrm{int}_K(g(K) \cap K)\big),\end{aligned}$$

proving the theorem. □

As an immediate consequence we get the following.

Corollary 4.6. *Let $K \subset \mathbb{R}^d$ be a self-similar set satisfying the strong separation condition, and let μ_1 and μ_2 be self-similar measure on K. Then for any similitude g of $\mathbb{R}^d$,*

$$\mu_1\big(g(K) \cap K\big) > 0 \Longleftrightarrow \mu_2\big(g(K) \cap K\big) > 0.$$

We also get the following fairly easily.

Corollary 4.7. *Let $K \subset \mathbb{R}^d$ be a self-similar set satisfying the strong separation condition, let A_K be the affine span of K and let μ be a self-similar measure on K. Then the set of those similitudes $g : A_K \to \mathbb{R}^d$ for which $\mu\big(g(K) \cap K\big) > 0$ is countably infinite.*

Proof. It is clear that there exist infinitely many similitudes g such that $\mu\big(g(K) \cap K\big) > 0$ since the elementary pieces of K are similar to K and have positive μ measure.

By Lemma 2.6, $\mu\big(g(K) \cap K\big) > 0$ implies that $g \in \mathcal{S}_K$ and, by Theorem 4.5, that $g(K)$ contains an elementary piece of K. Therefore it is enough to show that for each fixed elementary piece $a(K)$ of K there are only countably many $g \in \mathcal{S}_K$ such that $g(K) \supset a(K)$, which is the same as $(a^{-1} \circ g)(K) \supset K$. By the first part of Proposition 4.3 there are only countably many such $a^{-1} \circ g \in \mathcal{S}_K$, so there are only countably many such $g \in \mathcal{S}_K$. □

From the first part of Proposition 4.3 we get more results about those similarity maps that map a self-similar set into itself. These results will be used in the next section and they are also related to a theorem and a question of Feng and Wang [**8**] as it will be explained before Corollary 4.10.

LEMMA 4.8. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set with strong separation condition. There exists only finitely many similitudes g for which $g(K) \subset K$ holds and $g(K)$ intersects at least two first generation elementary pieces of K.*

Proof. The similarity ratios of these similitudes g are strictly separated from zero. Thus the similarity ratio of their inverses have some finite upper bound, and also $K \subset g^{-1}(K)$ holds. The set of similitudes with the latter property form a discrete and closed set according to the first part of Proposition 4.3.

Those $h \in S_K^*$ similarity maps (cf. (10)) whose similarity ratio is under some fixed bound and for which $h(K) \cap K \neq \emptyset$ holds form a compact set in S_K^* (see proof of Proposition 4.2). Since a discrete and closed subspace of a compact set is finite, the proof is finished. □

THEOREM 4.9. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set with strong separation condition and let λ be a similitude for which $\lambda(K) \subset K$. There exist an integer $k \geq 1$ and multi-indices I, J such that $\lambda^k \circ \varphi_I = \varphi_J$.*

Proof. For every integer $k \geq 1$ there exists a smallest elementary piece $\varphi_I(K)$ which contains $\lambda^k(K)$. For this multi-index I, $(\varphi_I^{-1} \circ \lambda^k)(K)$ is a subset of K and intersects at least two first generation elementary pieces of K. There are only finitely many similitudes with this property according to Lemma 4.8, hence there exist $k < k'$, I, I' such that $\varphi_I^{-1} \circ \lambda^k = \varphi_{I'}^{-1} \circ \lambda^{k'}$. By rearrangement we obtain $\varphi_{I'} \circ \varphi_I^{-1} = \lambda^{k'-k}$ and $\lambda^{k'-k} \circ \varphi_I = \varphi_{I'}$. □

Feng and Wang [**8**, Theorem 1.1 (The Logarithmic Commensurability Theorem)] proved that if $K = \varphi_1(K) \cup \ldots \cup \varphi_r(K)$ is a self-similar set in $\mathbb{R}$ satisfying the open set condition with Hausdorff dimension less than 1 and such that each similarity map φ_i is of the form $\varphi_i(x) = bx + c_i$ with a fixed b and $aK + t \subset K$ for some $a, t \in \mathbb{R}$ then $\log |a| / \log |b| \in \mathbb{Q}$. They also posed the problem (Open Question 2) of generalizing this result to higher dimensions. If we assume the strong open set condition instead of the open set condition then the above Theorem 4.9 tells much more about the maps $\varphi_1, \ldots, \varphi_r$ and $ax + t$ and immediately gives the following higher dimensional generalization of the Logarithmic Commensurability Theorem of Feng and Wang, in which we can also allow non-homogeneous self-similar sets.

COROLLARY 4.10. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set with strong separation condition and suppose that λ is a similitude for which $\lambda(K) \subset K$. If $a_1, \ldots, a_r$ and b denote the similarity ratios of $\varphi_1, \ldots, \varphi_r$ and λ, respectively, then $\log b$ must be a linear combination of $\log a_1, \ldots, \log a_r$ with rational coefficients.*

5. *Isometry invariant measures*

In this section all self-similar sets we consider will satisfy the strong separation condition.

Before we start to study and characterize the isometry invariant measures on a self-similar set of strong separation condition, we have to pay some attention to the connection of a self-similar set and the self-similar measures living on it.

We have called a compact set K self-similar with the SSC if $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ holds for some similitudes $\varphi_1, \ldots, \varphi_r$. A *presentation* of K is a finite collection of similitudes $\{\psi_1, \ldots, \psi_s\}$, such that $K = \psi_1(K) \cup^* \ldots \cup^* \psi_s(K)$ and $s \geq 2$. Clearly, a self-similar set with SSC has many different presentations. For example, if $\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ is a presentation of K, then $\{\varphi_i \circ \varphi_j : 1 \leq i, j \leq r\}$ is also a presentation.

As we will see in the next section, it is possible that a self-similar set has no "smallest" presentation. We say that a presentation $\mathcal{F}_1 = \{\psi_1, \psi_2, \ldots, \psi_s\}$ is smaller than the presentation $\mathcal{F} = \{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ if for every $1 \leq i \leq r$ there exists a multi-index I such that $\varphi_i = \psi_I$. This defines a partial order on the presentations; let us write $\mathcal{F}_1 \leq \mathcal{F}$ when $\mathcal{F}_1$ is smaller than $\mathcal{F}$. We call a presentation *minimal*, if there is no smaller presentation (excluding itself). We call a presentation *smallest*, if it is smaller than any other presentation.

There exists a self-similar set with the SSC which has more than one minimal presentations; that is, it has no smallest presentation (see Section 6).

The notion of a self-similar measure on a self-similar set depends on the presentation. Thus, when we say that μ is a self-similar measure on K, we always mean that μ is self-similar with respect to the given presentation of K. Clearly if $\mathcal{F}_1 \leq \mathcal{F}$, then there are less self-similar measures with respect to $\mathcal{F}_1$ than to $\mathcal{F}$. It will turn out that the isometry invariant self-similar measures are the same independently of the presentations.

Notation 5.1. For the sake of simplicity, for a similitude λ with $\lambda(K) \subset K$ let $\mu(\lambda)$ denote $\mu\big(\lambda(K)\big)$. In the composition of similitudes we might omit the mark $\circ$, so $g_1 g_2$ stands for $g_1 \circ g_2$, and by g^k we will mean the composition of k many g's.

Clearly, given any self-similar measure μ, $\mu \circ \varphi_I = \mu(\varphi_I) \cdot \mu$ holds for the similitudes φ_I arising from the presentation of K. According to the next proposition, if for a given self-similar measure μ the congruent elementary pieces are of equal measure, then the same holds for any similitude λ satisfying $\lambda(K) \subset K$; that is, we have $\mu \circ \lambda = \mu(\lambda) \cdot \mu$ as well.

Proposition 5.2. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set with strong separation condition, and μ be a self-similar measure on K for which the congruent elementary pieces are of equal measure.*

1. *Then for every similitude λ with $\lambda(K) \subset K$, $\mu \circ \lambda = \mu\big(\lambda(K)\big) \cdot \mu$ holds; that is, for any Borel set $H \subset K$ we have $\mu\big(\lambda(H)\big) = \mu\big(\lambda(K)\big) \cdot \mu\big(H\big)$.*
2. *For every elementary piece $\varphi_I(K)$ and for every isometry g for which $g(\varphi_I(K)) \subset K$ holds, we have $\mu\big(\varphi_I(K)\big) = \mu\big(g(\varphi_I(K))\big)$.*

Proof. According to Lemma 4.8 there are only finitely many similitudes λ for which $\lambda(K) \subset K$ holds and $\lambda(K)$ intersects at least two first generation elementary pieces. Denote these by $\lambda_0, \lambda_1, \dots, \lambda_t$, where λ_0 should stand for the identity.

We claim that it is enough to prove the first part of the proposition only for these similitudes. Let λ be a similitude for which $\lambda(K) \subset K$. Let $\varphi_I(K)$ be the smallest elementary piece which contains $\lambda(K)$. Then the similitude $\varphi_I^{-1} \circ \lambda$ maps K into itself and the image intersects at least two first generation elementary pieces, hence it is equal to a similitude λ_i for some i. Thus $\lambda = \varphi_I \circ \lambda_i$. The measure μ being self-similar we have $\mu \circ \varphi_J = p_J \cdot \mu = \mu\big(\varphi_J(K)\big) \cdot \mu$ for every multi-index J, hence for any Borel set $H \subset K$ we obtain

$$\begin{aligned}\mu\big(\lambda(H)\big) = \mu\big((\varphi_I \circ \lambda_i)(H)\big) = \mu\big(\varphi_I(K)\big) \cdot \mu\big(\lambda_i(H)\big) = \mu\big(\varphi_I(K)\big) \cdot \mu\big(\lambda_i(K)\big) \cdot \mu\big(H\big) \\ = \mu\big((\varphi_I \circ \lambda_i)(K)\big) \cdot \mu\big(H\big) = \mu\big(\lambda(K)\big) \cdot \mu\big(H\big),\end{aligned}$$

as we stated.

According to Theorem 4.9, for every integer i with $0 \le i \le t$ there exist multi-indices I_i, J_i and a positive integer k_i, for which $\lambda_i^{k_i} \circ \varphi_{I_i} = \varphi_{J_i}$. Let $b_i \stackrel{\text{def}}{=} \varphi_{I_i}$, $c_i \stackrel{\text{def}}{=} \varphi_{J_i}$, hence $\lambda_i^{k_i} b_i = c_i$.

Let

$$\mu^*(\lambda_i) \stackrel{\text{def}}{=} \sqrt[k_i]{\frac{\mu(c_i)}{\mu(b_i)}}.$$

Our aim is to show that $\mu^*(\lambda_i) = \mu(\lambda_i)$.

For every integer i with $0 \le i \le t$ and for every multi-index I there exists an integer j, $0 \le j \le t$, and a multi-index J such that $\lambda_i \circ \varphi_I = \varphi_J \circ \lambda_j$ (let $\varphi_J(K)$ be the smallest elementary piece which contains $(\lambda_i \circ \varphi_I)(K)$).

We define the *congruency* equivalence relation among similitudes: for similitudes g_1 and g_2 let $g_1 \approx g_2$ denote that $g_1 \circ g_2^{-1}$ is an isometry; that is, for every set H the sets $g_1(H)$ and $g_2(H)$ are congruent. This is the same as that the similarity ratio of g_1 and g_2 are equal. Hence congruency is independent of the order of the composition, so $g_1 \circ g_2 \approx g_3 \iff g_2 \circ g_1 \approx g_3$. Using the equalities $\lambda_i \varphi_I = \varphi_J \lambda_j$, $\lambda_i^{k_i} b_i = c_i$ and $\lambda_j^{k_j} b_j = c_j$ we obtain

$$\begin{gathered}\underbrace{\lambda_i^{k_i k_j}\, \varphi_I^{k_i k_j}}_{\approx \varphi_J^{k_i k_j} \lambda_j^{k_i k_j}}\, b_i^{k_j}\, b_j^{k_i} \;\approx\; \varphi_J^{k_i k_j}\, b_i^{k_j}\, \underbrace{\lambda_j^{k_i k_j}\, b_j^{k_i}}_{\approx c_j^{k_i}} \;\approx\; \varphi_J^{k_i k_j}\, b_i^{k_j}\, c_j^{k_i}, \\ \lambda_i^{k_i k_j}\, \varphi_I^{k_i k_j}\, b_i^{k_j}\, b_j^{k_i} \;\approx\; \underbrace{\lambda_i^{k_i k_j}\, b_i^{k_j}}_{\approx c_i^{k_j}}\, \varphi_I^{k_i k_j}\, b_j^{k_i} \;\approx\; c_i^{k_j}\, \varphi_I^{k_i k_j}\, b_j^{k_i}.\end{gathered}$$

Comparing these we get

$$\varphi_J^{k_i k_j}\, b_i^{k_j}\, c_j^{k_i} \;\approx\; c_i^{k_j}\, \varphi_I^{k_i k_j}\, b_j^{k_i}.$$

Since all the similitudes b_i, b_j, c_i, c_j are some composition of similitudes of the presentation, the elementary pieces $\big(\varphi_J^{k_i k_j}\, b_i^{k_j}\, c_j^{k_i}\big)(K)$ and $\big(c_i^{k_j}\, \varphi_I^{k_i k_j}\, b_j^{k_i}\big)(K)$ are congruent, so they are of equal measure. The measure is self-similar, thus

$$\mu(\varphi_J)^{k_i k_j}\, \mu(b_i)^{k_j}\, \mu(c_j)^{k_i} = \mu(c_i)^{k_j}\, \mu(\varphi_I)^{k_i k_j}\, \mu(b_j)^{k_i},$$

hence by the definition of μ^* we get

$$\mu(\varphi_J)^{k_ik_j}\mu^*(\lambda_j)^{k_ik_j} = \mu^*(\lambda_i)^{k_ik_j}\mu(\varphi_I)^{k_ik_j},$$
$$\mu^*(\lambda_j)\mu(\varphi_J) = \mu^*(\lambda_i)\mu(\varphi_I).$$

Therefore

$$\mu(\lambda_i\varphi_I) = \mu(\varphi_J\lambda_j) = \mu(\varphi_J)\mu(\lambda_j) = \frac{\mu^*(\lambda_i)\mu\big(\varphi_I\big)}{\mu^*\big(\lambda_j\big)}\mu\big(\lambda_j\big).$$

Altering this we get the following: for every i and I there exists j such that

$$\mu(\lambda_i\varphi_I) = \mu^*(\lambda_i)\frac{\mu(\lambda_j)}{\mu^*(\lambda_j)}\mu(\varphi_I).$$

Note that $\mu^*(\lambda_j) \neq 0$.

Let m be an index for which

$$\frac{\mu(\lambda_m)}{\mu^*(\lambda_m)} \mathrel{\dot{\leq}} \frac{\mu(\lambda_i)}{\mu^*(\lambda_i)}$$

for every index $0 \leq i \leq t$. (We label some inequalities with a dot so we can refer to them later.) Then for any φ_I,

$$\mu(\lambda_m\varphi_I) = \mu^*(\lambda_m)\frac{\mu(\lambda_j)}{\mu^*(\lambda_j)}\mu(\varphi_I) \mathrel{\dot{\geq}} \mu^*(\lambda_m)\frac{\mu(\lambda_m)}{\mu^*(\lambda_m)}\mu(\varphi_I) = \mu(\lambda_m)\mu(\varphi_I)$$

for some index j with $0 \leq j \leq t$.

Let $\{\varphi_{I_i}(K)\}$ be a finite partition of K with elementary pieces such that the partition includes $\varphi_I(K)$. Then

$$\mu\big(\lambda_m(K)\big) = \mu\Big(\lambda_m\left(\bigcup\nolimits^* \varphi_{I_i}(K)\right)\Big) = \mu\Big(\bigcup\nolimits^* \lambda_m(\varphi_{I_i}(K))\Big) = \sum \mu(\lambda_m\varphi_{I_i})$$
$$\mathrel{\dot{\geq}} \sum \mu(\lambda_m)\mu(\varphi_{I_i}) = \mu(\lambda_m),$$

hence equality holds everywhere, so $\mu(\lambda_m\varphi_I) = \mu(\lambda_m)\mu(\varphi_I)$ for every multi-index I.

Let $H \subset K$ be a Borel set. By the definition of the measure μ, there exist elementary pieces $a_{ij}(K)$ for which $H \subset \bigcap_j \bigcup_i^* a_{ij}(K)$ and $\mu(H) = \inf_j \mu\big(\bigcup_i^* a_{ij}(K)\big) = \mu\big(\bigcap_j \bigcup_i^* a_{ij}(K)\big)$ hold. Then

$$\mu(\lambda_m(H)) \leq \mu\Big(\lambda_m\Big(\bigcap_j \bigcup_i{}^* a_{ij}(K)\Big)\Big) = \mu\Big(\bigcap_j \bigcup_i{}^* \lambda_m(a_{ij}(K))\Big)$$
$$\leq \inf_j \mu\Big(\bigcup_i{}^* \lambda_m(a_{ij}(K))\Big) = \inf_j \sum_i \mu(\lambda_m a_{ij}) = \inf_j \sum_i \mu(\lambda_m)\mu(a_{ij})$$
$$= \mu(\lambda_m)\inf_j \sum_i \mu(a_{ij}) = \mu(\lambda_m)\mu\Big(\bigcap_j \bigcup_i{}^* a_{ij}(K)\Big) = \mu(\lambda_m)\mu(H).$$

Repeating this argument for $H^c \overset{\text{def}}{=} K \setminus H$ we obtain $\mu\big(\lambda_m(H^c)\big) \leq \mu(\lambda_m)\mu(H^c)$. Summing these we get $\mu\big(\lambda_m(H)\big)+\mu\big(\lambda_m(H^c)\big) \leq \mu(\lambda_m)\mu(H)+\mu(\lambda_m)\mu(H^c)$, in fact this is an equality, so we have $\mu\big(\lambda_m(H)\big) = \mu(\lambda_m)\mu(H)$. Thus $\mu \circ \lambda_m = \mu(\lambda_m)\cdot\mu$.

From this we obtain that for any Borel set $H \subset K$,

$$\mu\big(\lambda_m^n(H)\big) = \mu\big(\lambda_m(\lambda_m^{n-1}(H))\big) = \mu(\lambda_m)\mu\big(\lambda_m^{n-1}(H)\big),$$

and by induction we get that $\mu\big(\lambda_m^n(H)\big) = \mu(\lambda_m)^n\mu(H)$. Hence $\mu(\lambda_m^n) = \mu(\lambda_m)^n$.

Therefore $\mu(\lambda_m^{k_m} b_m) = \mu(\lambda_m)^{k_m}\mu(b_m)$ holds. From the definition of $\mu^*(\lambda_m)$ we have $\mu(c_m) = \mu^*(\lambda_m)^{k_m}\mu(b_m)$ and $c_m = \lambda_m^{k_m} b_m$, thus

$$\mu(\lambda_m)^{k_m}\mu(b_m) = \mu(\lambda_m^{k_m} b_m) = \mu(c_m) = \mu^*(\lambda_m)^{k_m}\mu(b_m).$$

Since $\mu(b_m) > 0$, we get $\mu(\lambda_m) = \mu^*(\lambda_m)$. Since m was chosen to be that index i for which $\frac{\mu(\lambda_i)}{\mu^*(\lambda_i)}$ is minimal, we get that $\mu^*(\lambda_i) \mathrel{\dot{\leq}} \mu(\lambda_i)$ for every $0 \leq i \leq t$.

Now we can repeat the whole argument for such an index m for which $\frac{\mu(\lambda_m)}{\mu^*(\lambda_m)} \geq \frac{\mu(\lambda_i)}{\mu^*(\lambda_i)}$ holds for every index i $(0 \leq i \leq t)$. We just have to reverse the inequalities labelled with a dot, and we obtain that for every index i $(0 \leq i \leq t)$, $\mu^*(\lambda_i) \geq \mu(\lambda_i)$ holds. Thus for every i $(0 \leq i \leq t)$ we have $\mu^*(\lambda_i) = \mu(\lambda_i)$. Therefore we could choose any i $(0 \leq i \leq t)$ as m, so for every i the equality $\mu \circ \lambda_i = \mu(\lambda_i) \cdot \mu$ holds. By the observation we made at the beginning of the proof we get that for every similitude λ with $\lambda(K) \subset K$, $\mu \circ \lambda = \mu(\lambda) \cdot \mu$ holds, thus $\mu \circ \lambda^n = \mu(\lambda)^n \cdot \mu$ holds as well for any positive integer n.

Now we shall prove the second part of the proposition. Suppose that the isometry g maps the elementary piece $\varphi_L(K)$ into K, so $g(\varphi_L(K)) \subset K$. By Theorem 4.9 there exist multi-indices I, J and a positive integer k such that $(g \circ \varphi_L)^k \circ \varphi_I = \varphi_J$. Using the first part of this proposition (which is already proven) we get

$$\mu(\varphi_J) = \mu\big((g \circ \varphi_L)^k \circ \varphi_I\big) = \mu(g \circ \varphi_L)^k \mu(\varphi_I). \tag{14}$$

Clearly $\varphi_J = (g \circ \varphi_L)^k \circ \varphi_I \approx (\varphi_L)^k \varphi_I$, thus

$$\mu(\varphi_J) = \mu\big((\varphi_L)^k \varphi_I\big) = \mu(\varphi_L)^k \mu(\varphi_I). \tag{15}$$

By (14) and (15) we obtain

$$\mu(g \circ \varphi_L)^k \mu(\varphi_I) = \mu(\varphi_L)^k \mu(\varphi_I),$$

$$\mu(g \circ \varphi_L) = \mu(\varphi_L),$$

which proves the proposition. □

Theorem 5.3 (Characterization of isometry invariant measures)
Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set with the strong separation condition and μ a self-similar measure on K for which congruent elementary pieces are of equal measure. Then μ is an isometry invariant measure on K.

Proof. We have to show that for any isometry g and Borel set $H \subset K$ if $g(H) \subset K$ then $\mu(H) = \mu\big(g(H)\big)$.

Let $c < 1$ be the constant given by Theorem 4.1. At first consider a set $H \subset K$ of positive measure. Applying Lemma 2.12 for the set H with $\varepsilon = 1 - c$ we obtain that there exists an elementary piece $a(K)$ for which $\mu\big(H \cap a(K)\big) > c \cdot \mu\big(a(K)\big)$. Since

$H \subset g^{-1}(K)$, we have $\mu\big(g^{-1}(K) \cap a(K)\big) > c \cdot \mu\big(a(K)\big)$, so applying Theorem 4.1 $a(K) \subset g^{-1}(K)$, $g(a(K)) \subset K$. Put $\lambda = g \circ a$. According to the second part of Proposition 5.2 we have $\mu(\lambda) = \mu(a)$ (where $\mu(\lambda)$ is an abbreviation of $\mu\big(\lambda(K)\big)$), and putting $H_0 \stackrel{\text{def}}{=} a^{-1}(a(K) \cap H)$ we have $\mu\big(\lambda(H_0)\big) = \mu(\lambda)\mu(H_0)$, thus

$$\begin{aligned} 0 < c \cdot \mu\big(a(K)\big) < \mu\big(a(K) \cap H\big) = \mu\big(a(H_0)\big) = \mu(a)\mu(H_0) = \mu(\lambda)\mu(H_0) \\ = \mu\big(\lambda(H_0)\big) = \mu\big(g(a(H_0))\big) = \mu\big(g(a(K) \cap H)\big) \le \mu\big(g(H)\big), \end{aligned}$$

so $g(H)$ is of positive measure. Thus a congruent copy of a set of positive measure is of positive measure, and a congruent copy of a negligible set is also negligible.

Now let $H \subset K$ be any Borel set, g an isometry, for which $g(H) \subset K$. Apply Lemma 2.13 with some $0 < \varepsilon < 1 - c$. We obtain elementary pieces $a_i(K)$ such that

$$\mu\big(H \cap a_i(K)\big) > (1 - \varepsilon) \cdot \mu\big(a_i(K)\big) \quad \text{and} \quad \mu\Big(H \setminus \bigcup_i^* a_i(K)\Big) = 0.$$

Then $H \subset g^{-1}(K)$, therefore $\mu\big(g^{-1}(K) \cap a_i(K)\big) > (1 - \varepsilon) \cdot \mu\big(a_i(K)\big)$. According to Theorem 4.1, $g^{-1}(K) \supset a_i(K)$, so $g(a_i(K)) \subset K$. By the second part of Proposition 5.2 we get $\mu\big(g(a_i(K))\big) = \mu\big(a_i(K)\big)$, and using the fact that a congruent copy of a set of zero measure is also of zero measure,

$$\begin{aligned} \mu\big(g(H)\big) &= \mu\Big(g\Big(H \cap \bigcup_i^* a_i(K)\Big)\Big) + \mu\Big(g\Big(H \setminus \bigcup_i^* a_i(K)\Big)\Big) = \mu\Big(g\Big(H \cap \bigcup_i^* a_i(K)\Big)\Big) \\ &= \sum_i \mu\big(g(H \cap a_i(K))\big) \le \sum_i \mu\big(g(a_i(K))\big) = \sum_i \mu\big(a_i(K)\big) \\ &\le \frac{1}{1 - \varepsilon} \cdot \sum_i \mu\big(H \cap a_i(K)\big) = \frac{1}{1 - \varepsilon} \cdot \mu\Big(H \cap \bigcup_i^* a_i(K)\Big) = \frac{1}{1 - \varepsilon} \cdot \mu(H). \end{aligned}$$

This is true for any $0 < \varepsilon < 1 - c$, hence $\mu\big(g(H)\big) \le \mu\big(H\big)$. Repeating this argument for $g(H)$ instead of H and for g^{-1} instead of g gives $\mu\big(H\big) \le \mu\big(g(H)\big)$, hence $\mu\big(H\big) = \mu\big(g(H)\big)$. Thus μ is isometry invariant. □

REMARK 5.4. Using this theorem it is relatively easy to decide whether a self-similar measure is isometry invariant or not. Denote the similarity ratio of the similitude φ_i by α_i. It is clear that two elementary pieces are congruent if and only if they are images of K by similitudes of equal similarity ratio. Thus a self-similar measure μ is isometry invariant if and only if the equality $p_{i_1} p_{i_2} \dots p_{i_n} = p_{j_1} p_{j_2} \dots p_{j_m}$ holds (for the weights of the measure μ) whenever $\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_n} = \alpha_{j_1} \alpha_{j_2} \dots \alpha_{j_m}$. By switching from the similarity ratios α_i and weights p_i to the negative of their logarithm we get a system of linear equations for the variables $-\log p_i$. The solutions of this system and the additional equation $\sum_i p_i = 1$ give those weight vectors which define isometry invariant measures on K.

For example, it is easy to see that if the positive numbers $-\log \alpha_i$ $(i = 1, \dots, r)$ are linearly independent over $\mathbb{Q}$, then every self-similar measure is isometry invariant.

So, to the r dimensional vectors, formed by the $-\log p_i$ weights of the isometry invariant measures, correspond the intersection of a linear subspace of $\mathbb{R}^r$ and the hypersurface corresponding to $\sum_i p_i = 1$. That this subspace is of dimension at least 1 and intersects the positive part of the space $\mathbb{R}^r$ we know from the existence of Hausdorff measure. (Or rather from the fact that the weights $p_i = \alpha_i^s$ automatically satisfy all the equalities.)

The notion of a self-similar measure depended on the choice of the presentation. However, the converse is true for the notion of isometry invariant self-similar measure.

THEOREM 5.5. *Let K be self-similar with the strong separation condition and $\{\varphi_1, \varphi_2, \dots, \varphi_r\}$ a presentation of it. Let μ be isometry invariant and self-similar with respect to this presentation. Then μ is self-similar with respect to any presentation of K. Thus the class of isometry invariant self-similar measures is independent of the choice of the presentation.*

Proof. Let $\{\psi_1, \dots, \psi_s\}$ be an other presentation of K. According to Theorem 4.9 there exist positive integer k and elementary pieces φ_I, φ_J such that $\psi_i^k \circ \varphi_I = \varphi_J$, so applying the first part of Proposition 5.2 we get

$$0 < \mu(\varphi_J) = \mu(\psi_i^k \circ \varphi_I) = \mu(\psi_i)^k \mu(\varphi_I),$$

that is, $\mu(\psi_i) > 0$ for every $1 \le i \le s$.

According to the first part of Proposition 5.2, $\mu \circ \psi_i = \mu(\psi_i) \cdot \mu$, and since $\sum \mu\big(\psi_i(K)\big) = 1$ holds, this means exactly that μ is a self-similar measure with respect to the presentation $\{\psi_1, \dots, \psi_s\}$. $\square$

DEFINITION 5.6. Let $K = \varphi_1(K) \cup^* \dots \cup^* \varphi_r(K)$ be a self-similar set with the strong separation condition. Put $S = \{-\log \alpha_i : 1 \le i \le r\}$, where α_i is the similarity ratio of φ_i. The *algebraic dependence number* (of this presentation) is the dimension over $\mathbb{Q}$ of the vectorspace generated by S minus one.

By Remark 5.4 it is easy to see that the algebraic dependence number of a presentation is exactly the same as the topological dimension of the surface corresponding to the isometry invariant self-similar measures on K. Thus, by Theorem 5.5, one can prove the following.

THEOREM 5.7. *The algebraic dependence number of a self-similar set does not depend on the presentation we choose.*

We mention that it is easy to show that the algebraic dependence number is the same for two presentations $\mathcal{F}_1$ and $\mathcal{F}_2$ if $\mathcal{F}_1 \le \mathcal{F}_2$; that is, when one of them extends the other in the trivial way we defined at the beginning of this section. However, there are self-similar sets with two presentations which have no common extension and they are not an extension of the same third presentation (see Theorem 6.4). Thus we have no direct (or trivial) proof for Theorem 5.7.

An easy consequence of the characterization theorem is the following.

Corollary 5.8. *Let $K = \varphi_1(K) \cup^* \ldots \cup^* \varphi_r(K)$ be a self-similar set with strong separation condition, μ be a self-similar measure on K. Then if μ is invariant under orientation preserving isometries, then it is invariant under all isometries.*

Proof. According to Theorem 5.3 it is enough to show that congruent elementary pieces are of equal measure. Let $\varphi_I(K)$ and $\varphi_J(K)$ be congruent elementary pieces. Then $\varphi_I^2(K)$ and $\varphi_J^2(K)$ are also congruent elementary pieces, and φ_I^2 and φ_J^2 are orientation preserving similitudes. This implies that $\varphi_I^2 \circ \varphi_J^{-2}$ is an orientation preserving isometry. Therefore, by the assumption $\mu\big(\varphi_I^2(K)\big) = \mu\big(\varphi_J^2(K)\big)$. Since μ is self-similar, $\mu\big(\varphi_I^2(K)\big) = \mu\big(\varphi_I(K)\big)^2$ and $\mu\big(\varphi_J^2(K)\big) = \mu\big(\varphi_J(K)\big)^2$, thus $\mu\big(\varphi_I(K)\big) = \mu\big(\varphi_J(K)\big)$. This proves the statement. □

6. *Minimal presentations*

First we give an example of a self-similar set on the line (with the strong separation condition) that has no smallest presentation, that is, it has more than one minimal presentations. Set $\varphi_1(x) = \frac{x}{3}$, $\varphi_2(x) = \frac{x}{3} + \frac{2}{3}$, let K be the compact set for which $K = \varphi_1(K) \cup \varphi_2(K)$, apparently this is the triadic Cantor set. Set $\psi_1(x) = -\frac{x}{3} + \frac{1}{3}$. Then $K = \psi_1(K) \cup^* \varphi_2(K)$ as well, and it is clear, that both of these two different presentations are minimal, since they consist of only two similitudes.

However, these two presentations are not "essentially different": the sets $\{\varphi_1(K), \varphi_2(K)\}$ and $\{\psi_1(K), \varphi_2(K)\}$ coincide. On *essential presentation* we shall mean not the set of the similitudes but rather the set of the first generation elementary pieces. We shall say that the essential presentation $\{a_1(K), \ldots, a_r(K)\}$ is *briefer* than the essential presentation $\{b_1(K), \ldots, b_s(K)\}$, if for every $j = 1, \ldots, s$ there exists $1 \le i \le r$ such that $b_j(K) \subset a_i(K)$. We call an essential presentation minimal if the only briefer essential presentation is itself, and we call it the smallest if it is briefer than any other essential presentation. It is easy to check that the triadic Cantor set possesses a smallest essential presentation.

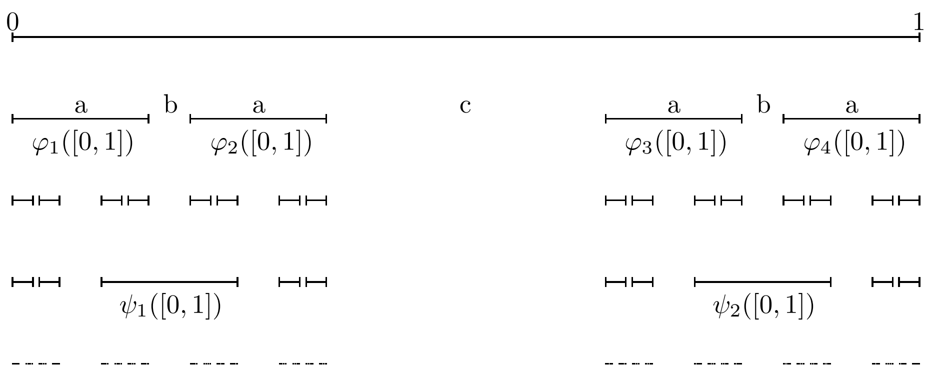
In what follows we shall present a self-similar set which has no smallest essential presentation, that is, it has more than one minimal essential presentations.

Remark 6.1. The following statement is true for many self-similar sets K: If λ_1 and λ_2 are similitudes for which $\lambda_1(K) \subset K$, $\lambda_2(K) \subset K$ and $\lambda_1(K) \cap \lambda_2(K) \ne \emptyset$, then $\lambda_1(K) \subset \lambda_2(K)$ or $\lambda_2(K) \subset \lambda_1(K)$. The proofs of Section 4 would have been much simpler if this statement held for every self-similar set satisfying the strong separation condition. However this statement does not hold generally as we shall show in our following construction. We note that this statement is not necessarily equivalent to that K has only one minimal essential presentation. See also the end of Section 9 and especially Question 9.3.

Theorem 6.2. *There exists a self-similar set K with the strong separation condition which has no smallest essential presentation. Moreover, there exist similitudes λ_1 and λ_2 such that $\lambda_1(K) \cap \lambda_2(K) \ne \emptyset$, but $\lambda_1(K) \not\subset \lambda_2(K)$ and $\lambda_2(K) \not\subset \lambda_1(K)$.*

Proof. We present a figure of our construction. One may check the proof of this theorem just by looking at that figure.

Let a, b, c be positive integers for which $a+b+a+c+a+b+a=1$ and $b=a\cdot c$. It is easy to see that for every $0<a<1/4$ there exist a unique b and c with these conditions. Let φ_1 be the orientation preserving similitude mapping the interval $[0,1]$ onto the interval $[0,a]$. Let φ_2 take the interval $[0,1]$ onto $[a+b, a+b+a]$, φ_3 onto $[1-a-b-a, 1-a-b]$, and φ_4 onto $[1-a, 1]$, all of them preserving the orientation. That is, $\varphi_1(x)=a\cdot x$, $\varphi_2(x)=a\cdot x+a+b$, $\varphi_3(x)=a\cdot x+1-a-b-a$, $\varphi_4(x)=a\cdot x+1-a$.



Let K be the unique compact set for which $K=\varphi_1(K)\cup^*\varphi_2(K)\cup^*\varphi_3(K)\cup^*\varphi_4(K)$. Thus the first generation elementary pieces of K are all of diameter a, and there are gaps between them of length b, c and b. It is clear that $K\subset[0,1]$ and K is symmetric about $\frac{1}{2}$.

The second row of the figure symbolizes this presentation of K, more precisely it shows the intervals $\varphi_i([0,1])$ (choosing $a=0.15,\ c=\frac{0.4}{1.3}$). The third row of the figure shows the intervals $\varphi_i(\varphi_j([0,1]))\quad(1\le i,j\le 4)$. The fifth row is an attempt to visualize K itself.

Set $\psi_1(x)=a\cdot x+a^2+a\cdot b+a^2+a\cdot c$ and $\psi_2(x)=a\cdot x+1-a-b-a^2-a\cdot b-a^2$. In the fourth row of the figure the images of the interval $[0,1]$ by the similitudes φ_1^2, $\varphi_1\circ\varphi_2$, ψ_1, $\varphi_2\circ\varphi_3$, $\varphi_2\circ\varphi_4$, $\varphi_3\circ\varphi_1$, $\varphi_3\circ\varphi_2$, ψ_2, $\varphi_4\circ\varphi_3$, φ_4^2 are shown.

We claim that $\psi_1(K)\subset K$ and $\psi_2(K)\subset K$, moreover

$$\{\varphi_1^2,\ \varphi_1\circ\varphi_2,\ \psi_1,\ \varphi_2\circ\varphi_3,\ \varphi_2\circ\varphi_4,\ \varphi_3\circ\varphi_1,\ \varphi_3\circ\varphi_2,\ \psi_2,\ \varphi_4\circ\varphi_3,\ \varphi_4^2\}$$

is a presentation of K (see the fourth row of the figure). For this it is sufficient to prove that $\psi_1\circ\varphi_1=\varphi_1\circ\varphi_3$, $\psi_1\circ\varphi_2=\varphi_1\circ\varphi_4$, $\psi_1\circ\varphi_3=\varphi_2\circ\varphi_1$, $\psi_1\circ\varphi_4=\varphi_2\circ\varphi_2$, and $\psi_2\circ\varphi_1=\varphi_3\circ\varphi_3$, $\psi_2\circ\varphi_2=\varphi_3\circ\varphi_4$, $\psi_2\circ\varphi_3=\varphi_4\circ\varphi_1$, $\psi_2\circ\varphi_4=\varphi_4\circ\varphi_2$. These can be easily checked, all equalities rely on the choice of $b=a\cdot c$.

Now we prove that there does not exist an essential presentation $\{\varrho_1(K),\dots,\varrho_r(K)\}$ of the self-similar set K which is briefer than both of the essential presentations corresponding to the original presentation $\{\varphi_1,\varphi_2,\varphi_3,\varphi_4\}$ and the presentation just defined above. This will prove that K has no unique minimal essential presentation. (In fact both of these essential presentations are minimal.) Suppose on the contrary that there exists an essential presentation $\{\varrho_1(K), \ldots, \varrho_r(K)\}$ of this kind. Since $\varphi_1(K) \cap \psi_1(K) \neq \emptyset$, for some i $\varphi_1(K) \cup \psi_1(K) \subset \varrho_i(K)$. For the same i we also have $\varphi_2(K) \cup \psi_1(K) \subset \varrho_i(K)$. Similarly there exists an index j such that $\varphi_3(K) \cup \psi_2(K) \cup \varphi_4(K) \subset \varrho_j(K)$. From this we conclude that $K = \varrho_i(K) \cup^* \varrho_j(K)$, but then the similitudes ϱ_i and ϱ_j could only be the ones mapping $[0,1]$ onto $[0, a+b+a]$ and $[1-a-b-a, 1]$. This yields $b = (a+b+a) \cdot c$, which contradicts $b = a \cdot c$.

The similitudes λ_1 and λ_2 promised above can be chosen to be φ_1 and ψ_1. □

Remark 6.3. This example (and many other results of the present article) is contained in the Master Thesis of the third author [**18**]. Independently, Feng and Wang in [**8**] exhibit an almost identical example. Moreover, much of their paper is devoted to the investigation of the structure of possible presentations of given self-similar sets; or, using their terminology, the structure of generating iterated function systems of self-similar sets. They also prove positive results (that is, when a smallest presentation does exist) under various assumptions.

Theorem 6.4. *There exists a self-similar set K with the strong separation condition and two (essential) presentations of K, $\mathcal{F}_1$ and $\mathcal{F}_2$, such that there is no presentation $\mathcal{G}$ which is a common extension of $\mathcal{F}_1$ and $\mathcal{F}_2$, nor there exists an (essential) presentation which is smaller (briefer) than $\mathcal{F}_1$ and $\mathcal{F}_2$.*

Thus, Theorem 5.7 cannot be proved in the trivial way (see our remarks after that theorem). We leave the proof of Theorem 6.4 to the reader, with the instructions that one should choose the self-similar set K constructed above, and the presentations of the second and fourth row of the figure should be chosen as $\mathcal{F}_1$ and $\mathcal{F}_2$.

7. *Intersection of translates of a self-affine Sierpiński sponge*

The following is the key lemma for all results of this section.

Proposition 7.1. *Let $K = K(M, D)$ and μ be like in Definition 2.14 and let $t \in \mathbb{R}^n$ be such that $\|M^k t\| > 0$ for every $k = 1, 2, \ldots$.*

Then $\mu\big(K \cap (K+t)\big) > 0$ implies that there exists a

$$w \in \{-1,0,1\} \times \ldots \times \{-1,0,1\} \setminus \{(0,\ldots,0)\}$$

such that $D + w = D$ modulo $(m_1, \ldots, m_n)$; that is,

$$D + w + M(\mathbb{Z}^n) = D + M(\mathbb{Z}^n) = D - w + M(\mathbb{Z}^n).$$

Proof. To make the argument intuitive and precise we shall present the same proof in an informal and in a formal way separately.

The informal proof: According to Lemma 2.16 and Lemma 2.12 we can find a k such that $M^k t$ is not very close to any point of $\mathbb{Z}^n$, and a $k-1$-th generation elementary piece S of K in which the density of $K+t$ is almost 1. Then in all the

r k-th generation elementary pieces of K that are in S the density of $K+t$ is still very close to 1.

Each of these subparts intersect some k-th generation elementary pieces of $K+t$. The key observation is that there are at most 2^n possible ways how these parts can intersect each other.

Since $M^k t$ is not very close to the lattice points, these intersections are intersections of sets similar to K such that one is always a not very close translate of the other. Hence Lemma 2.23 implies that they cannot have big intersection.

Since the density of $K+t$ is very close to 1 in all k-th generation elementary pieces of K that are in S, this implies that in the two directions for which the possible intersection has biggest measure, $K+t$ must have a k-th generation elementary piece.

Hence we get two periods of the pattern D such that their difference w is in $\{-1,0,1\} \times \ldots \times \{-1,0,1\}$.

The formal proof: Applying Lemma 2.23 for $\varepsilon = 1/(2\max(m_1,\ldots,m_n))$ we get a $0<\delta<1$ such that

$$\mu\big(K \cap (K+u)\big) \le 1-\delta \quad \text{whenever } |u| \ge \frac{1}{2\max(m_1,\ldots,m_n)}. \tag{16}$$

Applying Lemma 2.12 for $B=(K+t)\cap K$ and $\varepsilon = \frac{\delta}{2^n r}$ and Lemma 2.16 we get a $k \in \mathbb{N}$ and a $k-1$-th generation elementary piece S of K such that

$$\mu\big(S \cap (K+t)\big) > \frac{1-\frac{\delta}{2^n r}}{r^{k-1}} \tag{17}$$

and

$$\|M^k t\| > \frac{1}{2\max(m_1,\ldots,m_n)}. \tag{18}$$

Let Φ be the similarity map which maps S to $M(K)=K+D$; that is,

$$\begin{aligned}\Phi(x) &= M^k(x-(M^{-(k-1)}\alpha_{k-1}+\ldots+M^{-1}\alpha_1)) \\ &= M^k x - (M\alpha_{k-1}+M^2\alpha_{k-2}+\ldots+M^{k-1}\alpha_1),\end{aligned}$$

where $S = M^{k-1}(K) + M^{-(k-1)}\alpha_{k-1}+\ldots+M^{-1}\alpha_1$.

Using that $\Phi(S) = K+D = \cup_{j=1}^r K+d_j$, applying (3) and (17) we get

$$\begin{aligned}\tilde\mu\Big(\bigcup_{j=1}^r (K+d_j)\cap(\Phi(K+t))\Big) &= \tilde\mu\big(\Phi(S\cap(K+t))\big) = r^k\mu\big(S\cap(K+t)\big) \\ &> r^k\frac{1-\frac{\delta}{2^n r}}{r^{k-1}} = r - \frac{\delta}{2^n}.\end{aligned}$$

Since $\tilde\mu(K+d_j)=1$ $(j=1,\ldots,r)$ and the sets can intersect each other only at a set of $\tilde\mu$-measure zero this implies that

$$\tilde\mu\big((K+d_j)\cap\Phi(K+t)\big) > 1-\frac{\delta}{2^n} \quad \text{for every } j=1,\ldots,r. \tag{19}$$

Since $\Phi(K) = M^k(K) - (M\alpha_{k-1} + \ldots + M^{k-1}\alpha_1)$ and $M^k(K) \subset K + D + M(\mathbb{Z}^n)$, we have $\Phi(K) \subset K + D + M(\mathbb{Z}^n)$, and so $\Phi(K+t) \subset K + D + \Phi(t) + M(\mathbb{Z}^n)$. Thus

$$\begin{aligned}(K + d_j) \cap \Phi(K+t) \\ &\subset (K + d_j) \cap \big(K + D + \Phi(t) + M(\mathbb{Z}^n)\big) \\ &= \bigcup_{i=1}^{r} \big(K \cap (K + d_i + \Phi(t) - d_j + M(\mathbb{Z}^n))\big) + d_j.\end{aligned}$$

Combining this with (19) and (3) (for $l = 0$) we get

$$\begin{aligned}1 - \frac{\delta}{2^n} &< \tilde{\mu}\big((K + d_j) \cap \Phi(K+t)\big) \\ &\leq \sum_{i=1}^{r} \tilde{\mu}\big(\big(K \cap (K + d_i + \Phi(t) - d_j + M(\mathbb{Z}^n))\big) + d_j\big) \qquad (20) \\ &= \sum_{i=1}^{r} \mu\big(K \cap (K + d_i + \Phi(t) - d_j + M(\mathbb{Z}^n))\big) \qquad (j = 1, \ldots, r).\end{aligned}$$

Clearly, we have $\mu\big(K \cap (K + d_i + \Phi(t) - d_j + M(\mathbb{Z}^n))\big) = 0$ whenever

$$d_i + \Phi(t) - d_j \notin (-1,1) \times \ldots \times (-1,1) + M(\mathbb{Z}^n).$$

Hence there are at most 2^n vectors $v \in \mathbb{Z}^n$ such that $v + \Phi(t) \in (-1,1) \times \ldots \times (-1,1)$; let these vectors be $v_1, v_2, \ldots, v_p$, $(p \leq 2^n)$.

Thus, by omitting some zero terms on the right-hand side of (20) we can rewrite (20) as

$$1 - \frac{\delta}{2^n} < \sum_{l \,:\, (\exists i)\ d_i - d_j \in v_l + M(\mathbb{Z}^n)} \mu\big(K \cap (K + v_l + \Phi(t))\big) \qquad (j = 1, \ldots, r). \tag{21}$$

Let

$$\beta_l = \mu\big(K \cap (K + v_l + \Phi(t))\big) \qquad (l = 1, \ldots, p).$$

By rearranging $v_1, \ldots, v_p$ if necessary, we may assume that

$$\beta_1 \geq \beta_2 \geq \ldots \geq \beta_p. \tag{22}$$

Since $v_l \in \mathbb{Z}^n$ and $K \subset [0,1]^n$, the sets $K + v_l + \Phi(t)$ $(l = 1, \ldots, p)$ are pairwise disjoint and clearly $K = \cup_{l=1}^{p} K \cap (K + v_l + \Phi(t))$, we get

$$1 = \mu(K) = \sum_{l=1}^{p} \beta_l. \tag{23}$$

Since, using (18), $\|M^k t\| > \frac{1}{2\max(m_1, \ldots, m_n)}$, we have $|v_1 + \Phi(t)| > \frac{1}{2\max(m_1, \ldots, m_n)}$. Thus, by (16),

$$\beta_1 = \mu\big(K \cap (K + v_1 + \Phi(t))\big) \leq 1 - \delta. \tag{24}$$

Clearly (22), (23) and (24) implies that $\beta_1 \geq \beta_2 \geq \frac{\delta}{p-1} > \frac{\delta}{2^n}$ and so

$$\begin{aligned}\beta_1 + \beta_3 + \beta_4 + \ldots \beta_{2^n} &< 1 - \frac{\delta}{2^n} \qquad \text{and} \\ \beta_2 + \beta_3 + \beta_4 + \ldots \beta_{2^n} &< 1 - \frac{\delta}{2^n}.\end{aligned}$$

Combining this with (21) we get that for every $j \in \{1, \ldots, r\}$ there must be an i_1 such that $d_{i_1} - d_j \in v_1 + M(\mathbb{Z}^n)$ and an i_2 such that $d_{i_2} - d_j \in v_2 + M(\mathbb{Z}^n)$. Since $D = \{d_1, \ldots, d_r\}$, this means that for every $d \in D$ we must have $d + v_1, d + v_2 \in D + M(\mathbb{Z}^n)$.

Therefore $D + M(\mathbb{Z}^n) \supset D + v_1$ and so $D + M(\mathbb{Z}^n) \supset D + M(\mathbb{Z}^n) + v_1$. Applying this $m_1 \cdot \ldots \cdot m_n$ many times we get

$$\begin{aligned} D + M(\mathbb{Z}^n) \supset D + M(\mathbb{Z}^n) + v_1 \supset D + M(\mathbb{Z}^n) + 2v_1 \supset \ldots \\ \ldots \supset D + M(\mathbb{Z}^n) + m_1 \cdot \ldots \cdot m_n v_1 = D + M(\mathbb{Z}^n). \end{aligned} \tag{25}$$

Therefore $D + M(\mathbb{Z}^n) = D + M(\mathbb{Z}^n) + v_1$ and similarly $D + M(\mathbb{Z}^n) = D + M(\mathbb{Z}^n) + v_2$. Thus $D + M(\mathbb{Z}^n) + v_1 - v_2 = D + M(\mathbb{Z}^n) = D + M(\mathbb{Z}^n) + v_2 - v_1$. Noting that, by definition, $w = v_1 - v_2 \in \{-1, 0, 1\} \times \ldots \times \{-1, 0, 1\} \setminus \{(0, \ldots, 0)\}$, the proof is complete. □

In order to use Proposition 7.1 effectively we need a discrete lemma.

Lemma 7.2. *Let M and D be like in Definition 2.14, $l \in \{1, 2, \ldots, n\}$, $i \in \mathbb{N}$,*

$$D_i = M^{i-1}(D) + M^{i-2}(D) + \ldots + M(D) + D,$$

and suppose that

$$D_i + (\underbrace{1, \ldots, 1}_{l}, 0, \ldots, 0) + M^i(\mathbb{Z}^n) = D_i + M^i(\mathbb{Z}^n). \tag{26}$$

Then at least one of the following two statements hold.
(a) We have $m_1 = \ldots = m_l$ and $a_1 = \ldots = a_l$ for every $(a_1, \ldots, a_n) \in D$.
(b) For some $l' \in \{1, 2, \ldots, l-1\}$ we have

$$D_{i-1} + (\underbrace{1, \ldots, 1}_{l'}, 0, \ldots, 0) + M^{i-1}(\mathbb{Z}^n) = D_{i-1} + M^{i-1}(\mathbb{Z}^n).$$

Proof. Let $w = (\underbrace{1, \ldots, 1}_{l}, 0, \ldots, 0)$. From (26) we get

$$D_i + kw + M^i(\mathbb{Z}^n) = D_i + M^i(\mathbb{Z}^n) \qquad (k \in \mathbb{Z}). \tag{27}$$

First suppose that $a_1 = \ldots = a_l$ does not hold for some $a = (a_1, \ldots, a_n) \in D$. Then we can suppose that $a_1 = \ldots = a_j < a_{j+1} \leq \ldots \leq a_l$ for some $j \in \{1, \ldots, l-1\}$. Let $b = (b_1, \ldots, b_n) \in D_{i-1}$ be arbitrary. Then $Mb + a \in M(D_{i-1}) + D = D_i$. Thus applying (27) for $k = -(a_1 + 1)$ we get

$$Mb + a - (a_1 + 1)w \in D_i + M^i(\mathbb{Z}^n).$$

Rewriting both sides we get

$$\begin{aligned} &M((b_1 - 1, \ldots, b_j - 1, b_{j+1}, \ldots, b_n)) \\ &\quad + (m_1 - 1, \ldots, m_j - 1, a_{j+1} - a_1 - 1, \ldots, a_l - a_1 - 1, a_{l+1}, \ldots, a_n) \\ &\qquad\qquad \in M(D_{i-1} + M^{i-1}(\mathbb{Z}^n)) + D. \end{aligned}$$

Since the second term of the left-hand side is in $\{0, 1 \ldots, m_1-1\} \times \{0, 1, \ldots, m_n-1\}$, we must have

$$(b_1 - 1, \ldots, b_j - 1, b_{j+1}, \ldots, b_n) \in D_{i-1} + M^{i-1}(\mathbb{Z}^n).$$

Since $b = (b_1, \ldots, b_n) \in D_{i-1}$ was arbitrary we get that

$$D_{i-1} - (\underbrace{1, \ldots, 1}_{j}, 0, \ldots, 0) \subset D_{i-1} + M^{i-1}(\mathbb{Z}^n),$$

which implies, similarly like in (7), that

$$D_{i-1} + (\underbrace{1, \ldots, 1}_{j}, 0, \ldots, 0) + M^{i-1}(\mathbb{Z}^n) = D_{i-1} + M^{i-1}(\mathbb{Z}^n).$$

Thus we proved that if $a_1 = \ldots = a_l$ does not hold for some $(a_1, \ldots, a_l) \in D$ then the statement (b) must hold. Exactly the same way (but ordering so that $m_1 - a_1 \leq \ldots \leq m_n - a_n$ and applying (27) for $k = m_1 - a_1$ instead of $k = a_1$) we get that if $m_1 - a_1 = \ldots = m_l - a_l$ does not hold for some $(a_1, \ldots, a_n) \in D$ then again the statement (b) must hold. Therefore the negation of (a) implies (b), which completes the proof of the Lemma. □

LEMMA 7.3. *Let $K = K(M, D)$ be a self-affine Sierpiński sponge in $\mathbb{R}^n$ as described in Definition 2.14. Let $D_n = M^{n-1}(D) + M^{n-2}(D) + \ldots + M(D) + D$ and suppose that there exists a $w_n \in \{-1, 0, 1\} \times \ldots \times \{-1, 0, 1\} \setminus \{(0, \ldots, 0)\}$ such that*

$$D_n + w_n + M^n(\mathbb{Z}^n) = D_n + M^n(\mathbb{Z}^n).$$

Then K is of the form $K = L \times K_0$, where L is a diagonal of a cube $[0, 1]^l$, where $l \in \{1, 2, \ldots, n\}$ and K_0 is a smaller dimensional self-affine Sierpiński sponge.

Proof. Since every condition is invariant under any autoisometry of the cube $[0, 1]^n$ and by such a transformation we can map w_n to a vector of the form $(1, \ldots, 1, 0, \ldots, 0)$ we can suppose that

$$w_n = (\underbrace{1, \ldots, 1}_{l_n}, 0, \ldots, 0), \qquad \text{where } l_n \in \{1, 2, \ldots, n\}.$$

Now we can apply Lemma 7.2 for $i = n$, $l = l_n$. If statement (b) of Lemma 7.2 holds then let $l_{n-1} = l'$ and apply the lemma again for $i = n - 1$, $l = l_{n-1}$. If (b) holds again then we continue. Since $n \geq l_n > l_{n-1} > l_{n-2} > \ldots \geq 1$ we cannot repeat this for more than $n - 1$ times, hence for some $1 \leq i \leq n$ (a) of Lemma 7.2 must hold when we apply the lemma for $i, l = l_i$. This way we get $i, l \in \{1, \ldots, n\}$ such that (26) and (a) of Lemma 7.2 hold.

It is easy to see that (26) implies that

$$D + (\underbrace{1, \ldots, 1}_{l}, 0, \ldots, 0) + M(Z^n) = D + M(Z^n)$$

and also that this and (a) of Lemma 7.2 implies that D must be of the form

$$D = \{(\underbrace{a, \ldots, a}_{l}) \; : \; a \in \{0, 1, \ldots, m_1 - 1\}\} \times D',$$

where $D' \subset \{0, 1, \ldots, m_{l+1} - 1\} \times \ldots \times \{0, 1, \ldots, m_n - 1\}$ and $m_1 = \ldots = m_l$. Then $K = K(M, D)$ must be exactly of the claimed form, which completes the proof. □

Now we are ready to characterize those self-affine sponges for which $\mu(K \cap (K+t))$ can be positive for "irregular" translations.

THEOREM 7.4. *Let μ be the natural probability measure on a self-affine Sierpiński sponge $K = K(M, D) \subset \mathbb{R}^n$ (as described in Definition 2.14) and let $t \in \mathbb{R}^n$.*

Then $\mu\big(K \cap (K + t)\big) = 0$ holds except in the following two trivial exceptional cases:

(i) There exists two elementary pieces S_1 and S_2 of K such that $S_2 = S_1 + t$.

(ii) K is of the form $K = L \times K_0$, where L is a diagonal of a cube $[0, 1]^l$, where $l \in \{1, 2, \ldots, n\}$ and K_0 is a smaller dimensional self-affine Sierpiński sponge.

Proof. If $\|M^k t\| = 0$ for some $k \in \{0, 1, 2, \ldots\}$ then for any two k-th generation elementary pieces S_1 and S_2 of K, S_2 and $S_1 + t$ are either identical or $\mu\big((S_1 + t) \cap S_2\big) = 0$. Therefore in this case either (i) or $\mu\big(K \cap (K + t)\big) = 0$ holds, thus we can suppose that $\|M^k t\| > 0$ for every $k = 0, 1, 2, \ldots$ and $\mu\big(K \cap (K + t)\big) > 0$.

Let $D_i = M^{i-1}(D) + M^{i-2}(D) + \ldots + M(D) + D$. Notice that, by definition, $K(M, D) = K(M^i, D_i)$ for any $i \in \mathbb{N}$. Therefore we can apply Proposition 7.1 to (M^n, D_n) to obtain $w \in \{-1, 0, 1\}^n \setminus \{(0, \ldots, 0)\}$ such that

$$D_n + w_n + M^n(\mathbb{Z}^n) = D_n + M^n(\mathbb{Z}^n).$$

Then we can apply Lemma 7.3 to get that $K = K(M, D)$ must be exactly of the form as in (ii) of Theorem 7.4, which completes the proof. □

REMARK 7.5. Clearly, case (i) holds if and only if t is of the form $\sum_{j=1}^{k} M^{-j}(\alpha_j - \beta_j)$, where $k \in \{0, 1, 2, \ldots\}$ and $\alpha_1, \beta_1, \ldots, \alpha_k, \beta_k \in D$.

REMARK 7.6. It follows from the proof that in the coordinates of L every m_i must be the same. Thus in case (ii) we must have $l = 1$ if $m_1, \ldots, m_n$ are all distinct.

In particular, if $n = 1$ then (ii) means $K = [0, 1]$.

The following statement is the analogue of Theorem 4.5.

COROLLARY 7.7. *Let μ be the natural probability measure on a self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ (as described in Definition 2.14) and let $t \in \mathbb{R}^n$.*

The set $K \cap (K + t)$ has positive μ-measure if and only if it has non-empty interior (relative) in K.

Proof. If $K \cap (K + t)$ has non-empty interior in K then clearly $\mu\big(K \cap (K + t)\big) > 0$.

We shall prove the converse by induction. Assume that the converse is true for any smaller dimensional self-affine Sierpiński sponge. Suppose that $\mu\big(K \cap (K+t)\big) > 0$ and apply Theorem 7.4. If (i) of Theorem 7.4 holds then clearly $K \cap (K + t)$ has non-empty interior in K, so we can suppose that (ii) holds: $K = L \times K_0$, L is a diagonal of $[0,1]^l$ and K_0 is a smaller dimensional self-affine Sierpiński sponge. Then $\mu = c\lambda \times \mu_0$, where $1/c$ is the length of L (that is, $c = 1/\sqrt{l}$), λ is the (one-dimensional) Lebesgue measure on L and μ_0 is the natural probability measure on K_0.

Let $t_\alpha = (t_1, \dots, t_l)$ and $t_\beta = (t_{l+1}, \dots, t_n)$ and we suppose that the coordinates of L are the first l coordinates. Then

$$K \cap (K+t) = (L \times K_0) \cap ((L+t_\alpha) \times (K_0+t_\beta)) = (L \cap (L+t_\alpha)) \times (K_0 \cap (K_0+t_\beta)).$$

Therefore we have

$$0 < \mu\big(K \cap (K+T)\big) = c\lambda\big(L \cap (L+t_\alpha)\big) \cdot \mu_0\big(K_0 \cap (K_0+t_\beta)\big)$$

and so $\lambda\big(L \cap (L+t_\alpha)\big) > 0$ and $\mu_0\big(K_0 \cap (K_0+t_\beta)\big) > 0$. This implies that $L \cap (L+t_\alpha)$ has non-empty interior in L and, by our assumption, $K_0 \cap (K_0+t_\beta)$ has non-empty interior in K_0. Thus $K \cap (K+t) = (L \cap (L+t_\alpha)) \times (K_0 \cap (K_0+t_\beta))$ has non-empty interior in $K = L \times K_0$. □

For getting the analogue of Theorem 4.1 we need one more lemma.

PROPOSITION 7.8. *Let $K = K(M,D)$ and μ be like in Definition 2.14, and let $0 \neq t \in \mathbb{R}^n$ be such that $\mu\big(K \cap (K+t)\big) > 1 - \frac{1}{r^2}$.*

Then there exists a

$$w \in \{-1,0,1\} \times \ldots \times \{-1,0,1\} \setminus \{(0,\dots,0)\}$$

such that $D + w = D$ modulo $(m_1, \dots, m_n)$; that is,

$$D + w + M(\mathbb{Z}^n) = D + M(\mathbb{Z}^n).$$

Proof. By Proposition 7.1 we are done if $||M^k t|| > 0$ for every $k = 1, 2, \dots$. Thus we can suppose that this is not the case and choose a minimal $k \in \{1, 2, \dots\}$ such that $||M^k t|| = 0$. Then, letting $u = M^k t$, we have $u \in \mathbb{Z}^n \setminus M(\mathbb{Z}^n)$.

Let

$$D_k = M^{k-1}(D) + M^{k-2}(D) + \ldots + M(D) + D,$$

and define the measure μ_k so that $\mu_k(M^k A) = r^k \mu(A)$ for any Borel set $A \subset K$. Then by definition we have $M^k K = K + D_k$, and for each $d \in D_k$ we have $\mu_k(K+d) = 1$. Using the above facts and definitions and the condition $\mu\big(K \cap (K+t)\big) > 1 - \frac{1}{r^2}$, we get

$$\begin{aligned} r^{k-2}(r^2-1) = r^k\Big(1 - \frac{1}{r^2}\Big) < r^k \mu\big(K \cap (K+t)\big) &= \mu_k\big(M^k K \cap (M^k K + M^k t)\big) \\ &= \mu_k\big((K+D_k) \cap (K+D_k+u)\big) = \#(D_k \cap (D_k+u)), \end{aligned}$$

where $\#(.)$ denotes the number of the elements of a set.

Then by the pigeonhole principle there exists an $e \in M^{k-1}(D) + M^{k-2}(D) + \cdots + M^2(D) \subset M^2(\mathbb{Z}^n)$ such that $e + M(D) + D \subset D_k + u$. This implies that $M(D) + D + M^2(\mathbb{Z}^n) \subset D_k + u + M^2(\mathbb{Z}^n) = M(D) + D + u + M^2(\mathbb{Z}^n)$. Similarly, we can prove that $M(D)+D+u+M^2(\mathbb{Z}^n) \subset M(D)+D+M^2(\mathbb{Z}^n)$. Therefore we get

$$M(D)+D+u+M^2(\mathbb{Z}^n) = M(D)+D+M^2(\mathbb{Z}^n). \tag{28}$$

In particular, we have $D+u+M(\mathbb{Z}^n) = D+M(\mathbb{Z}^n)$.

Then, starting from arbitrary $f_0 \in D$ we can get a sequence $(f_i) \subset D$ so that

$$f_i + u + M(\mathbb{Z}^n) = f_{i+1} + M(\mathbb{Z}^n) \qquad (i = 0,1,2,\ldots). \tag{29}$$

Since $u \notin M(\mathbb{Z}^n)$ we have $f_i \neq f_{i+1}$ for each i. This and the fact that the sequence (f_i) is contained in a finite set imply that there must be a $j \in \mathbb{N}$ such that $f_{j+1} - f_j \neq f_j - f_{j-1}$.

Let $e \in D$ be arbitrary. Applying (28) and (29) we get that there exist $e', e'' \in D$ such that

$$Me' + f_{j-1} + u + M^2(\mathbb{Z}^n) = Me + f_j + M^2(\mathbb{Z}^n)$$

and

$$Me' + f_j + u + M^2(\mathbb{Z}^n) = Me'' + f_{j+1} + M^2(\mathbb{Z}^n),$$

which implies

$$(f_j - f_{j-1}) - (f_{j+1} - f_j) = M(e'' - e) + M^2(\mathbb{Z}^n).$$

Thus there exists a $w \in \mathbb{Z}^n$ such that

$$Mw = (f_j - f_{j-1}) - (f_{j+1} - f_j) = M(e'' - e) + M^2(\mathbb{Z}^n). \tag{30}$$

Since $e, e'', f_{j-1}, f_j, f_{j+1} \in D \subset \{0,1,\ldots,m_1-1\} \times \ldots \times \{0,1,\ldots,m_n-1\}$, (30) implies that

$$e + w + M(\mathbb{Z}^n) = e'' + M(\mathbb{Z}^n)$$

and

$$w \in \{-1,0,1\} \times \ldots \times \{-1,0,1\} \setminus \{(0,\ldots,0)\}.$$

Since $e \in D$ was arbitrary, $e'' \in D$ and w does not depend on e we get that

$$D + w + M(\mathbb{Z}^n) = D + M(\mathbb{Z}^n),$$

which completes the proof. □

THEOREM 7.9. *Let μ be the natural probability measure on a self-affine Sierpiński sponge $K = K(M,D) \subset \mathbb{R}^n$ (as described in Definition 2.14) and let $t \in \mathbb{R}^n$.*

Then $\mu\big(K \cap (K+t)\big) \leq 1 - \frac{1}{r^2}$ holds (where r denotes the number of elements in the pattern D) except in the following two trivial exceptional cases:
(i) $t = 0$.
(ii) K is of the form $K = L \times K_0$, where L is a diagonal of a cube $[0,1]^l$, where $l \in \{1,2,\ldots,n\}$ and K_0 is a smaller dimensional self-affine Sierpiński sponge.

Proof. Suppose that $t \neq 0$ and $\mu(K \cap (K+t)) > 1 - \frac{1}{r^2}$. For $D_n = M^{n-1}(D) + M^{n-2}(D) + \ldots + M(D) + D$, by definition, $K(M,D) = K(M^n, D_n)$. Therefore we can apply Proposition 7.8 to (M^n, D_n) to obtain $w_n \in \{-1,0,1\}^n \setminus \{(0,\ldots,0)\}$ such that

$$D_n + w_n + M^n(\mathbb{Z}^n) = D_n + M^n(\mathbb{Z}^n).$$

Then Lemma 7.3 completes the proof. □

8. *Translation invariant measures for self-affine Sierpiński sponges*

As an easy application of Theorem 7.4 (and Lemma 2.18) we get the following.

THEOREM 8.1. *For any self-affine Sierpiński sponge $K \subset \mathbb{R}^n$ ($n \in \mathbb{N}$) there exists a translation invariant Borel measure ν on $\mathbb{R}^n$ such that $\nu(K) = 1$.*

Proof. Let μ be the natural probability Borel measure on K (see Definition 2.14). We shall prove by induction that μ can be extended to $\mathbb{R}^n$ as a translation invariant Borel measure. Assume that this is true for any smaller dimensional self-affine Sierpiński sponge.

First suppose that K is of the form $K = K_1 \times K_2$, where K_1 and K_2 are smaller dimensional self-affine Sierpiński sponges. Then $\mu = \mu_1 \times \mu_2$, where μ_1 and μ_2 are the natural probability Borel measures on K_1 and K_2, respectively. Then, by our assumption, μ_1 and μ_2 has translation invariant extensions $\tilde{\mu}_1$ and $\tilde{\mu}_2$ and then one can easily check that $\tilde{\mu} = \tilde{\mu}_1 \times \tilde{\mu}_2$ is a translation invariant Borel measure on $\mathbb{R}^n$ and an extension of μ.

If K is not of the form $K = K_1 \times K_2$ then we shall check that condition (6) of Lemma 2.18 is satisfied, so then Lemma 2.18 will complete the proof. Fix $B \subset K$ and $t \in \mathbb{R}^n$ such that $B + t \subset K$. Then $B \subset K \cap (K - t)$ and $B + t \subset K \cap (K + t)$, so we have $\mu(B) = 0 = \mu(B + t)$ unless

$$\mu\big(K \cap (K + t)\big) > 0 \quad \text{or} \quad \mu\big(K \cap (K - t)\big) > 0. \tag{31}$$

By Theorem 7.4 and since case (ii) of Theorem 7.4 is already excluded, (31) implies (i) of Theorem 7.4. On the other hand, if (i) of Theorem 7.4 holds then the translation by t maps elementary pieces of B to elementary pieces of $B + t$ and then the condition (6) clearly holds.

Since we checked all cases, the proof is complete. □

We also show a more direct proof for the above theorem, which does not use Theorem 7.4 and which works for a slightly larger class of self-affine sets.

THEOREM 8.2. *Let φ be an injective affine map which is contractive (in some norm), $t_1, \dots, t_r \in \mathbb{R}^n$ and $K \subset \mathbb{R}^n$ the compact self-affine set such that $K = \cup_{i=1}^r \varphi(K) + t_i$. Suppose that the natural probability measure on K has the property that*

$$\mu\Big(K \cap \Big(\big((\varphi(K) + t_i) \cap (\varphi(K) + t_j)\big) + u\Big)\Big) = 0 \quad (\forall\ 1 \le i < j \le r,\ u \in \mathbb{R}^n). \tag{32}$$

(a) Then for any $t \in \mathbb{R}^n$ and elementary piece S of K we have

$$\mu\big(K \cap (S + t)\big) \le \mu(S).$$

(b) There exists a translation invariant Borel measure ν on $\mathbb{R}^n$ such that $\nu(K) = 1$. In fact, ν is an extension of μ.

Proof. First we prove (a). Suppose that S is a k-th generation elementary piece of K. Then K can be written as

$$K = \cup_{j=1}^{r^k} S + h_j$$

for some $h_1, \ldots, h_{r^k} \in \mathbb{R}^n$ and by (32) the sets $S + h_j$ are pairwise almost disjoint.

Using this and that $\mu(A) = \mu(A + h_j)$ for any Borel set $A \subset S$ we get that

$$\begin{aligned}
\mu\big(K \cap (S+t)\big) &= \mu\Big(\bigcup_{j=1}^{r^k} (S+h_j) \cap (S+t)\Big) \\
&= \sum_{j=1}^{r^k} \mu\Big((S+h_j) \cap (S+t)\Big) \\
&= \sum_{j=1}^{r^k} \mu\Big(\big(S \cap (S+t-h_j)\big) + h_j\Big) \\
&= \sum_{j=1}^{r^k} \mu\Big(S \cap (S+t-h_j)\Big). \qquad (33)
\end{aligned}$$

Using (32) we get that for any $i \neq j$ we have

$$\begin{aligned}
&\mu\Big(\big(S \cap (S+t-h_i)\big) \cap \big(S \cap (S+t-h_j)\big)\Big) \\
&\qquad = \mu\Big(S \cap \Big(\big((S+h_j) \cap (S+h_i)\big) + t - h_i - h_j\Big)\Big) = 0. \qquad (34)
\end{aligned}$$

Thus we can continue (33) as

$$\mu\big(K \cap (S+t)\big) = \sum_{j=1}^{r^k} \mu\big(S \cap (S+t-h_j)\big) = \mu\left(S \cap \bigcup_{j=1}^{r^k} (S+t-h_j)\right) \leq \mu(S),$$

which completes the proof of (a).

For proving (b) define

$$\nu(H) = \inf\left\{ \sum_{j=1}^{\infty} \mu(S_j) : H \subset \cup_{j=1}^{\infty} S_j + u_j,\ S_j \text{ is an elem. piece of } K, u_j \in \mathbb{R}^n \right\}$$

for any $H \subset \mathbb{R}^n$. Then ν is clearly a translation invariant outer measure on $\mathbb{R}^n$.

We claim that ν is a metric outer measure; that is, $\nu(A \cup B) = \nu(A) + \nu(B)$ if $A, B \subset \mathbb{R}^n$ have positive distance. Indeed, in this case in the cover $A \cup B \subset \cup_{j=1}^{\infty} S_j + u_j$ in the definition of $\nu(A \cup B)$ we can replace replace each S_j by its small elementary pieces such that each small elementary piece covers only at most one of A and B. Since this transformation does not change $\sum_{j=1}^{\infty} \mu(S_j)$ this implies that $\nu(A \cup B) \geq \nu(A) + \nu(B)$. Since ν is an outer measure we get that $\nu(A \cup B) = \nu(A) + \nu(B)$.

It is well known (see e.g. in [**6**]) that restricting a metric outer measure to the Borel sets we get a Borel measure.

So it is enough to prove that $\nu(K) = 1$. The definition of $\nu(K)$ implies that $\nu(K) \leq \mu(K) = 1$.

For proving $\nu(K) \geq 1$ let $K \subset \cup_{j=1}^{\infty} S_j + u_j$ be an arbitrary cover such that each S_j is an elementary piece of K and $u_j \in \mathbb{R}^n$. Then, using the already proved (a) part we get that

$$\sum_{j=1}^{\infty}\mu(S_j) \geq \sum_{j=1}^{\infty}\mu\big(K\cap(S_j+u_j)\big) \geq \mu\Big(\bigcup_{j=1}^{\infty}(K\cap(S_j+u_j))\Big) = \mu(K),$$

which completes the proof of (b). □

Using Lemma 2.9, the above theorem has the following consequence.

COROLLARY 8.3. *Let $K = \varphi_1(K)\cup\ldots\cup\varphi_r(K)$ be a self-affine set with the convex open set condition and suppose that $\varphi_1(K),\ldots,\varphi_r(K)$ are translates of each other.*

Then the natural probability measure on K can be extended as a translation invariant measure on $\mathbb{R}^n$. □

9. *Concluding remarks*

Our results might be true for much larger classes of self-similar or self-affine sets. We have no counter-example even for the strongest very naive conjecture that the intersection of any two affine copies of *any* self-affine set is of positive measure (according to any self-affine measure on one of the copies) if and only if it contains a set which is open in both copies.

We do not even know whether this very naive conjecture holds at least for two isometric copies of a self-affine Sierpiński sponge. (Note that if we allow only translated copies then Corollary 7.7 provides an affirmative answer.) For generalizing our results about Sierpiński sponges from translates to isometries the following statement could help.

CONJECTURE 9.1. *If K is a self-affine sponge, μ is the natural probability measure on it, φ is an isometry and $\mu\big(K\cap\varphi(K)\big)>0$ then there exists a translation t such that $K\cap\varphi(K) = K\cap(K+t)$.*

This conjecture and the above mentioned Corollary 7.7 would clearly imply that Corollary 7.7 holds for isometric copies of self-affine Sierpiński sponges as well. Then, in the same way as Theorem 8.1 is proved, we could get an *isometry* invariant Borel measure ν for an arbitrary Sierpiński sponge K such that $\nu(K)=1$.

For getting this stronger version of Theorem 8.1 the other natural way could be a generalization of Theorem 8.2 for isometries at least for self-affine Sierpinśki sponges. Since part (b) of Theorem 8.2 follows from (a) for isometries as well it would be enough to show (a), that is, it would be enough to show that $\mu\big(K\cap\varphi(S)\big)\leq\mu(S)$, for any elementary piece S of any self-affine Sierpiński sponge K with natural measure μ. We do not know whether this last mentioned statement holds or not.

As we saw in Theorem 7.9, the instability results are not true for arbitrary self-affine sets, not even for self-similar sets with the open set condition: the simplest counter-example is $K = C\times[0,1]$, where C denotes the classical triadic Cantor set. Then K is self-similar (with six similitudes of ratio 1/3), the open set condition clearly holds and if μ is the evenly distributed self-similar measure on K (that is, $p_1=\ldots=p_6$) then $\mu\big(K\cap(K+(0,\varepsilon)\big) = 1-\varepsilon$. The instability results might be true for totally disconnected (which means that each connected component is a singleton) self-affine sets.

In the definition of self-affine sets we allowed only contractive affine maps. If we allowed non-contractive affine maps as well then the above $K = C \times [0,1]$ set would be a self-affine set (with two affine maps) with the strong separation condition, so it would be a counter-example for both theorems (Theorem 3.2 and Theorem 3.5) about self-affine sets.

We do not know whether the analogues of Theorem 4.1, Theorem 4.5 and Corollary 4.7 hold for self-affine sets with the strong separation condition. Although Theorem 3.5 says that for self-affine sets and isometries the analogue of Theorem 4.1 holds, and Theorem 4.5 was proved from Theorem 4.1, we cannot get the same way that for self-affine sets and at least for isometries the analogue of Theorem 4.5 holds. This is because in the proof of Theorem 4.5 it was important that the maps $\varphi_1, \dots, \varphi_r$ that generated the self-similar sets were also in the group (in this case the group of similitudes) for which we had Theorem 4.1. In order to get any analogue of Theorem 4.5 for self-affine sets in the same way we need to prove a self-affine analogue of Theorem 4.1 for a group of transformation containing the affine maps $\varphi_1, \dots, \varphi_r$ that generates the self-affine set.

From a positive answer for the following question we could get fairly easily that the self-affine analogue of Theorem 4.1 holds at least for affine maps from any compact subset of the space of affine maps. Then, if we could also show that we can assume that the affine maps are from a compact set (as in Proposition 4.2 for similitudes) then we would get that all the main results of Section 4 also hold for self-affine sets and affine maps as well.

QUESTION 9.2. *Let $K \subset \mathbb{R}^d$ be a self-affine set satisfying the strong separation condition and let f be an affine map such that $f(K) \subset K$. Does this imply that $f(K)$ is a relative open set in K?*

Note that for $f(K)$ being a relative open set in K means that it is the union of countably many pairwise disjoint elementary pieces of K, and since $f(K)$ is compact this means that $f(K)$ is a finite union of elementary pieces of K.

A positive answer at least for the following self-similar special case of the above question could make the proof of Theorem 4.1 simpler. However, we cannot answer this question even for $d = 1$.

QUESTION 9.3. *Let $K \subset \mathbb{R}^d$ be a self-similar set satisfying the strong separation condition and let f be a similitude such that $f(K) \subset K$. Does this imply that $f(K)$ is a relative open set in K (or in other words $f(K)$ is a finite union of elementary pieces of K)?*

Note that in Section 6 we saw that self-similar set (even in $\mathbb{R}$) may contain similar copies of itself in non-trivial ways.

REFERENCES

[1] C. Bandt and S. Graf, Self-similar sets. VII. A characterization of self-similar fractals with positive Hausdorff measure, *Proc. Amer. Math. Soc.* **114** (1992), no. 4, 995–1001.

[2] T. Bedford, *Crinkly curves, Markov partitions and box dimension in self-similar sets*, Ph. D. Thesis, University of Warwick, 1984.

[3] R. O. Davies, Sets which are null or non-sigma-finite for every translation-invariant measure, *Mathematika* **18** (1971), 161–162.

[4] M. Elekes and T. Keleti, Borel sets which are null or non-σ-finite for every translation invariant measure, *Adv. Math.* **201** (2006), 102-115.

[5] K. J. Falconer, Classes of sets with large intersection, *Mathematika* **32** (1985), no. 2, 191–205.

[6] K. J. Falconer, *The geometry of fractal sets.* Cambridge Tracts in Mathematics No. 85, Cambridge University Press, 1985.

[7] K. J. Falconer, *Techniques in Fractal Geometry*, John Wiley & Sons, 1997.

[8] D. Feng and Y. Wang, *On the structures of generating iterated function systems of Cantor sets*, preprint.

[9] H. Furstenberg, Intersections of Cantor sets and transversality of semigroups, *Problems in analysis (Sympos. Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*, pp. 41-59, Princeton Univ. Press, 1970.

[10] D. Gatzouras and S. Lalley (1992), Hausdorff and box dimensions of certain self-affine fractals, *Indiana University Math. J.* **41** (1992), 533–568.

[11] J. E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747.

[12] K. Igudesman, Lacunary self-similar fractal sets and its application to intersection of Cantor sets, *Lobachevskii J. Math.* **12** (2003), 41–50.

[13] M. Järvenpää, Hausdorff and packing dimensions, intersection measures, and similarities, *Ann. Acad. Sci. Fenn. Math.* **24** (1999), no. 1., 165–186.

[14] T. Keleti, A 1-dimensional subset of the reals that intersects each of its translates in at most a single point, *Real Analysis Exchange* **24** (1998/99), 843–844.

[15] R. Kenyon and Y. Peres, Measures of full dimension on affine-invariant sets, *Ergodic Theory Dynamical Syst.* **16** (1996), 307–323.

[16] J. C. Lagarias and Y. Wang, Self-affine tiles in R^n, *Adv. Math.* **121** (1996), no. 1, 21–49.

[17] W. Li and D. Xiao, Intersection of translations of Cantor triadic set, *Acta Math. Sci. (English Ed.)* **19** (1999), no. 2, 214–219.

[18] A. Máthé, *Önhasonló halmazok egybevágóság-invariáns mértékeiről* (On isometry invariant measures of self-similar sets) (In Hungarian), Master Thesis, Eötvös Loránd University, 2005 (`http://www.cs.elte.hu/math/diploma/math`).

[19] P. Mattila, On the structure of self-similar fractals, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **7** (1982), no. 2., 189–195.

[20] P. Mattila, Hausdorff dimensions and capacities of intersections of sets in n-space, *Acta Math.* **152** (1984), no. 1-2, 77–105.

[21] P. Mattila, On the Hausdorff dimension and capacities of intersections, *Mathematika* **32** (1985), no. 2, 213–217.

[22] C. McMullen, The Hausdorff dimension of general Sierpiński carpets, *Nagoya Math. J.* **96** (1984), 1–9.

[23] C. G. T. de A. Moreira, Stable intersections of Cantor sets and homoclinic bifurcations, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13**, no. 6, 741–781.

[24] C. G. T. de A. Moreira and J-C. Yoccoz, Stable intersections of regular Cantor sets with large Hausdorff dimensions, *Ann. of Math. (2)* **154** (2001), no. 1, 45–96.

[25] F. Nekka and J. Li, Intersection of triadic Cantor sets with their translates. I. Fundamental properties, *Chaos Solitons Fractals* **13** (2002), no. 9, 1807–1817.

[26] Y. Peres and B. Solomyak, Self-similar measures and intersections of Cantor sets, *Trans. Amer. Math. Soc.* **350** (1998), no. 10, 4065–4087.

[27] Y. Peres, The packing measure of self-affine carpets, *Math. Proc. Cambridge Philos. Soc.* **115** (1994), no. 3, 437–450.

[28] Y. Peres, The self-affine carpets of McMullen and Bedford have infinite Hausdorff measure,

Math. Proc. Cambridge Philos. Soc. **116** (1994), 513–526.

[29] W. Rudin, *Real and complex analysis*, Third Edition, McGraw-Hill Book Company, 1987.

[30] A. Schief, Spearation properties for self-similar sets, *Proc. Amer. Math. Soc.* **122** (1994), no. 1, 111–115.

Available online at www.sciencedirect.com

ScienceDirect

J. Math. Anal. Appl. 337 (2008) 1394–1403

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

Periodic decomposition of measurable integer valued functions ☆

Tamás Keleti

Department of Analysis, Eötvös Loránd University, Pázmány Péter sétány 1/c, H-1117 Budapest, Hungary

Received 15 January 2007

Available online 18 April 2007

Submitted by B. Bongiorno

Abstract

We study those functions that can be written as a sum of (almost everywhere) integer valued periodic measurable functions with given periods. We show that being (almost everywhere) integer valued measurable function and having a real valued periodic decomposition with the given periods is not enough. We characterize those periods for which this condition is enough. We also get that the class of bounded measurable (almost everywhere) integer valued functions does not have the so-called decomposition property. We characterize those periods $a_1, \dots, a_k$ for which an almost everywhere integer valued bounded measurable function f has an almost everywhere integer valued bounded measurable $(a_1, \dots, a_k)$-periodic decomposition if and only if $\Delta_{a_1} \cdots \Delta_{a_k} f = 0$, where $\Delta_a f(x) = f(x+a) - f(x)$.

Keywords: Periodic functions; Measurable functions; Periodic decomposition; Integer valued functions; Almost everywhere integer valued functions; Real valued functions; Difference operator; Decomposition property

0. Introduction

In [7] those functions were studied that can be written as a sum of periodic integer valued functions with given periods $a_1, \dots, a_k$. Clearly these functions must be integer valued and they can be written as a sum of periodic real valued functions with given periods $a_1, \dots, a_k$. Several results were proved about the question whether the converse is true or false; that is, whether the existence of a real valued $(a_1, \dots, a_k)$-periodic decomposition of an integer valued function implies the existence of an integer valued $(a_1, \dots, a_k)$-periodic decomposition. Among others the following result were proved:

Theorem 0.1. *(See [7, Theorem 2.1].) Suppose that an integer valued function $f : \mathbb{Z} \to \mathbb{Z}$ can be written as $f = g_1 + \cdots + g_k$, where each g_j is a real valued a_j-periodic function for some $a_j \in \mathbb{Z}$.*

Then f can be also written as $f = h_1 + \cdots + h_k$, where each h_j is an integer valued a_j-periodic function.

☆ Supported by the Hungarian Scientific Foundation Grants Nos. F 43620 and T 49786. This research started when the author was a visitor at the Alfréd Rényi Institute of Mathematics of the Hungarian Academy of Science.

E-mail address: elek@cs.elte.hu.

URL: http://www.cs.elte.hu/anal/keleti.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2007.04.013

The question whether the same result holds for functions defined on $\mathbb{R}$ is still open:

Question 0.2. *(See [7, Question 5.1].) Is it true for any $a_1, \dots, a_k \in \mathbb{R}$ that if an integer valued function $f : \mathbb{R} \to \mathbb{Z}$ has a real valued $(a_1, \dots, a_k)$-periodic decomposition, then f also has an integer valued $(a_1, \dots, a_k)$-periodic decomposition?*

Added in proof: A positive answer to this question was meanwhile proved in [4].

There are some positive partial results if we have some assumptions about the periods $a_1, \dots, a_k$ (see [7] and [4]).

For bounded decomposition of bounded functions the following counter-example was found for the analogue question:

Theorem 0.3. *(See [7, Theorem 3.1].) There exists a function $u : \mathbb{Z} \times \mathbb{Z} \to \{0, 1\}$ that can be written as a sum of a $(0, 1)$-periodic, a $(1, 0)$-periodic and a $(1, 1)$-periodic bounded $\mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ function, can be written also as the sum of three periodic $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ functions with the same periods but cannot be written as a sum of three periodic bounded $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ functions with the same periods.*

Note that by repeating this construction on each coset one can get a similar counter-example on any Abelian group that contains a $\mathbb{Z} \times \mathbb{Z}$ subgroup, so in particular there exist $a_1, a_2, a_3 \in \mathbb{R}$ and a function $f : \mathbb{R} \to \{0, 1\}$ that has a bounded real valued (a_1, a_2, a_3)-periodic decomposition but does not have a bounded integer valued (a_1, a_2, a_3)-periodic decomposition.

In the above cited results no measurability was assumed. In this paper we will study what happens if we allow only measurable functions. First we give a negative answer (Theorem 1.2) to the measurable analogue of Question 0.2 and then we characterize (Theorem 2.5) those periods for which we have positive result, at least if we are happy with almost everywhere integer valued decompositions. Everywhere integer valued measurable decompositions are studied in Section 3. It turns out that the question whether we can get integer valued decompositions instead of almost everywhere integer valued decompositions depends on the answers to the nonmeasurable questions like the above mentioned Question 0.2.

The characterization of those functions that can be written as a sum of periodic functions with given period has a much longer history. It started in the seventies with some unpublished work of I.Z. Ruzsa and continued among others in [1,2,4–6,8–13] and [14]. If $f = f_1 + \dots + f_n$ is an $(a_1, \dots, a_n)$-periodic decomposition of f, then

$$\Delta_{a_1} \Delta_{a_2} \cdots \Delta_{a_n} f = 0, \quad \text{where } \Delta_{a_j} f(x) = f(x + a_j) - f(x), \tag{1}$$

since the difference operators Δ_{a_j} commute. A class of functions $\mathcal{F}$ is said to have the *decomposition property* if every $f \in \mathcal{F}$ that satisfies (1) has an $(a_1, \dots, a_k)$-periodic decomposition in $\mathcal{F}$. Since for the identity function $f(x) = x$ we clearly have $\Delta_1 \Delta_1 f = 0$ but it is not the sum of two 1-periodic functions, many natural classes of functions (e.g., all $\mathbb{R} \to \mathbb{R}$ functions, continuous $\mathbb{R} \to \mathbb{R}$ functions) do not have the decomposition property. However, many classes of functions do have the decomposition property: for example the class of all bounded continuous $\mathbb{R} \to \mathbb{R}$ functions [10], the class of all bounded $\mathbb{R} \to \mathbb{R}$ functions [2,11], the class of all bounded measurable $\mathbb{R} \to \mathbb{R}$ functions [11] and the class of all bounded real valued functions on an arbitrary Abelian group [11].

For integer valued functions it was proved in [7] that the class of bounded $\mathbb{Z} \to \mathbb{Z}$ functions has the decomposition property but the class of bounded $\mathbb{R} \to \mathbb{Z}$ functions does not have the decomposition property. In fact, among the torsion free Abelian groups only the additive subgroups of $\mathbb{Q}$ are those on which the class of bounded integer valued functions has the decomposition property [7, Corollary 3.5].

In this note we get (Corollary 1.3) that on $\mathbb{R}$ the classes of bounded measurable (almost everywhere) integer valued functions and (almost everywhere) integer valued L_∞ functions do not have the decomposition property. We characterize (Theorem 2.5) those periods $a_1, \dots, a_k$ for which for any bounded measurable $\mathbb{R} \to \mathbb{Z}$ function the existence of a bounded measurable almost everywhere integer valued $(a_1, \dots, a_k)$-periodic decomposition is equivalent to (1). We show (Proposition 3.4) that this characterization is not valid for everywhere integer valued $(a_1, \dots, a_k)$-periodic decompositions.

Meanwhile, as a spin-off, we also characterize (Corollary 2.4) those periods $a_1, \dots, a_k$ for which the measurable $(a_1, \dots, a_k)$-periodic decomposition of an $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ function is essentially unique. This characterization turns out to be different for $\mathbb{R} \to \mathbb{R}$ functions (Lemma 1.1).

1. A negative result

The following fact is known, see, e.g., [8].

Lemma 1.1. *Let $a_1, \dots, a_k \in \mathbb{R} \setminus \{0\}$ such that $a_i/a_j \notin \mathbb{Q}$ for any $i \neq j$ and suppose that $f_1 + \dots + f_k = g_1 + \dots + g_k$ and for each j, f_j and g_j are a_j-periodic measurable $\mathbb{R} \to \mathbb{R}$ functions.*

Then $f_j - g_j$ is almost everywhere constant for every $j = 1, \dots, k$.

It is easy to see that the condition $a_i/a_j \notin \mathbb{Q}$ for $i \neq j$ is also necessary. We will see (Corollary 2.4) that for $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ functions the necessary and sufficient condition for $a_1, \dots, a_k$ is stronger.

The following theorem shows that the existence of a measurable real valued $(a_1, \dots, a_n)$-periodic decomposition for an $\mathbb{R} \to \mathbb{Z}$ function does not always implies the existence of a measurable integer valued $(a_1, \dots, a_n)$-periodic decomposition, not even the existence of a measurable, almost everywhere integer valued $(a_1, \dots, a_n)$-periodic decomposition.

Theorem 1.2. *There exists an integer valued bounded Lebesgue measurable function on the real line that can be written as a sum of three real valued bounded measurable periodic functions but cannot be written as a sum of three almost everywhere integer valued measurable periodic functions with the same periods.*

Proof. Let $t \in \mathbb{R} \setminus \mathbb{Q}$ be arbitrary and let

$$f(x) = \{tx\} + \big\{(1-t)x\big\} + \{-x\}, \tag{2}$$

where $\{.\}$ denotes the fractional part; that is, $\{a\} = a - [a]$. Then f is clearly measurable and it is the sum of a $\frac{1}{t}$-periodic, a $\frac{1}{1-t}$-periodic and a 1-periodic bounded measurable function. Noting that f can be also written as

$$f(x) = tx - [tx] + (1-t)x - \big[(1-t)x\big] - x - [-x] = -[tx] - \big[(1-t)x\big] - [-x],$$

we get that f is integer valued.

Suppose that $f = g_1 + g_2 + g_3$ and g_1, g_2, g_3 are measurable almost everywhere integer valued periodic measurable functions with periods $\frac{1}{t}$, $\frac{1}{1-t}$ and 1, respectively. Since $t \notin \mathbb{Q}$ and by adding a constant to $\{-x\}$ we cannot get an almost everywhere integer valued function, applying Lemma 1.1 for $\{tx\} + \{(1-t)x\} + \{-x\} = g_1 + g_2 + g_3$, we get a contradiction. □

Corollary 1.3. *The following classes of functions do not have the decomposition property*: $\{f : \mathbb{R} \to \mathbb{Z}:\ f \in L_\infty\}$, $\{f : \mathbb{R} \to \mathbb{Z}:\ f$ *is bounded and measurable*$\}$, $\{f : \mathbb{R} \to \mathbb{R}:\ f \in L_\infty$ *and f is almost everywhere integer valued*$\}$, *and* $\{f : \mathbb{R} \to \mathbb{R}:\ f$ *is bounded, measurable and almost everywhere integer valued*$\}$.

Proof. Let t and f be as in the above proof. Then f is contained in all the above classes, $\Delta_1 \Delta_{1/t} \Delta_{1/(1-t)} f = 0$ but, as we saw it in the above proof, f cannot be written as a sum of a $\frac{1}{1-t}$-periodic, a $\frac{1}{t}$ and a 1-periodic measurable almost everywhere integer valued function. □

2. Almost everywhere integer valued decompositions

The following lemma might be known but for completeness we present a proof.

Lemma 2.1. *If $E_j \subset \mathbb{R}$ is an a_j-periodic measurable set with positive measure for each $j = 1, \dots, k$ and $\frac{1}{a_1}, \dots, \frac{1}{a_k}$ are linearly independent over $\mathbb{Q}$, then $E_1 \cap \dots \cap E_k$ has positive measure.*

Proof. By applying the Lebesgue's Density Theorem for each E_j, we can find $\delta > 0$ and $d_1, \dots, d_k \in \mathbb{R}$ such that

$$\lambda\big((d_j - 2\delta, d_j + 2\delta) \setminus E_j\big) < \frac{2\delta}{k} \quad (j = 1, \dots, k).$$

For each $j = 1, \dots, k$, using that E_j is a_j-periodic, we get that for any $m_j \in \mathbb{Z}$,

$$t \in (m_j a_j + d_j - \delta, m_j a_j + d_j + \delta) \quad \Rightarrow \quad \lambda\big((t - \delta, t + \delta) \setminus E_j\big) < \frac{2\delta}{k}. \tag{3}$$

One form of Kronecker's theorem (see, e.g., [3, Theorem 444]) states that if $b_1, \dots, b_k \in \mathbb{R}$ are linearly independent over $\mathbb{Q}$, $c_1, \dots, c_k \in \mathbb{R}$ and $\varepsilon > 0$, then there exist $t \in \mathbb{R}$ and $m_1, \dots, m_k \in \mathbb{Z}$ such that $|b_j t - m_j - c_j| < \varepsilon$ for every $j = 1, \dots, k$.

Applying the above mentioned form of Kronecker's theorem for $b_j = \frac{1}{a_j}$, $c_j = \frac{d_j}{a_j}$ and $\varepsilon = \frac{\delta}{a_j}$ we get $t \in \mathbb{R}$ such that $|t - m_j a_j - d_j| < \delta$ for every j.

Then by (3),

$$\lambda\big((t - \delta, t + \delta) \setminus E_j\big) < \frac{2\delta}{k}$$

for every $j = 1, \dots, k$, which implies that $\lambda(E_1 \cap \dots \cap E_k \cap (t - \delta, t + \delta)) > 0$. $\square$

We remark that Lemma 2.1 easily implies the following statement. In fact, the converse implication is also easy.

Corollary 2.2. *If $f_1, \dots, f_k : \mathbb{R} \to (0, \infty)$ are periodic measurable functions such that the reciprocals of the periods are linearly independent over $\mathbb{Q}$, then*

$$\|f_1 + \dots + f_k\|_\infty = \|f_1\|_\infty + \dots + \|f_k\|_\infty.$$

The following theorem shows that if $\frac{1}{a_1}, \dots, \frac{1}{a_k}$ are linearly independent over $\mathbb{Q}$, then the almost everywhere integer valued measurable functions have only trivial measurable $(a_1, \dots, a_k)$-periodic decompositions.

Theorem 2.3. *Let $a_1, \dots, a_k \in \mathbb{R} \setminus \{0\}$ such that $\frac{1}{a_1}, \dots, \frac{1}{a_k}$ are linearly independent over $\mathbb{Q}$. Suppose that $f_j : \mathbb{R} \to \mathbb{R}$ is an a_j-periodic measurable function for each $j = 1, \dots, k$ and that $f = f_1 + \dots + f_k$ is an almost everywhere integer valued function.*

Then each fractional part $\{f_j\}$ is constant almost everywhere.

Proof. Let

$$F_j = \bigcup \big\{ f_j^{-1}((r, q)) \colon r, q \in \mathbb{Q},\ \lambda\big(f_j^{-1}((r, q))\big) = 0 \big\} \quad (j = 1, \dots, k),$$

$$E = \big\{ x \in \mathbb{R} \colon f(x) \in \mathbb{Z} \big\} \setminus \bigcup_{j=1}^{k} F_j.$$

Then $\lambda(\mathbb{R} \setminus E) = 0$, so it is enough to prove that for any fixed $u, v \in E$ and $\varepsilon > 0$, we have $\|f_1(u) - f_1(v)\| < \varepsilon$, where $\|.\|$ denotes the distance from the nearest integer; that is, $\|x\| = \min(\{x\}, \{1 - x\})$.

Let

$$E_1 = \left\{ x \in \mathbb{R} \colon \big|f_1(x) - f_1(u)\big| < \frac{\varepsilon}{k} \right\}$$

and

$$E_j = \left\{ x \in \mathbb{R} \colon \big|f_j(x) - f_j(v)\big| < \frac{\varepsilon}{k} \right\} \quad (j = 2, 3, \dots, k).$$

For each $j = 1, \dots, k$, the set E_j is measurable and a_j-periodic since f_j is measurable and a_j-periodic, and $\lambda(E_j) > 0$ since $u, v \in E$ and so $u, v \notin F_j$. Hence by Lemma 2.1, $\lambda(E_1 \cap \dots \cap E_k) > 0$, so there exists $t \in E \cap E_1 \cap \dots \cap E_k$.

Then

$$\left|f_1(u)-f_1(t)\right|<\frac{\varepsilon}{k}, \tag{4}$$

$$\left|f_j(v)-f_j(t)\right|<\frac{\varepsilon}{k} \quad (j=2,\dots,k). \tag{5}$$

Since $f_1(v)+f_2(v)+\dots+f_k(v)=f(v)\in\mathbb{Z}$ and $f_1(t)+f_2(t)+\dots+f_k(t)=f(t)\in\mathbb{Z}$, (5) implies that $\|f_1(t)-f_1(v)\|<(k-1)\frac{\varepsilon}{k}$. Combining this with (4) we get that $\|f_1(u)-f_1(v)\|<\varepsilon$, which completes the proof. □

Now we can characterize those periods for which the measurable decomposition of an $\mathbb{R}\to\mathbb{R}/\mathbb{Z}$ function is unique up to additive constants. Note that, by Lemma 1.1, the characterization is different for $\mathbb{R}\to\mathbb{R}$ functions.

Corollary 2.4. *For any $a_1,\dots,a_k\in\mathbb{R}\setminus\{0\}$ the following two statements are equivalent.*

(i) *If $f_1+\dots+f_k=g_1+\dots+g_k$ and for each j, f_j and g_j are a_j-periodic measurable $\mathbb{R}\to\mathbb{R}/\mathbb{Z}$ functions, then f_j-g_j is almost everywhere constant for every $j=1,\dots,k$.*
(ii) *$\frac{1}{a_1},\dots,\frac{1}{a_k}$ are linearly independent over $\mathbb{Q}$.*

Proof. (i) ⇒ (ii). Suppose that (ii) is false, so there exists $(m_1,\dots,m_k)\in\mathbb{Z}\times\dots\times\mathbb{Z}\setminus\{0,\dots,0\}$ such that

$$\frac{m_1}{a_1}+\dots+\frac{m_k}{a_k}=0.$$

Then $f_j(x)=\frac{m_j}{a_j}x \bmod 1$ for $j=1,\dots,k$ and $g_1=\dots=g_k=0$ shows that (i) is also false.

(ii) ⇒ (i). This follows simply from Theorem 2.3. □

Now we can characterize those periods for which the existence of a (bounded) measurable real valued $(a_1,\dots,a_k)$-periodic decomposition of an integer valued or almost everywhere integer valued function implies the existence of a (bounded) measurable almost everywhere integer valued $(a_1,\dots,a_k)$-periodic decomposition. We will get the same characterization for those periods for which an integer valued or almost everywhere integer valued bounded measurable function has a bounded measurable almost everywhere integer valued decomposition if and only if $\Delta_{a_1}\cdots\Delta_{a_k}f=0$, where $\Delta_a f(x)=f(x+a)-f(x)$.

Theorem 2.5. *For any $a_1,\dots,a_k\in\mathbb{R}\setminus\{0\}$ the following seven statements are equivalent.*

(i)/(i′) *If an everywhere/almost everywhere integer valued measurable function f on $\mathbb{R}$ can be decomposed as $f=f_1+\dots+f_k$ such that each f_j is an a_j-periodic measurable $\mathbb{R}\to\mathbb{R}$ function, then f can be also decomposed as $f=g_1+\dots+g_k$ such that each g_j is an a_j-periodic almost everywhere integer valued measurable function.*
(ii)/(ii′) *If an everywhere/almost everywhere integer valued bounded measurable function f on $\mathbb{R}$ can be decomposed as $f=f_1+\dots+f_k$ such that each f_j is an a_j-periodic bounded measurable $\mathbb{R}\to\mathbb{R}$ function, then f can be also decomposed as $f=g_1+\dots+g_k$ such that each g_j is an a_j-periodic almost everywhere integer valued bounded measurable function.*
(iii)/(iii′) *An everywhere/almost everywhere integer valued bounded measurable function f on $\mathbb{R}$ can be decomposed as $f=g_1+\dots+g_k$ such that each g_j is an a_j-periodic almost everywhere integer valued bounded measurable function if and only if $\Delta_{a_1}\cdots\Delta_{a_k}f=0$.*
(iv) *If $B_1,\dots,B_n$ are the equivalence classes of $\{a_1,\dots,a_k\}$ with respect to the relation $a\sim b\Leftrightarrow a/b\in\mathbb{Q}$, and b_j denotes the smallest common multiple of the numbers in B_j (for each $j=1,\dots,n$), then $\frac{1}{b_1},\dots,\frac{1}{b_n}$ are linearly independent over $\mathbb{Q}$.*

Proof. (i) ⇔ (i′), (ii) ⇔ (ii′), (iii) ⇔ (iii′). The ⇐ implications are clear. Now we prove (i) ⇒ (i′), the other ⇒ implications can be proved in the same way.

Suppose that $f:\mathbb{R}\to\mathbb{R}$, $H\subset\mathbb{R}$ has measure zero, f is integer valued on $\mathbb{R}\setminus H$ and $f=f_1+\cdots+f_k$ is a real valued $(a_1,\dots,a_k)$-periodic decomposition of f. We need to find an almost everywhere integer valued $(a_1,\dots,a_k)$-periodic decomposition of f.

By replacing H by $\{h+n_1a_1+\cdots+n_ka_k\colon h\in H,\ n_1,\dots,n_k\in\mathbb{Z}\}$ if necessary we can suppose that the characteristic functions χ_H and $\chi_{\mathbb{R}\setminus H}$ are a_j-periodic for every $j=1,\dots,k$.

Applying (i) to the integer valued function $\chi_{\mathbb{R}\setminus H}\cdot f$ we get an almost everywhere integer valued $(a_1,\dots,a_k)$-periodic decomposition $\chi_{\mathbb{R}\setminus H}\cdot f=g_1+\cdots+g_k$. Multiplying by $\chi_{\mathbb{R}\setminus H}$ we get

$$\chi_{\mathbb{R}\setminus H}\cdot f=\chi_{\mathbb{R}\setminus H}\cdot g_1+\cdots+\chi_{\mathbb{R}\setminus H}\cdot g_k. \tag{6}$$

By adding $\chi_H\cdot f=\chi_H\cdot f_1+\cdots+\chi_H\cdot f_k$ to (6) we get an almost everywhere integer valued $(a_1,\dots,a_k)$-periodic decomposition

$$f=(\chi_{\mathbb{R}\setminus H}\cdot g_1+\chi_H\cdot f_1)+\cdots+(\chi_{\mathbb{R}\setminus H}\cdot g_k+\chi_H\cdot f_k).$$

(i) $\Rightarrow$ (iv), (ii) $\Rightarrow$ (iv). We prove that if (iv) is false, then (i) and (ii) are also false. Suppose that $\frac{1}{b_1},\dots,\frac{1}{b_n}$ are not linearly independent over $\mathbb{Q}$. For each b_j choose an a_{i_j} such that b_j is a multiple of a_{i_j}. Then $\frac{1}{a_{i_1}},\dots,\frac{1}{a_{i_n}}$ are also linearly dependent over $\mathbb{Q}$, so there exists $(m_1,\dots,m_n)\in\mathbb{Z}\times\cdots\times\mathbb{Z}\setminus\{0,\dots,0\}$ such that

$$\frac{m_1}{a_{i_1}}+\cdots+\frac{m_n}{a_{i_n}}=0. \tag{7}$$

Let

$$f_{i_j}(x)=\left\{\frac{m_j}{a_{i_j}}x\right\}\quad (j=1,\dots,n),$$
$$f_i(x)=0\quad \bigl(i\in\{1,\dots,k\}\setminus\{i_1,\dots,i_n\}\bigr)$$

and $f=f_1+\cdots+f_k$.

Then, using (7), we get

$$f(x)=\sum_{i=1}^k f_i=\sum_{j=1}^n f_{i_j}=\sum_{j=1}^n\frac{m_j}{a_{i_j}}x-\sum_{j=1}^n\left[\frac{m_j}{a_{i_j}}x\right]=-\sum_{j=1}^n\left[\frac{m_j}{a_{i_j}}x\right]\in\mathbb{Z}.$$

Clearly each f_i is a bounded measurable a_i-periodic $\mathbb{R}\to\mathbb{R}$ function, so the conditions of (i) and (ii) are satisfied.

Suppose that f can be also written as $f=g_1+\cdots+g_k$ such that each g_i is an a_i-periodic measurable almost everywhere integer valued function. For each $j=1,\dots,n$ let h_j be the sum of those g_i's for which b_j is a multiple of a_i. Then $f=h_1+\cdots+h_n$ and each h_j is an almost everywhere integer valued measurable b_j-periodic function. On the other hand, $f=f_{i_1}+\cdots+f_{i_n}$ and each f_{i_j} is a measurable b_j-periodic function. Since $b_j/b_{j'}\notin\mathbb{Q}$ for any $j\neq j'$, Lemma 1.1 implies that $f_{i_j}-h_j$ is constant almost everywhere. Since $f_{i_j}=\{\frac{m_j}{a_{i_j}}x\}$, h_j is almost everywhere integer valued and at least one of $m_1,\dots,m_n$ is not zero, this is a contradiction.

(ii) $\Leftrightarrow$ (iii). As we already mentioned in the introduction, it is proved in [11] that the class of bounded measurable functions has the decomposition property; that is, a bounded measurable function $f:\mathbb{R}\to\mathbb{R}$ can be decomposed as $f=f_1+\cdots+f_k$ such that each f_j is an a_j-periodic real valued bounded measurable function if and only if $\Delta_{a_1}\cdots\Delta_{a_k}f=0$. On the other hand, by (ii), for integer valued functions the existence of a real valued bounded measurable $(a_1,\dots,a_k)$-periodic decomposition is equivalent with the existence of an almost everywhere integer valued bounded measurable $(a_1,\dots,a_k)$-periodic decomposition.

(iv) $\Rightarrow$ (i), (iv) $\Rightarrow$ (ii). First consider the case when $a_i/a_j\in\mathbb{Q}$ for every i,j. Then we can clearly assume that $a_j\in\mathbb{Z}$ for every j.

For $t\in[0,1)$ and $n\in\mathbb{Z}$ let $F_t(n)=f(n+t)$ and $F_{j,t}(n)=f_j(n+t)$ $(j=1,\dots,k)$. Then, applying Theorem 0.1 for each $t\in[0,1)$ for the decomposition $F_t=F_{1,t}+\cdots+F_{k,t}$ we get functions $G_{j,t}:\mathbb{Z}\to\mathbb{Z}$ such that $F_t=G_{1,t}+\cdots+G_{k,t}$ and each $G_{j,t}$ is a_j-periodic. Letting $g_j(n+t)=G_{j,t}(n)$ for each $j=1,\dots,k$, $n\in\mathbb{Z}$, $t\in[0,1)$ we get that $g_1,\dots,g_k$ have all the desired properties except measurability and boundedness.

Let N be the smallest common multiple of $a_1,\dots,a_k$. For every $n_0,n_1,\dots,n_{N-1}\in\mathbb{Z}$ let

$$A_{n_0,\dots,n_{N-1}} = \{x \in [0,1): f(x) = n_0,\ f(x+1) = n_1, \dots, f(x+N-1) = n_{N-1}\}.$$

Note that if $t, t' \in A_{n_0,\dots,n_{N-1}}$ for some $n_0, n_1, \dots, n_{N-1} \in \mathbb{Z}$, then $F_t = F_{t'}$, so we may guarantee $G_{j,t} = G_{j,t'}$ for each j in this case. Since every $A_{n_0,\dots,n_{N-1}}$ is measurable this guarantees that $g_1, \dots, g_k$ are measurable.

If $f_1, \dots, f_k$ are bounded, then f is also bounded, so $A_{n_0,\dots,n_{N-1}}$ is nonempty only for finitely many sequences $n_0, n_1, \dots, n_{N-1} \in \mathbb{Z}$. Since each $G_{j,t}$ is clearly bounded the previous paragraph guarantees that $g_1, \dots, g_k$ are also bounded.

Finally we prove (iv) $\Rightarrow$ (i) and (iv) $\Rightarrow$ (ii) in the general case. For each equivalence class E_i ($i = 1, \dots, n$) of $\{a_1, \dots, a_k\}$ with respect to the relation $a \sim b \Leftrightarrow a/b \in \mathbb{Q}$ let h_i be the sum of those f_j's for which $a_j \in E_i$. This way we get a decomposition $f = h_1 + \dots + h_n$ such that each h_i is a b_i-periodic measurable $\mathbb{R} \to \mathbb{R}$ function and if every f_j is bounded, then so is every h_i.

Since by (iv), $1/b_1, \dots, 1/b_n$ are linearly independent over $\mathbb{Q}$, Theorem 2.3 implies that every $\{h_i\}$ is constant almost everywhere. By adding constants to some of the functions $f_1, \dots, f_k$ we can guarantee that each $\{h_i\} = 0$ almost everywhere, which means that we can suppose that each h_i is almost everywhere integer valued. Since $h_i = \sum_{a_j \in E_i} f_j$ and $a_j/a_{j'} \in \mathbb{Q}$ if $a_j, a_{j'} \in E_i$, the first considered case can be applied for each h_i. $\square$

3. Integer valued decompositions

It is natural to ask whether we can get (everywhere) integer valued decompositions in (i)–(iii) of Theorem 2.5 or not. We will see that this depends on the answers to some questions about the general (nonmeasurable) case.

Question 3.1. *Can one add the following statement to the list of equivalent statements of Theorem* 2.5?

(i$''$) *If an integer valued function $f : \mathbb{R} \to \mathbb{Z}$ can be decomposed as $f = f_1 + \dots + f_k$ such that each f_j is an a_j-periodic measurable $\mathbb{R} \to \mathbb{R}$ function, then f can be also decomposed as $f = g_1 + \dots + g_k$ such that each g_j is an a_j-periodic integer valued measurable function.*

We shall prove that this question is actually equivalent with Question 0.2. For proving the equivalence of these questions we need the following lemma, which might be known.

Lemma 3.2. *For every $l = 1, 2, \dots$ there exists an additive subgroup A_l of $\mathbb{R}$ such that*

(a) *A_l is isomorphic to $\mathbb{Z}^l = \mathbb{Z} \times \dots \times \mathbb{Z}$, and*
(b) *whenever $k \in \mathbb{N}$, $t_1, \dots, t_k \in A_l \setminus \{0\}$ and $t_i/t_j \notin \mathbb{Q}$ for every $i \neq j$, then $\frac{1}{t_1}, \dots, \frac{1}{t_k}$ are linearly independent over $\mathbb{Q}$.*

Proof. We prove by induction. For $l = 1$ we can choose $A_1 = \mathbb{Z}$.

Now we construct A_{l+1} from A_l. For fixed $k \in \mathbb{N}$, $a = (a_1, \dots, a_k) \in A_l^k$, $m = (m_1, \dots, m_k) \in \mathbb{Z}^k$ and $n = (n_1, \dots, n_k) \in (\mathbb{Z} \setminus \{0\})^k$ let

$$\Phi_{a,m,n}(x) = \frac{n_1}{a_1 + m_1 x} + \dots + \frac{n_k}{a_k + m_k x}.$$

For some polynomial $P_{a,m,n}(x)$, the function $\Phi_{a,m,n}(x)$ can be also written as

$$\Phi_{a,m,n}(x) = \frac{P_{a,m,n}(x)}{(a_1 + m_1 x) \cdot \dots \cdot (a_k + m_k x)}.$$

Choose $y \in \mathbb{R}$ such that $y \notin A_l$ and y is not the root of any of those polynomials $P_{a,m,n}(x)$ that are not identically zero. This is possible since A_l is countable and we have only countably many polynomials and each has finitely many roots.

We claim that $A_{l+1} = A_l + \mathbb{Z}y$ has the required properties. Since $y \notin A_l$, A_{l+1} is indeed isomorphic to $\mathbb{Z}^{l+1}$. Suppose that $t_1, \dots, t_k \in A_{l+1} \setminus \{0\}$ and $t_i/t_j \notin \mathbb{Q}$ for every $i \neq j$ but $\frac{1}{t_1}, \dots, \frac{1}{t_k}$ are not linearly independent over $\mathbb{Q}$. Then there exist $a = (a_1, \dots, a_k) \in A_l^k$ and $m = (m_1, \dots, m_k) \in \mathbb{Z}^k$ such that $t_i = a_i + m_i y$ for each i and there exists

$n=(n_1,\dots,n_k)\in\mathbb{Z}^k\setminus\{0,\dots,0\}$ such that $\frac{n_1}{t_1}+\dots+\frac{n_k}{t_k}=0$. Then $\Phi_{a,m,n}(y)=0$, so $P_{a,m,n}(y)=0$. We can suppose that $n=(n_1,\dots,n_k)\in(\mathbb{Z}\setminus\{0\})^k$ since we may simply omit every superfluous t_j. By the definition of y, this implies that $P_{a,m,n}$ is identically zero, so $\Phi_{a,m,n}(x)=0$ for every x such that $a_i+m_ix\neq 0$ for every i.

We cannot have $m_1=\dots=m_k=0$ since then we would have $t_1,\dots,t_k\in A_l$, which would mean that A_l does not satisfy property (b). Let i be such that $m_i\neq 0$. Then on one hand

$$\lim_{x\to-\frac{a_i}{m_i}}\left|\frac{n_i}{a_i+m_ix}\right|=\infty,$$

on the other hand, $\lim_{x\to-\frac{a_i}{m_i}}\Phi_{a,m,n}=0$, so there must be an other term $\frac{n_j}{a_j+m_jx}$ of $\Phi_{a,m,n}$ with similar property. This means that there exists $j\neq i$ such that $a_j/m_j=a_i/m_i$. But this implies that

$$\frac{t_i}{t_j}=\frac{a_i+m_ix}{a_j+m_jx}=\frac{m_i}{m_j}\in\mathbb{Q},$$

which is a contradiction. □

Proposition 3.3. *The answers to Questions* 3.1 *and* 0.2 *must be the same.*

Added in proof: Since in [4] a positive answer was given to Question 0.2, it turned out that the answer to Question 3.1 is also affirmative.

Proof. First suppose that the answer is affirmative to Question 3.1 and suppose that an integer valued function $f:\mathbb{R}\to\mathbb{Z}$ can be written as $f=g_1+\dots+g_k$, where each g_j is a real valued a_j-periodic function for some $a_j\in\mathbb{R}$. We want to show that f can be also written as $f=h_1+\dots+h_k$, where each h_j is an integer valued a_j-periodic function. It is enough to find such h_j's on the additive group A generated by $a_1,\dots,a_k$ since then we can define every h_j the same way on every coset of A. Then we can also suppose that $f,g_1,\dots,g_k$ are defined also only on A, which is isomorphic to a group of the form $\mathbb{Z}^l$ for some l.

By Lemma 3.2 there exists an additive subgroup A_l of $\mathbb{R}$ that is isomorphic to $\mathbb{Z}^l$ (and so to A as well) and satisfies (b) of Lemma 3.2. Hence we may assume that $f,g_1,\dots,g_k$ are defined on A_l. For every $x\in\mathbb{R}$ let

$$F(x)=\begin{cases} f(x) & \text{if } x\in A_l,\\ 0 & \text{if } x\notin A_l,\end{cases}$$

and for each $j=1,\dots,k$,

$$G_j(x)=\begin{cases} g_j(x) & \text{if } x\in A_l,\\ 0 & \text{if } x\notin A_l.\end{cases}$$

Then for each j the function G_j is a_j-periodic and also measurable since A_j has measure 0. Since $a_1,\dots,a_k\in A_l$ and A_l satisfies (b) of Lemma 3.2, (iv) of Theorem 2.5 holds for $a_1,\dots,a_k$. On the other hand, affirmative answer to Question 3.1 means that (iv) of Theorem 2.5 implies (i″) of Question 3.1. Therefore we may apply (i″) of Question 3.1 for $F=G_1+\dots+G_k$ to get a decomposition $F=H_1+\dots+H_k$ such that each $H_j:\mathbb{R}\to\mathbb{Z}$ is a_j-periodic. Then the restriction h_j of H_j to A_l ($j=1,\dots,k$) gives a suitable decomposition of f on A.

Now we suppose that the answer to Question 0.2 is affirmative and we prove that then (i″) of Question 3.1 is equivalent to (i), (i′), (ii), (ii′), (iii), (iii′) and (iv) of Theorem 2.5. Since in the (i) ⇒ (iv) proof of Theorem 2.5 integer valued f is constructed that the same argument also proves (i″) ⇒ (iv). So it is enough to prove (i) ⇒ (i″).

Suppose that (i) holds, $f:\mathbb{R}\to\mathbb{Z}$, $f=f_1+\dots+f_k$ and each f_j is a measurable a_j-periodic function. By (i), there exists a decomposition $f=g_1+\dots+g_k$ such that each g_j is a measurable a_j-periodic almost everywhere integer valued function. All we have to do is replacing g_j's by integer valued measurable functions.

Let

$$E_j=\{x:\ g_j(x)\notin\mathbb{Z}\}\quad\text{and}\quad E=\left(\bigcup_{j=1}^k E_j\right)+a_1\mathbb{Z}+\dots+a_k\mathbb{Z}.$$

Then E is a set of measure zero and it is a_j-periodic for every j. Thus for each j the function $g_j\chi_{\mathbb{R}\setminus E}$ is a_j-periodic. Then, by the assumption that the answer to Question 0.2 is affirmative, $f\chi_E = g_1\chi_E + \cdots + g_k\chi_E$ implies that there exists a decomposition

$$f\chi_E = F_1 + \cdots + F_k$$

such that each F_j is an integer valued a_j-periodic function.

For each j let

$$G_j(x) = g_j\chi_{\mathbb{R}\setminus E} + F_j\chi_E.$$

Then G_j is clearly a_j-periodic and integer valued. It is also measurable since g_j is measurable and E is of measure zero. Since

$$f(x) = G_1(x) + \cdots + G_k(x)$$

clearly holds both for $x \in E$ and $x \in \mathbb{R} \setminus E$, we obtained a decomposition we wanted. □

Now we prove that we cannot guarantee everywhere integer valued decompositions in (ii) and (iii) of Theorem 2.5.

Proposition 3.4. *There exists $a_1, a_2, a_3 \in \mathbb{R}$ such that $\frac{1}{a_1}$, $\frac{1}{a_2}$ and $\frac{1}{a_3}$ are linearly independent over $\mathbb{Q}$ and a function $f : \mathbb{R} \to \{0, 1\}$ that has a bounded measurable real valued (a_1, a_2, a_3)-periodic decomposition but does not have a bounded measurable integer valued (a_1, a_2, a_3)-periodic decomposition.*

Consequently one cannot replace "almost everywhere integer valued" by "integer valued" in (ii) *and* (iii) *of Theorem* 2.5.

Proof. Choose a_1, a_2 and a_3 so that $a_1 + a_2 = a_3$ but $\frac{1}{a_1}$, $\frac{1}{a_2}$ and $\frac{1}{a_3}$ are linearly independent over $\mathbb{Q}$, which is possible for example by taking $a_1, a_2 \in A_2$ such that $a_1/a_2 \notin \mathbb{Q}$, where A_2 is the additive subgroup of $\mathbb{R}$ obtained by Lemma 3.2.

By Theorem 0.3 there exists a function $u : \mathbb{Z} \times \mathbb{Z} \to \{0, 1\}$ that has a decomposition $u = u_1 + u_2 + u_3$ such that u_1, u_2 and u_3 are bounded real valued periodic functions with periods $(1, 0)$, $(0, 1)$ and $(1, 1)$, respectively, but u has no decomposition $u = v_1 + v_2 + v_3$ such that v_1, v_2 and v_3 are bounded integer valued periodic functions with the same periods.

Let $E = a_1\mathbb{Z} + a_2\mathbb{Z}$,

$$f(x) = \begin{cases} u(n, m) & \text{if } x = na_1 + ma_2 \ (n, m \in \mathbb{Z}), \\ 0 & \text{if } x \notin E, \end{cases}$$

and for each $j = 1, 2, 3$,

$$f_j(x) = \begin{cases} u_j(n, m) & \text{if } x = na_1 + ma_2 \ (n, m \in \mathbb{Z}), \\ 0 & \text{if } x \notin E. \end{cases}$$

Then clearly f maps to $\{0, 1\}$, $f = f_1 + f_2 + f_3$ and each f_j is a_j-periodic, bounded and measurable (since almost everywhere zero).

But f cannot have a decomposition $f = g_1 + g_2 + g_3$ such that each g_j is a_j-periodic, bounded and integer valued since then $v_j(n, m) = g_j(na_1 + ma_2)$ $(j = 1, 2, 3)$ would give an integer valued bounded decomposition of u with periods a_1, a_2, a_3, which is impossible. □

Finally we pose two problems.

Problem 3.5. Characterize those periods $a_1, \ldots, a_k \in \mathbb{R} \setminus \{0\}$ for which the existence of a bounded measurable real valued $(a_1, \ldots, a_k)$-periodic decomposition of an integer valued function implies the existence of a bounded measurable integer valued $(a_1, \ldots, a_k)$-periodic decomposition.

Theorem 2.5 implies that (iv) of Theorem 2.5 is a necessary condition but Proposition 3.4 shows that it is not sufficient. The proofs of Propositions 3.3 and 3.4 indicate that this problem must be also related to the analogue nonmeasurable problem, which seems to be also open.

Problem 3.6. Characterize those periods $a_1, \ldots, a_k \in \mathbb{R} \setminus \{0\}$ for which the existence of a bounded real valued $(a_1, \ldots, a_k)$-periodic decomposition of an integer valued function implies the existence of a bounded integer valued $(a_1, \ldots, a_k)$-periodic decomposition.

As we mentioned after Theorem 0.3, it is proved in [7] that some restriction on the periods is necessary.

References

[1] B. Farkas, Sz.Gy. Révész, Decomposition as the sum of invariant functions with respect to commuting transformations, Aequationes Math., in press.

[2] Z. Gajda, Note on decomposition of bounded functions into the sum of periodic terms, Acta Math. Hungar. 59 (1–2) (1992) 103–106.

[3] G.H. Hardy, E.M. Wright, An Introduction to the Theory of Numbers, Oxford, 1960.

[4] B. Farkas, V. Harangi, T. Keleti, Sz.Gy. Révész, Invariant decomposition of functions with respect to commuting invertible transformations, Proc. Amer. Math. Soc., submitted for publication.

[5] V.M. Kadets, S.B. Shumyatskiy, Averaging technique in the periodic decomposition problem, Mat. Fiz. Anal. Geom. 7 (2) (2000) 184–195.

[6] V.M. Kadets, S.B. Shumyatskiy, Additions to the periodic decomposition theorem, Acta Math. Hungar. 90 (4) (2001) 293–305.

[7] Gy. Károlyi, G. Kós, T. Keleti, I.Z. Ruzsa, Periodic decomposition of integer valued functions, Acta Math. Hungar., in press.

[8] T. Keleti, On the differences and sums of periodic measurable functions, Acta Math. Hungar. 75 (4) (1997) 279–286.

[9] T. Keleti, Difference functions of periodic measurable functions, Fund. Math. 157 (1998) 15–32.

[10] M. Laczkovich, Sz.Gy. Révész, Periodic decompositions of continuous functions, Acta Math. Hungar. 54 (3–4) (1989) 329–341.

[11] M. Laczkovich, Sz.Gy. Révész, Decompositions into periodic functions belonging to a given Banach space, Acta Math. Hungar. 55 (3–4) (1990) 353–363.

[12] S. Mortola, R. Peirone, The sum of periodic functions, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) 2 (1999) 393–396.

[13] T. Natkaniec, W. Wilczyński, Sums of periodic Darboux functions and measurability, Atti Semin. Mat. Fis. Univ. Modena 51 (2003) 369–376.

[14] M. Wierdl, Continuous functions that can be represented as the sum of finitely many periodic functions, Mat. Lapok 32 (1984) 107–113 (in Hungarian).