

MÖBIUS NUMBER SYSTEMS

Alexandr Kazda¹

Abstract. Möbius number systems represent the extended real line or, equivalently, the unit complex circle by sequences of Möbius transformations. A Möbius number system consists of an iterative system of Möbius transformations and a subshift.

In this paper we give an overview of the area of Möbius number systems. We are particularly interested in the conditions, under which a Möbius number system does or does not exist. We give an overview of known sufficient and necessary conditions on the iterative system and then introduce a necessary condition for the subshift.

As Möbius number systems use subshifts instead of the whole symbolic space, we can ask what is the language complexity of these subshifts. We present a more user-friendly version of an already known condition for a number system to be sofic.

AMS Mathematics Subject Classification (2000): 40A05, 37B10, 51M09

Key words and phrases: Möbius transformation, numeral system, subshift, convergence

1. Introduction

Numeral systems are recipes for expressing numbers in symbols. The most common are positional systems (usually with base ten, two, eight or sixteen). However, these systems are by no means the only possibility. Various historical systems used different approaches (consider for example the Roman numerals).

Modern numeration theorists typically study positional systems with real base (such as the golden mean or -2) or various modifications of continued fractions. It turns out that numeration theory has connections to various other fields, namely the study of fractals and tilings, symbolic dynamics, ergodic theory, computability theory and even cryptography.

In this paper, we study Möbius number systems as introduced in [8]. A Möbius number system represents numbers as sequences of Möbius transformations obtained by composing a finite starting set of Möbius transformations.

Möbius number systems display complicated dynamical properties and have connections to other kinds of numeral systems. In particular, Möbius number systems can generalize continued fractions as shown in [10].

¹Department of Algebra, Charles University, Prague, Czech Republic,
e-mail: alexak@atrey.karlin.mff.cuni.cz

The first half of the paper consists of a survey of already known theory (particularly from [4] and [7]). While in the other half we branch our study in several directions: We present the results of a computer experiment suggesting that known conditions on Möbius iterative systems are quite close to being sufficient and necessary, then prove that a nontrivial class of subshifts does not admit Möbius number systems and finally formulate a sufficient condition for a number system to be sofic.

2. Preliminaries

2.1. Metric spaces and words

Denote by \mathbb{T} the unit circle and by \mathbb{D} the closed unit disc in the complex plane.

Let A be a finite alphabet. Any sequence of elements of A is a *word* over A . Let λ be the empty word. Denote by A^* the monoid of all finite words over A , by A^+ the set $A^* \setminus \{\lambda\}$ and by A^ω the set of all one-sided infinite words over A . Let $|w|$ denote the length of the word w . If n is finite, let A^n be the set of all words over A of length precisely n . We use the notation $w = w_0w_1w_2\cdots$ and $w_{[i,j]} = w_iw_{i+1}\cdots w_j$. When u is a finite word and v any word we can define the *concatenation* of u and v as $uv = u_0u_1\cdots u_{|u|-1}v_0v_1\cdots$.

Let $v \in A^*$ be a word of length n . Then we write

$$[v] = \{w \in A^\omega : w_{[0,n-1]} = v\}$$

and call the resulting subset of A^ω the *cylinder* of v . A word u is a *factor* of a word v if there exist i, j such that $u = v_{[i,j]}$.

Let X be a metric space. We denote by ρ the metric function of X , by $\text{Int}(V)$ the interior of the set V and by $B_r(x)$ the open ball of radius r centered at x . If I is an interval, denote by $|I|$ the length of I .

We equip \mathbb{C} with the metric $\rho(x, y) = |x - y|$ and \mathbb{T} with the circle distance metric (i.e. metric measuring distances along the circle). The shift space A^ω of one-sided infinite words comes equipped with the metric $\rho(u, v) = \max(\{2^{-k} : u_k \neq v_k\} \cup \{0\})$. A *subshift* $\Sigma \subseteq A^\omega$ is a set that is both topologically closed and invariant under the *shift map* $\sigma(w)_i = w_{i+1}$ (i.e. $\sigma(\Sigma) \subseteq \Sigma$).

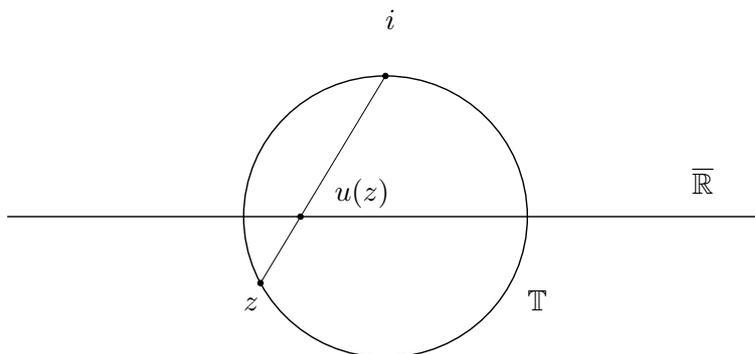
As shown in [6, pages 5 and 179], subshifts of A^ω are precisely the subsets of A^ω that can be defined by some set of forbidden factors. More precisely, Σ is a subshift iff there exists $F \subseteq A^+$ such that

$$\Sigma = \{w \in A^\omega : \forall v \in F, v \text{ is not a factor of } w\}.$$

We are going to occasionally define shifts using some such set of forbidden factors.

The *language* of a subshift $\mathcal{L}(\Sigma)$ is the set of all the words $v \in A^*$ for which there exists $w \in \Sigma$ such that v is a factor of w . See [6] for a more detailed treatment of this topic.

We will mainly consider symbolic representations of \mathbb{T} , although representations of the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ will make an appearance as well.

Figure 1: The stereographic projection of \mathbb{T} onto $\overline{\mathbb{R}}$

Note that $\overline{\mathbb{R}}$ is homeomorphic to \mathbb{T} via the stereographic projection (see Fig. 1).

$$u : \mathbb{T} \rightarrow \overline{\mathbb{R}}, \quad u : z \mapsto \frac{-iz + 1}{z - i}.$$

Therefore, as long as we are not interested in arithmetics, representing \mathbb{T} is equivalent to representing the extended real line.

2.2. Möbius transformations

A Möbius transformation (MT for short) of the complex sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ is any map of the form

$$F : z \mapsto \frac{az + b}{cz + d}$$

where $(a, b), (c, d)$ are linearly independent vectors from \mathbb{C}^2 .

Note that the stereographic projection u as defined above is actually a Möbius transformation. Therefore, if we represent \mathbb{T} using the system $\{F_a : a \in A\}$ of MTs, we can represent $\overline{\mathbb{R}}$ in the same way with the system $\{u \circ F_a \circ u^{-1} : a \in A\}$.

To every regular 2×2 complex matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we can associate the MT defined by $F_A(z) = \frac{az+b}{cz+d}$. While the map $A \mapsto F_A$ is surjective, every MT has many preimages: if A is a matrix for F then so is cA for any $c \in \mathbb{C}, c \neq 0$. Even normalizing the matrices by demanding $\det A = 1$ is not enough, as it leaves two preimages A and $-A$ for each F .

This ambiguity is, however, a small price to pay: An easy calculation shows that composition of MTs corresponds to multiplying their respective matrices: $F_A \circ F_B = F_{A \cdot B}$. This is why we will often think of MTs as of matrices. It follows that the set of all MTs together with the operation of composition is a group (isomorphic to $SL(2, \mathbb{C})/\{E, -E\}$). In particular, MTs are bijective on $\mathbb{C} \cup \{\infty\}$.

Unless noted otherwise, we will assume all Möbius transformations to be *disc preserving*, i.e. demand that they map \mathbb{D} onto itself. Obviously, disc preserving

transformations form a subgroup of the group of all MTs. It turns out that F is disc preserving iff it has the form

$$F = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}$$

with the normalizing condition $|\alpha|^2 - |\beta|^2 = 1$. We provide a detailed proof in the Appendix as Lemma 21.

The geometrical theory of MTs is quite rich and has a strong link to hyperbolic geometry (see [3]). In this paper, we will need only a handful of basic fragments of this theory. We will use the fact that MTs take circles and lines to circles and lines (possibly turning a circle into a line or vice versa) and the observation that disc preserving transformations also preserve orientation of intervals on the circle (clockwise versus counterclockwise), so the image of the interval $[x, y]$ is the interval $[F(x), F(y)]$ (as compared to $[F(y), F(x)]$).

We can establish a taxonomy of disc preserving MTs by considering the trace of the normalized matrix representing F . While $\text{Tr } F$ does not have a well defined sign, the number $(\text{Tr } F)^2$ is unique and real for each disc preserving F .

Definition 1. Let $F \neq \text{id}$ be a disc preserving MT. We call F :

1. *elliptic* if $(\text{Tr } F)^2 < 4$,
2. *parabolic* if $(\text{Tr } F)^2 = 4$,
3. *hyperbolic* if $(\text{Tr } F)^2 > 4$.

To better understand this classification, consider the fixed points of F . We claim (Lemma 22 in the Appendix) that:

1. F is elliptic iff it has one fixed point inside and one fixed point outside of \mathbb{T} (the outside point might be ∞),
2. F is parabolic iff it has a single fixed point which lies on \mathbb{T} ,
3. F is hyperbolic iff it has two distinct fixed points, both lying on \mathbb{T} .

Remark. Let F be a hyperbolic transformation with fixed points x_1, x_2 . Then one of these points (say, x_1) is *stable* and the other is *unstable*. It is $F'(x_1) < 1 < F'(x_2)$ and for every $z \in \overline{\mathbb{C}}, z \neq x_2$, we have $\lim_{n \rightarrow \infty} F^n(z) = x_1$.

Similarly, when F is parabolic with the fixed point x , we have $F^n(z) \rightarrow x$ for all $z \in \overline{\mathbb{C}}$ and $F'(x) = 1$. See Lemmas 23 and 24 in the Appendix for proofs of these facts.

We will show the significance of this classification in the following sections.

2.3. Number representation

Möbius number systems assign numbers to sequences of mappings. This principle is actually less exotic than it appears to be. Consider the usual binary representation of the interval $[0, 1]$. Let $A = \{0, 1\}$ be our alphabet. We want to assign to each word $w \in A^\omega$ the number $\Phi(w) = 0.w$ and so obtain the map $\Phi : A^\omega \rightarrow [0, 1]$. We need to use some sort of limit process: Taking longer and longer prefixes of w , we obtain better and better approximations, ending with the unique number $0.w$.

The usual construction of the binary system involves letting $\Phi(w)$ to be equal to the limit of the sequence $\{0.w_{[0,k]}\}_{k=1}^\infty$. However, we can also define binary numbers in the language of mappings.

Consider the two maps

$$\begin{aligned} F_0 : x &\mapsto x/2 \\ F_1 : x &\mapsto (x+1)/2. \end{aligned}$$

For $v \in A^n$ let $F_v = F_{v_0} \circ F_{v_1} \circ \cdots \circ F_{v_{n-1}}$. Both maps F_0, F_1 are continuous and, more importantly, contractions on the interval $[0, 1]$: For each $x, y \in [0, 1]$ and each $i = 0, 1$ we have $|F_i(x) - F_i(y)| = \frac{1}{2}|F_i(x) - F_i(y)|$. Therefore, for any $w \in A^\omega$, the set $\bigcap_{k=1}^\infty F_{w_{[0,k]}}[0, 1]$ is a singleton. What is more, a proof by induction reveals that $F_{w_{[0,k]}}[0, 1]$ is actually precisely the set of all the real numbers whose binary expansion begins with $0.w_{[0,k]}$. We have obtained that $\bigcap_{k=1}^\infty F_{w_{[0,k]}}[0, 1] = \{\Phi(w)\} = \{0.w\}$. If we wished, we could go on to prove that Φ is continuous and surjective, both very desirable properties for a number system.

We would like to do the same for Möbius transformations in place of F_0, F_1 and call the result a Möbius number system. However, as MTs are bijective on the complex sphere, we cannot use the contraction property like we did above. To fix this, [8] defined Φ using convergence of measures. We will see that there are other (equivalent) definitions in Theorem 4 but let us give the original definition first.

Denote $m(\mathbb{T})$ the set of all Borel probability measures on \mathbb{T} . If ν is a Borel measure on \mathbb{T} and $F : \mathbb{T} \rightarrow \mathbb{T}$ an MT, we define the measure $F\nu$ by $F\nu(E) = \nu(F^{-1}(E))$ for all measurable sets E on \mathbb{T} . The Dirac measure centered at point x is the measure δ_x such that

$$\delta_x(E) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise} \end{cases}$$

for any E measurable subset of \mathbb{T} . It is a quite straightforward idea to identify δ_x with the point x itself.

Before we define what does it mean for a sequence of MTs to represent a point, let us give some brief background. Denote by $C(\mathbb{T}, \mathbb{R})$ the vector space of all continuous functions from \mathbb{T} to \mathbb{R} (with the supremum norm). Finite Borel measures act on $C(\mathbb{T}, \mathbb{R})$ as continuous linear functionals: Measure ν assigns

to $f \in C(\mathbb{T}, \mathbb{R})$ the number $\int f d\nu$ and if $\nu \neq \nu'$ then the two measures define different functionals by the Riesz representation theorem (see [1, page 184]).

We have the embedding $m(\mathbb{T}) \subseteq C(\mathbb{T}, \mathbb{R})^*$, where $C(\mathbb{T}, \mathbb{R})^*$ is the dual space to $C(\mathbb{T}, \mathbb{R})$. There are three usual topologies on $C(\mathbb{T}, \mathbb{R})^*$ (listed in the order of strength): The norm topology, the weak topology and the weak* topology.

Definition 2. Denote by μ the uniform probability measure on \mathbb{T} . Let $\{F_n\}_{n=1}^\infty$ be a sequence of Möbius transformations. We say that the sequence $\{F_n\}_{n=1}^\infty$ represents the point $x \in \mathbb{T}$ if and only if $\lim_{n \rightarrow \infty} F_n \mu = \delta_x$. Here μ is the uniform probability measure on \mathbb{T} and the convergence of measures is taken in the weak* topology, i.e. $\nu_n \rightarrow \nu$ if and only if for all $f : \mathbb{T} \rightarrow \mathbb{R}$ continuous we have $\int f d\nu_n \rightarrow \int f d\nu$.

Remark. The reader might wonder why did we choose weak* topology here instead of any of the two other common topologies.

One answer is that this is the usual way to define convergence of measures in fields such as ergodic theory. Another answer is that even weak topology is too strong to provide any representation of points at all: Consider any sequence $\{F_n\}_{n=1}^\infty$ of MTs. To obtain $\lim_{n \rightarrow \infty} F_n \mu = \delta_x$ in the weak topology, we would have to satisfy $\alpha(F_n \mu) = \alpha(\delta_x)$ for any continuous linear functional $\alpha \in C(\mathbb{T}, \mathbb{R})^{**}$.

By the Riesz representation theorem, the space $C(\mathbb{T}, \mathbb{R})^*$ can be identified with the space of all Radon signed measures on \mathbb{T} . For λ Radon signed measure on \mathbb{T} , define $\alpha(\lambda) = \lambda(\{x\})$. This is a continuous linear functional on $C(\mathbb{T}, \mathbb{R})^*$ (see Lemma 25 in the Appendix). Obviously, $\alpha(\delta_x) = 1$, while $\alpha(F_n \mu) = \mu(\{F_n^{-1}(x)\}) = 0$, so the sequence $\{F_n\}_{n=1}^\infty$ does not represent x . As this is true for all sequences and all values of x , Definition 2 would be meaningless in the weak topology and the same is true in the norm topology (which is even stronger than the weak topology).

3. Representing \mathbb{T} and $\overline{\mathbb{R}}$

As in the rest of this paper, we will assume all MTs to be disc preserving. If we wish to represent $\overline{\mathbb{R}}$, we can use MTs which preserve the upper half plane and obtain similar results.

Let $\alpha \in [0, 2\pi)$, $r \in [1, \infty)$. Call the transformation $R_\alpha(z) = e^{i\alpha}z$ a *rotation*. It turns out that rotations are precisely those disc preserving MTs whose matrices are diagonal.

Denote by $F^\bullet(x)$ the modulus of the derivative of F at x . Direct calculation gives us that when $F = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$, $\det F = 1$ then

$$F^\bullet(x) = |F'(x)| = \frac{1}{|\beta x + \alpha|^2}.$$

This number measures whether and how much F expands or contracts the neighborhood of x .

Definition 3. Let F be a Möbius transformation. Then, inspired by [8] and [3], we define the *expansion interval of F^{-1}* as $V = \{x \in \mathbb{T} : (F^{-1})^\bullet(x) > 1\}$ and *expansion set of F^{-1}* by $D = \{x \in \mathbb{C} : (F^{-1})^\bullet(x) \geq 1\}$.

A straightforward calculation gives us that if F is not a rotation then V is an interval on \mathbb{T} and D is a disc in the complex plane.

Recall that a sequence of Möbius transformations $\{F_n\}_{n=1}^\infty$ represents the point $x \in \mathbb{T}$ if and only if $\lim_{n \rightarrow \infty} F_n \mu = \delta_x$ in the weak* topology. We will now list several equivalent definitions of what does it mean to represent a point. Our list is essentially Theorem 9 in [4].

Theorem 4. Let $\{F_n\}_{n=1}^\infty$ be a sequence of disc preserving MTs only finitely many of which are rotations. Denote by V_n the expansion interval of F_n^{-1} , by D_n the expansion set of F_n^{-1} and by d_n the center of D_n . Then the following statements are equivalent:

1. The sequence $\{F_n\}_{n=1}^\infty$ represents $x \in \mathbb{T}$.
2. For every open interval I on \mathbb{T} containing x we have $\lim_{n \rightarrow \infty} (F_n \mu)(I) = 1$.
3. There exists a number $c > 0$ such that for every open interval I on \mathbb{T} containing x it is true that $\liminf_{n \rightarrow \infty} (F_n \mu)(I) > c$.
4. $\lim_{n \rightarrow \infty} d_n = x$
5. $\lim_{n \rightarrow \infty} D_n = \{x\}$
6. $\lim_{n \rightarrow \infty} \bar{V}_n = \{x\}$
7. For all $K \subseteq \text{Int}(\mathbb{D})$ compact we have $\lim_{n \rightarrow \infty} F_n(K) = \{x\}$.
8. For all $z \in \text{Int}(\mathbb{D})$ we have $\lim_{n \rightarrow \infty} F_n(z) = x$.
9. There exists $z \in \text{Int}(\mathbb{D})$ such that $\lim_{n \rightarrow \infty} F_n(z) = x$.
10. The sequence $\{F_n\}_{n=1}^\infty$ converges to the constant map $c_x : z \mapsto x$ in measure, that is

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mu(\{z : \rho(F_n(z), x) > \varepsilon\}) = 0.$$

Here, μ is the uniform probability measure on \mathbb{T} and ρ the metric on \mathbb{T} . In (5), (6) and (7), we take convergence in the Hausdorff metric on the space of nonempty compact subsets of \mathbb{C} , \mathbb{T} and \mathbb{D} respectively. In particular, $E_n \rightarrow \{x\}$ if and only if for every $\varepsilon > 0$ there exists n_0 such that $\forall n > n_0$ it is $E_n \subseteq B_\varepsilon(x)$.

As an easy corollary of Theorem 4, we can prove that two intuitive ideas are true.

Corollary 5. *Let $\{F_n\}_{n=1}^\infty$ be a sequence of MTs representing the point x . Let M be a disc preserving MT. Then*

1. *The sequence $\{F_n \circ M\}_{n=1}^\infty$ represents x .*
2. *The sequence $\{M \circ F_n\}_{n=1}^\infty$ represents $M(x)$.*

Proof. In both cases, we use the fact that if $G_n(z) \rightarrow x$ for some $z \in \text{Int}(\mathbb{D})$ then the sequence $\{G_n\}_{n=1}^\infty$ represents x .

1. As $M(0)$ lies inside \mathbb{D} , we have $F_n(M(0)) \rightarrow x$, therefore $(F_n \circ M)(0) \rightarrow x$.
2. As $F_n(0) \rightarrow x$ and M is continuous, we have $M(F_n(0)) \rightarrow M(x)$. \square

We now have enough tools to show, like in [8], how do the three classes of MTs behave with respect to point representation:

1. Let F be an elliptic disc preserving transformation. Then the sequence $\{F^n\}_{n=1}^\infty$ does not represent any point.
2. Let F be a parabolic disc preserving transformation. Then the sequence $\{F^n\}_{n=1}^\infty$ represents the fixed point of F .
3. Let F be a hyperbolic disc preserving transformation. Then the sequence $\{F^n\}_{n=1}^\infty$ represents the stable fixed point of F .

For all three claims, we will need part (8) of Theorem 4.

To prove (1), recall that if F is elliptic, there exists a fixed point of F inside \mathbb{T} . Denote this point by x . Then for all n , $F^n(x) = x \notin \mathbb{T}$, so $\{F^n\}_{n=1}^\infty$ can not represent anything.

In the parabolic and hyperbolic case, denote by x the (stable) fixed point of F . We now use Lemma 24 in the Appendix to obtain that for all $z \in \text{Int}(\mathbb{D})$ we have $F^n(z) \rightarrow x$. Therefore, $\{F^n\}_{n=1}^\infty$ represents x , proving (2) and (3).

4. Möbius number systems

In order to introduce Möbius number system, we need some sort of connection between words in A^ω and sequences of MTs. The notion of Möbius iterative system allows us to draw such a connection.

Definition 6. Let A be an alphabet. Assume we assign to every $a \in A$ a Möbius transformation F_a . The set $\{F_a : a \in A\}$ is then called a *Möbius iterative system*.

Given an iterative system, we assign to each word $v \in A^n$ the mapping $F_v = F_{v_0} \circ F_{v_1} \circ \cdots \circ F_{v_{n-1}}$.

Definition 7. Given $w \in A^\omega$, we define $\Phi(w)$ as the point $x \in \mathbb{T}$ such that the sequence $\{F_{w_{[0,n]}}\}_{n=1}^\infty$ represents x . If $\{F_{w_{[0,n]}}\}_{n=1}^\infty$ does not represent any point in \mathbb{T} , let $\Phi(w)$ be undefined. Denote the domain of the resulting map Φ by \mathbb{X}_F .

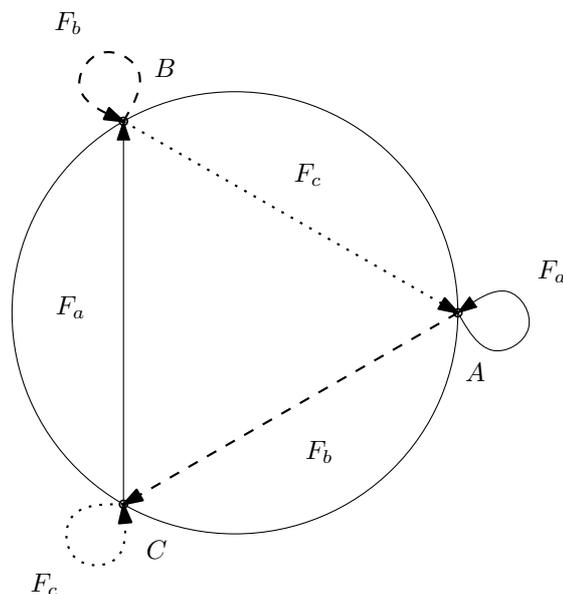


Figure 2: The three parabolic maps system

We are finally ready to say what a Möbius number system is:

Definition 8. The subshift $\Sigma \subseteq A^\omega$ is a *Möbius number system* for a given Möbius iterative system if $\Sigma \subseteq \mathbb{X}_F$, $\Phi(\Sigma) = \mathbb{T}$ and $\Phi|_\Sigma$ is continuous.

Using Corollary 5, we observe that if $\Phi(w) = x$ then $\Phi(\sigma(w)) = F_{w_0}^{-1}(x)$. We will use this simple property later.

We now give three examples of Möbius number systems, although the proof that they indeed are Möbius number systems will have to wait until Section 6 when we have suitable tools.

Example 9. Let A, B, C be three vertices of an equilateral triangle inscribed in \mathbb{T} . Take F_a, F_b, F_c the three parabolic transformations satisfying.

$$\begin{aligned} F_a(A) &= A, & F_a(C) &= B \\ F_b(B) &= B, & F_b(A) &= C \\ F_c(C) &= C, & F_c(B) &= A. \end{aligned}$$

See Fig. 2. A quick calculation reveals that F_a, F_b, F_c are in fact uniquely determined by the triangle ABC .

Let us define the shift Σ by the three forbidden factors ac, ba, cb . We claim that Σ is a Möbius number system for the iterative system $\{F_a, F_b, F_c\}$.

The following two examples are due to Petr Kůrka, see [10].

Words $00, 1\bar{1}$ and $\bar{1}1$ correspond to the identity maps while $\Phi((01)^\infty)$ and $\Phi((0\bar{1})^\infty)$ are not defined (as F_{01} and $F_{0\bar{1}}$ are parabolic). This is why we define the shift Σ by the set of forbidden words $00, 1\bar{1}, \bar{1}1, 101, \bar{1}0\bar{1}$.

It turns out that Σ is the regular continued fraction system as depicted in Fig. 3. In this picture, the labelled points represent the images of the point 0 under the corresponding sequence of transformations, while curves connect images of 0 that are next to each other in a given sequence. Observe that the images of 0 converge to the boundary of the disc, ensuring convergence.

A slight complication not present in the usual continued fraction system is that we need to juggle with signs, using the transformation $-1/x$ instead of $1/x$ because the latter does not preserve the unit disc (the map $x \mapsto 1/x$ preserves the unit circle but turns the disc inside out). Otherwise, the function $\Phi|_\Sigma : \Sigma \rightarrow \mathbb{T}$ mirrors the usual continued fraction numeration process.

Remark. Even a quick glance on Fig. 3 reveals that parts of the circle seem to be missing. While Φ is indeed surjective, the convergence of the images of 0 is sometimes quite slow in this system and so the depth used in the computer graphics was not enough to get near certain points. We can improve the speed of convergence by adding more transformations like in [10].

Example 11. As a last example, we obtain a circle variant of the signed binary number system. Take the following four upper half plane preserving transformations:

$$\begin{aligned}\hat{F}_{\bar{1}}(x) &= (x-1)/2 \\ \hat{F}_0(x) &= x/2 \\ \hat{F}_1(x) &= (x+1)/2 \\ \hat{F}_2(x) &= 2x.\end{aligned}$$

Again, we conjugate these MTs with the stereographic projection to be disc preserving:

$$\begin{aligned}F_{\bar{1}}(x) &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 3-i & -1-i \\ -1+i & 3+i \end{pmatrix} \\ F_0(x) &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 3 & -i \\ i & 3 \end{pmatrix} \\ F_1(x) &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 3+i & 1-i \\ 1+i & 3-i \end{pmatrix} \\ F_2(x) &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}\end{aligned}$$

We take these transformations as our iterative system and then define the shift $\Sigma \subseteq \{0, 1, \bar{1}, 2\}^\omega$ by forbidding the words $20, 02, 12, \bar{1}2, 1\bar{1}$ and $\bar{1}1$.

Why are we forbidding these words? The reason for disallowing 2 and 0 next to each other is that these transformations are inverse to each other. The first

subshift. An *interval system almost compatible* with $\{F_a : a \in A\}$ and Σ is any family of sets $\mathcal{W} = \{W_v : v \in A^*\}$ such that:

1. Each W_v is a finite union of disjoint open intervals on \mathbb{T} .
2. We have $W_v = \mathbb{T}$ iff $v = \lambda$.
3. For all u, v words we have $W_{uv} = W_u \cap F_u(W_v)$.
4. For all words v we have $\overline{W}_v = \bigcup \{\overline{W}_{va} : va \in \mathcal{L}(\Sigma)\}$.

Observe that thanks to (2) and (3), any interval system almost compatible with a given iterative system is uniquely described by the sets W_a , $a \in A$, as we must have:

$$W_v = W_{v_0} \cap F_{v_0}(W_{v_1}) \cap \cdots \cap F_{v_{|0, n-1|}}(W_{v_{n-1}}).$$

Observation 13. *If Σ is the full shift A^ω then condition (4) in the definition of almost compatibility can be replaced by $\mathbb{T} = \bigcup_{a \in A} \overline{W}_a$.*

Proof. Observe that if (2) and (4) hold then, letting $v = \lambda$, we obtain $\mathbb{T} = \bigcup_{a \in A} \overline{W}_a$.

On the other hand, assume that we have a family $\mathcal{W} = \{W_v : v \in A^*\}$ satisfying (1–3) and $\mathbb{T} = \bigcup_{a \in A} \overline{W}_a$. We want to prove that then $\overline{W}_v = \bigcup \{\overline{W}_{va} : a \in A\}$. A little thought gives us the chain of equalities

$$\bigcup_{va \in \mathcal{L}(\Sigma)} (\overline{W}_v \cap F_v(\overline{W}_a)) = \overline{W}_v \cap F_v \left(\bigcup_{a \in A} \overline{W}_a \right) = \overline{W}_v \cap F_v(\mathbb{T}) = \overline{W}_v.$$

Now observe that the sets $\overline{W}_{va} = \overline{W}_v \cap F_v(\overline{W}_a)$ and $\overline{W}_v \cap F_v(\overline{W}_a)$ differ only in a finite number of points (because all members of \mathcal{W} are finite unions of intervals). Therefore $D = \bigcup \{\overline{W}_{va} : a \in A\} \setminus \overline{W}_v$ is finite. But then D is an open set in the space $\bigcup \{\overline{W}_{va} : a \in A\}$ which is a disjoint union of (nondegenerate) intervals. But the only finite open subset of $\bigcup \{\overline{W}_{va} : a \in A\}$ is the empty set. Therefore, $D = \emptyset$ and $\overline{W}_v = \bigcup \{\overline{W}_{va} : a \in A\}$ and we are done. \square

Definition 14. Let $\{F_a : a \in A\}$ be a Möbius iterative system, $\Sigma \subseteq A^\omega$ a subshift and \mathcal{W} an interval system almost compatible with $\{F_a : a \in A\}$ and Σ . We then define the *expansion subshift* of Σ and \mathcal{W} by

$$\Sigma_{\mathcal{W}} = \{u \in \Sigma : \forall n \in \mathbb{N}, W_{u_{|0, n|}} \neq \emptyset\}.$$

The following theorem can be found in [5] as Corollary 27, however we have changed the notation to that of [7]:

Theorem 15. *Let $\{F_a : a \in A\}$ be a Möbius iterative system, Σ a subshift and $B \subseteq B^+$ a finite set of words. Assume that \mathcal{W} is such an interval system almost compatible with Σ that $W_b \subseteq V_b$ for every $b \in B$ and each $w \in \Sigma_{\mathcal{W}}$ contains as a prefix some $b \in B$.*

Then $\Sigma_{\mathcal{W}}$ is a Möbius number system for the iterative system $\{F_a : a \in A\}$. Moreover, $\Phi([v] \cap \Sigma_{\mathcal{W}}) = \overline{W}_v$ for all $v \in A^$.*

6. Examples revisited

We now return to the three number systems presented at the end of Section 4 and prove that they indeed are Möbius number systems. The main practical advantage of using Theorem 15 is that it turns verifying convergence, continuity and surjectivity of $\Phi|_\Sigma$ into combinatorial problems (finding \mathcal{W} and Σ , describing $\Sigma_{\mathcal{W}}$ and finding B so that $W_b \subseteq V_b$).

First, let us revisit **Example 9**. We have the three parabolic transformations F_a, F_b and F_c . Observe that $V_a = (A, B)$, $V_b = (B, C)$ and $V_c = (C, A)$. Let \mathcal{W} be the interval system generated by $W_a = V_a, W_b = V_b$ and $W_c = V_c$, i.e.

$$W_u = W_{u_0} \cap F_{u_0}(W_{u_1}) \cap \cdots \cap F_{u_{|u|-1}}(W_{|u|-1}).$$

then \mathcal{W} is almost compatible with $\{F_a, F_b, F_c\}$.

What is more, the expansion subshift Σ of A^ω and \mathcal{W} is precisely the shift Σ obtained by forbidding the words ac, ba, cb . One way to show this is to first show that $W_{ac} = W_{ba} = W_{cb} = \emptyset$ and then verify that whenever $u \in A^n$ does not contain any forbidden factor then $W_u = F_{u_{[0, n-1]}}(W_{u_{n-1}})$.

We can prove the last equality by induction on n : For $n = 1$ the claim is trivial, while for $n = 2$ we can examine all the (finitely many) cases. Assume that the claim is true for some n and let $u \in A^{n+1}$. Then:

$$\begin{aligned} W_u &= W_{u_{[0, n]}} \cap F_{u_{[0, n]}}(W_{u_n}) = F_{u_{[0, n-1]}}(W_{u_{n-1}}) \cap F_{u_{[0, n]}}(W_{u_n}) \\ &= F_{u_{[0, n-1]}}(W_{u_{n-1}} \cap F_{u_{n-1}}(W_{u_n})) = F_{u_{[0, n-1]}}(W_{u_{[n-1, n]}}). \end{aligned}$$

Now $W_{u_{[n-1, n]}} = F_{u_{n-1}}(W_{u_n})$ by the induction hypothesis for $n = 2$ and so:

$$W_u = F_{u_{[0, n-1]}}(F_{u_{n-1}}(W_{u_n})) = F_{u_{[0, n]}}(W_{u_n}).$$

To finish the proof, we let $B = \{a, b, c\}$ and apply Theorem 15.

In the case of the continued fraction system from **Example 10**, let

$$W_{\bar{1}} = (i, -1), W_0 = (-1, 1) \text{ and } W_1 = (1, i).$$

and again consider the interval system \mathcal{W} generated by these sets.

It is straightforward to see that then $(A^\omega)_{\mathcal{W}}$ is precisely the subshift Σ defined by forbidding $00, 1\bar{1}, \bar{1}1, 101, 10\bar{1}$. It remains to choose the set $B = \{01, 0\bar{1}, 1, \bar{1}\}$ (which contains a prefix of every word $w \in \Sigma$) and verify that $W_b \subseteq V_b$ for each $b \in B$.

By Theorem 15, we again conclude that Σ together with $\{F_{\bar{1}}, F_0, F_1\}$ is a Möbius number system.

Finally, we analyze the *signed binary system* from **Example 11**. Proving that this system is indeed a Möbius number system requires a reasonable amount of computation which we have decided to skip here, presenting only the main points of the proof.

Define the shift Σ_0 by forbidding the words $02, 20, 12$ and $\bar{1}2$. Then let $v : \bar{\mathbb{R}} \rightarrow \mathbb{T}$ denote the inverse of the stereographic projection and consider the interval almost cover \mathcal{W} generated by

$$W_{\bar{1}} = (-1, q^-), W_0 = (h^-, h^+), W_1 = (q^+, 1), W_2 = (h^-, h^+),$$

where

$$\begin{aligned} q^- &= v(-1/4) = \frac{-8 - 15i}{17} \\ q^+ &= v(1/4) = \frac{8 - 15i}{17} \\ h^- &= v(-1/2) = \frac{-4 - 3i}{5} \\ h^+ &= v(1/2) = \frac{4 - 3i}{5}. \end{aligned}$$

Next, we should show that \mathcal{W} is compatible with Σ_0 and that $(\Sigma_0)_{\mathcal{W}} = \Sigma$ is the subshift defined by the forbidden words $20, 02, 12, \bar{1}2, 1\bar{1}, \bar{1}1$. This follows from the set of identities:

$$\begin{aligned} F_2(\bar{W}_2) \cup F_2(\bar{W}_1) \cup F_2(\bar{W}_{\bar{1}}) &= \bar{W}_2 \\ F_0(\bar{W}_0) \cup F_0(\bar{W}_1) \cup F_0(\bar{W}_{\bar{1}}) &= \bar{W}_0 \\ F_1(\bar{W}_0) \cup F_1(\bar{W}_1) &= \bar{W}_1 \\ F_{\bar{1}}(\bar{W}_0) \cup F_{\bar{1}}(\bar{W}_{\bar{1}}) &= \bar{W}_{\bar{1}} \end{aligned}$$

together with

$$F_1(W_{\bar{1}}) \cap W_1 = F_{\bar{1}}(W_1) \cap W_{\bar{1}} = \emptyset.$$

It remains to take $B = \{0, 1, \bar{1}, 21, 2\bar{1}, 22\}$ and check the requirements of Theorem 15. It is easy to see that each $w \in \Sigma$ contains as a prefix a member of B . Finally, a detailed calculation (which we omit here) will verify that for every $b \in B$ we have $W_b \subseteq V_b$, therefore the signed binary system is a Möbius number system by Theorem 15.

7. Existence results

The most basic existence question one might ask is whether there exists any Möbius number system at all for a given iterative system. This problem is not solved yet, however, we can offer a partial answer based on Theorem 9 in [10] and Theorem 15.

Theorem 16. *Let $F : A^+ \times \mathbb{T} \rightarrow \mathbb{T}$ be a Möbius iterative system.*

1. *If $\overline{\bigcup_{u \in A^+} V_u} \neq \mathbb{T}$ then $\Phi(\mathbb{X}_F) \neq \mathbb{T}$ and so there is no Möbius number system for $\{F_a : a \in A\}$.*
2. *If there exists a finite set $B \subseteq A^+$ such that $\{\bar{V}_u : u \in B\}$ is a cover of \mathbb{T} then there exists a subshift Σ that is a Möbius number system for $\{F_a : a \in A\}$.*

While Theorem 16 gives a necessary condition for a Möbius number system to exist, this condition is not very comfortable to use. In the spirit of [9], we offer a condition that is easier to check.

Let $\{F_a : a \in A\}$ be a Möbius iterative system. A nonempty closed set $W \subseteq \mathbb{T}$ is *inward* if $\bigcup_{a \in A} F_a(W) \subseteq \text{Int}(W)$. All iterative systems have the trivial inward set \mathbb{T} and some systems have nontrivial inward sets as well.

Theorem 17. *Let $\{F_a : a \in A\}$ be an iterative system with a nontrivial inward set. Then there is no Möbius number system for $\{F_a : a \in A\}$.*

Proof. Assume that W is a nontrivial inward set. Let $z \notin W$. Because W^c is open there exists an open interval I disjoint with W and containing z . Assume that $\Phi(w) = z$. Then $\lim_{k \rightarrow \infty} |F_{w_{[0,k]}}^{-1}(I)| = 2\pi$. However, $\text{Int } W$ is nonempty so there exists an open interval $J \subseteq \text{Int } W$. Now for all k we have $F_{w_{[0,k]}}(J) \subseteq W$ so $J \cap F_{w_{[0,k]}}^{-1}(I) = \emptyset$. But then $|F_{w_{[0,k]}}^{-1}(I)| \leq 2\pi - |J|$, a contradiction. \square

Searching for the solution to the existence problem, we have used a numerical simulation to obtain insight in the behavior of MTs. The results suggest that closing the gap in Theorem 16 is an achievable task.

We have studied the behavior of the iterative system $\{F_a, F_b\}$ consisting of two hyperbolic transformations. The transformation F_a has fixed points 1 (stable) and $-i$ (unstable), while the transformation F_b has fixed points -1 (stable) and i (unstable). We have parameterized F_a, F_b by the values q_a, q_b of $(F_i)^\bullet$ at stable points, so we have:

$$\begin{aligned} F_a &= \frac{1}{2\sqrt{q_a}} \begin{pmatrix} 1 + q_a - i(1 - q_a) & 1 - q_a + i(1 - q_a) \\ 1 - q_a - i(1 - q_a) & 1 + q_a + i(1 - q_a) \end{pmatrix} \\ F_b &= \frac{1}{2\sqrt{q_b}} \begin{pmatrix} 1 + q_b - i(1 - q_b) & -1 + q_b - i(1 - q_b) \\ -1 + q_b + i(1 - q_b) & 1 + q_b + i(1 - q_b) \end{pmatrix}. \end{aligned}$$

Denote

$$Y = \{(q_a, q_b) : q_a, q_b \in (0, 1), \text{ there exists a Möbius number syst. for } \{F_a, F_b\}\}.$$

We wrote a C program that tries various pairs (q_a, q_b) , constructs F_a, F_b , then computes the intervals $\{\bar{V}_v : |v| \leq m\}$ (where m is the number of iterations to consider) and finally checks whether these intervals cover the whole \mathbb{T} . If they do, the program puts a white dot on the corresponding place in the graph, otherwise we leave it black.

As we are interested in characterization, we have plotted (in gray) a second set in the graph: The set U of all choices of (q_a, q_b) such that the iterative system $\{F_a, F_b\}$ has a nontrivial inward set. The formula for U , as shown in [9], is $U = \bigcup_{n \in \mathbb{Z}} U_n$, where for $n > 0$ we have:

$$\begin{aligned} U_0 &= U_{ab} \cap \left(0, \frac{1}{2}\right) \times \left(0, \frac{1}{2}\right) \\ U_n &= U_{a^n b} \cap U_{a^{n+1} b} \cap \left(\frac{1}{\sqrt[2]{n}}, \frac{1}{\sqrt[2]{n+1}}\right) \times \left(0, \frac{1}{2}\right) \\ U_{-n} &= U_{ab^n} \cap U_{ab^{n+1}} \cap \left(0, \frac{1}{2}\right) \times \left(\frac{1}{\sqrt[2]{n}}, \frac{1}{\sqrt[2]{n+1}}\right) \end{aligned}$$

and $U_v = \{(q_a, q_b) : F_v \text{ is hyperbolic}\}$ for $v \in A^*$. For practical reasons, we have only drawn the sets U_n with $|n| \leq m$ (the same m as the number of iterations in

the first part of the program). The result for $m = 10$ and resolution 1000×1000 is shown in Fig. 5.

By Theorem 17, $U \cap Y = \emptyset$. We are interested in the size of the complement of $U \cup Y$ in $(0, 1)^2$. It turns out that $U \cup Y$ covers most of the unit square and U and Y appear to fit rather well together. The area between the two sets is likely to get smaller and smaller as we let m grow, possibly shrinking to zero in the limit. However, there might still exist points that do not belong either to U or to Y . We conjecture, that $U \cup Y$ is equal to the whole $(0, 1)^2$ perhaps up to some countable set of exceptional points.

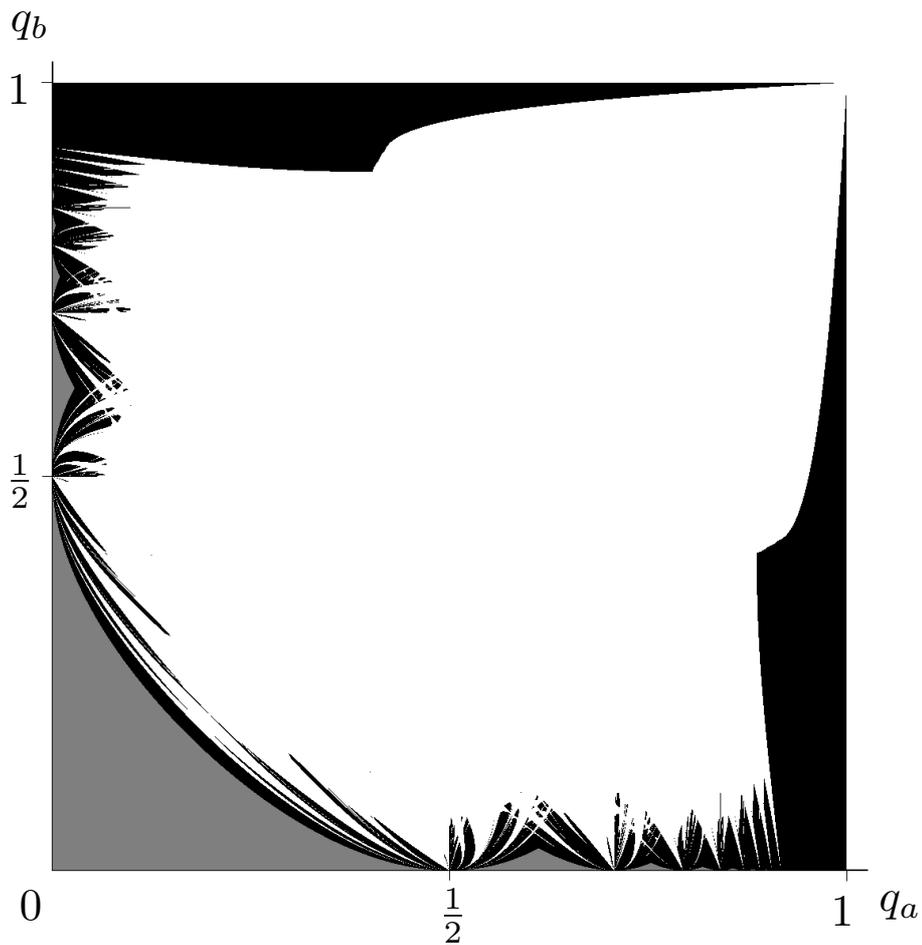


Figure 5: The graph of U and Y for the depth $m = 10$ and resolution 1000×1000 .

8. Subshifts not admitting Möbius number systems

An interesting question one can pose is whether a given subshift $\Sigma \subseteq A^\omega$ admits a Möbius number system, i.e. whether there exists an iterative system $\{F_a : a \in A\}$ such that Σ is a Möbius number system for $\{F_a : a \in A\}$.

Trivially, the cardinality of Σ must be precisely continuum, as for smaller Σ there is no projection from Σ onto \mathbb{T} . We offer a less trivial necessary condition.

A (non-erasing) *substitution* is any mapping $\psi : B \rightarrow A^+$. We can extend ψ to a map from B^ω to A^ω in a natural way, expanding each letter and gluing thus obtained words together. As there is no risk of confusion, we will denote the resulting map by ψ as well and call it a *substitution map*. Observe that $\psi : B^\omega \rightarrow A^\omega$ is continuous in the product topology.

Theorem 18. *Let $\Sigma \subseteq A^\omega$ be a Möbius number system for an iterative system $\{F_a : a \in A\}$. Then for all alphabets B and all substitution maps ψ we have $\Sigma \neq \psi(B^\omega)$.*

Proof. Assume that there exist B and ψ for which the claim is false. Then B^ω together with the maps $\{G_b : b \in B\}$ such that $G_b = F_{\psi(b)}$ is a Möbius number system. Denote by Φ the resulting projection of B^ω to \mathbb{T} and observe that $\Phi = \Phi_\Sigma \circ \psi$ where $\Phi_\Sigma : \Sigma \rightarrow \mathbb{T}$ is the number system on Σ . We see that Φ is surjective and continuous on B^ω .

Let us choose any $x \in \mathbb{T}$ and $u \in B^\omega$. There exists $v \in B^\omega$ with $\Phi(v) = x$. Consider the sequence $\{\Phi(u_{[0,k]}v)\}_{k=1}^\infty$. We have

$$\Phi(u_{[0,k]}v) = G_{u_{[0,k]}}(\Phi(v)) = G_{u_{[0,k]}}(x).$$

We know that Φ is continuous, so $\Phi(u_{[0,k]}v)$ tends to $\Phi(u)$ when k tends to infinity. Therefore, $G_{u_{[0,k]}}(x) \rightarrow \Phi(u)$. As x, u were arbitrary, we have shown that for every $u \in B^\omega$ and each point x of \mathbb{T} , the sequence $\{G_{u_{[0,k]}}(x)\}_{k=1}^\infty$ converges to $\Phi(u)$.

Let u be periodic with some period $w \in B^+$. Then $\Phi(u)$ is the (stable) fixed point of G_w . Therefore G_w may not be elliptic ($\Phi(u)$ would not be defined) nor hyperbolic (the images of the unstable fixed point of G_w would not converge). This means that for all $w \in B^+$ the transformation G_w must be parabolic.

If all the transformations $G_b, b \in B$ were parabolic with the same fixed point x then all the transformations $G_w, w \in B^+$ would be parabolic with fixed point x and we would not be able to represent anything except x . Therefore, there exist $a, b \in B$ such that G_a, G_b have different fixed points. The rest of the proof consists of a straightforward (but technical) calculation that such a situation is impossible.

Choose $a, b \in B$ so that G_a, G_b have different fixed points. We know that G_{ab} and G_{aab} must both be parabolic. Without loss of generality assume that G_a is similar to the matrix $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Write $G_a = MJM^{-1}$ where M is a regular matrix. Recall that a transformation F is parabolic iff $\text{Tr}(F)^2 = 4$ and observe that:

$$\text{Tr}(G_{ab}) = \text{Tr}(MJM^{-1}G_b) = \text{Tr}(JM^{-1}G_bM)$$

and

$$\mathrm{Tr}(G_{aab}) = \mathrm{Tr}(MJ^2M^{-1}G_b) = \mathrm{Tr}(J^2M^{-1}G_bM),$$

where we used the equality $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Let

$$M^{-1}G_bM = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and observe that $(a+d)^2 = \mathrm{Tr}(G_b)^2 = 4$.

We now have

$$JM^{-1}G_bM = \begin{pmatrix} a+c & b+d \\ c & d \end{pmatrix}, \quad J^2M^{-1}G_bM = \begin{pmatrix} a+2c & b+2d \\ c & d \end{pmatrix}.$$

Calculating the traces, we obtain the equalities

$$(a+2c+d)^2 = (a+c+d)^2 = (a+d)^2 = 4$$

that can only be satisfied when $c = 0$. But then matrices J and $M^{-1}G_bM$ share the eigenvector $(1, 0)^T$ and so $G_a = MJM^{-1}$ and G_b share the eigenvector $M \cdot (1, 0)^T$. However, eigenvectors of matrices are in one to one correspondence with fixed points of Möbius transformations (see the proof of Lemma 22 in the Appendix) and so G_a, G_b have the same fixed point, a contradiction. \square

Theorem 18 tells us in particular that there are no Möbius number systems on the full shift. We must always forbid some words in order to get rid of undesirable concatenations.

A rather unfortunate flaw of Theorem 18 is that the set $\psi(B^\omega)$ need not be a subshift: while $\psi(B^\omega)$ is always closed, σ -invariance is not guaranteed.

However, there are cases when $\psi(B^\omega)$ is a nontrivial subshift. Consider the Fibonacci shift Σ_F defined on the alphabet $\{0, 1\}$ by forbidding the factor 11. It is easy to see that $\Sigma_F = \psi(\{0, 1\}^\omega)$ under the substitution

$$\psi : 0 \mapsto 0, 1 \mapsto 10.$$

Therefore, Σ_F can never be a Möbius number system.

9. Sofic Möbius number systems

In this section, we will explore another facet of Möbius number systems. As every number system is a subshift, we can ask how complicated (in the sense of formal language theory, not information theory) is the language of this subshift.

A subshift Σ is of *finite type* if Σ can be defined using a finite set of forbidden words (note that this was the case in all our example subshifts). A subshift Σ is called *sofic* if and only if the language of Σ is regular (recognizable by a finite automaton). It is straightforward to show that subshifts of finite type are always sofic. Sofic subshifts and subshifts of finite type are quite popular in practice, as they are easier to manipulate than general subshifts. There are numerous results and algorithms available for sofic subshifts and subshifts of finite type.

The papers [8] and [10] contain several examples of Möbius number systems that are subshifts of finite type. Furthermore, Proposition 5 in [10] states a sufficient condition for a number system to be of finite type, while Theorem 24 in [4] offers a rather complicated condition for a number system to be sofic. We now present two similar but easier to understand conditions, one sufficient and one necessary.

Theorem 19. *Let $\{F_a : a \in A\}$ be an iterative system, Σ be a sofic subshift and \mathcal{W} such an interval system almost compatible with $\{F_a : a \in A\}$ that the set $\{F_v^{-1}(W_v) : v \in A^*\}$ is finite. Then $\Sigma_{\mathcal{W}}$ is sofic.*

Proof. We construct a finite automaton \mathcal{A} that recognizes all the words $v \in A^*$ such that $W_v \neq \emptyset$. We then intersect the resulting regular language with the language $\mathcal{L}(\Sigma)$ to obtain $\mathcal{L}(\Sigma_{\mathcal{W}})$. Because regular languages are closed under intersection, $\mathcal{L}(\Sigma_{\mathcal{W}})$ is regular.

The states of our automaton will be all the sets $Z_v = F_v^{-1}(W_v)$, $v \in A^*$. We let Z_λ to be the initial state and all states except \emptyset to be accepting states. A transition labelled by the letter a leads from Z_v to Z_{va} for every $v \in A^*$ and every $a \in A$.

We verify that when $Z_v = Z_u$ then $Z_{va} = Z_{ua}$, so the definition of our automaton is correct:

$$Z_{va} = F_a^{-1}F_v^{-1}(W_v \cap F_v(W_a)) = F_a^{-1}(F_v^{-1}(W_v)) \cap F_a^{-1}(W_a) = F_a^{-1}(Z_v) \cap Z_a$$

Similarly, $Z_{ua} = F_a^{-1}(Z_u) \cap Z_a$ and as $Z_u = Z_v$ we obtain $Z_{va} = Z_{ua}$.

To finish the proof, we observe that the automaton \mathcal{A} accepts the word v iff $Z_v \neq \emptyset$. Because $Z_v \neq \emptyset$ iff $W_v \neq \emptyset$, \mathcal{A} recognizes precisely those $v \in A^*$ with $W_v \neq \emptyset$. \square

Under an additional assumption, we can prove the converse of Theorem 19:

Theorem 20. *Assume that \mathcal{W} is an interval almost cover compatible with the subshift Σ . Let the subshift $\Sigma_{\mathcal{W}}$ be a sofic Möbius number system such that $\Phi([v] \cap \Sigma_{\mathcal{W}}) = \overline{W}_v$ for every word v . Then the set $\{F_v^{-1}(W_v) : v \in A^*\}$ is finite.*

Proof. To prove this theorem we define a chain of several finite sets, each obtained from the previous, with the final set being $\{F_v^{-1}(W_v) : v \in A^*\}$.

Denote by $\mathcal{F}(v)$ the *follower set* of v in $\Sigma_{\mathcal{W}}$, i.e. the set of all words $u \in A^*$ such that $vu \in \mathcal{L}(\Sigma_{\mathcal{W}})$. By the Myhill-Nerode theorem, we have that if $\Sigma_{\mathcal{W}}$ is sofic, then $\{\mathcal{F}(v) : v \in A^*\}$ is a finite set. Let $v \in A^*$ and denote $\mathcal{F}^\omega(v) = \{w \in A^\omega : \forall k, w_{[0,k]} \in \mathcal{F}(v)\}$. The set $\{\mathcal{F}^\omega(v) : v \in A^*\}$ is finite as each $\mathcal{F}^\omega(v)$ depends only on $\mathcal{F}(v)$.

A little thought gives us that $\mathcal{F}^\omega(v) = \{w \in A^\omega : vw \in \Sigma_{\mathcal{W}}\}$. Finally, denote $Z_v = \Phi(\mathcal{F}^\omega(v))$ and observe again that the set $\{Z_v : v \in A^*\}$ is finite. It remains to notice that

$$Z_v = \Phi(\{w \in A^\omega : vw \in \Sigma_{\mathcal{W}}\}) = F_v^{-1}(\Phi([v] \cap \Sigma_{\mathcal{W}})) = F_v^{-1}(W_v)$$

to see that the set $\{F_v^{-1}(W_v) : v \in A^*\}$ must be finite. \square

Observe that Theorems 19 and 20 give us that if Σ is sofic then the interval shift considered in Theorem 15 is sofic if and only if $\{F_v^{-1}(W_v) : v \in A^*\}$ is a finite set. Therefore, we have obtained a tool to decide whether $\Sigma_{\mathcal{W}}$ is sofic.

10. Conclusions and open problems

In the whole paper we have explored various topics in the theory of Möbius number systems. We have reviewed tools to prove that a subshift is a Möbius number system for a given iterative system as well as various existence results and criteria for sofic number systems. However, there remain quite a few open problems, practical as well as theoretical, in this area.

Expansion subshifts seem to be useful when dealing with concrete examples. Theorem 15 offers practical tools to prove that a given subshift is a Möbius number system for a given iterative system. We wonder how the available toolbox for this kind of proof could be further improved. We see numerous areas open to incremental improvements.

For the sake of examples and applications, we would like to have a sufficient and necessary condition for the existence of a Möbius number system for a given iterative system. Ideally, this condition should be effectively verifiable (for reasonable iterative systems, say when real and imaginary parts of coefficients are rational). While we doubt that a general effective algorithm exists, improvements in the tools for proving the existence of Möbius iterative systems would be welcome indeed.

Another, perhaps less practical, but combinatorially interesting problem is when a given subshift Σ can be a Möbius number system. So far, we have some sufficient and some necessary conditions and a large gap in between.

To manipulate number systems, it would be nice to have a sofic Möbius number system. Theorems 19 and 20 offer useful checks to perform when verifying if a number system is sofic. What is completely missing is a condition, similar to Theorem 15, for the existence of a sofic number system for a given iterative set. We are hoping that obtaining such a result is possible but the current amount of knowledge on sofic number systems is rather small. For example, we can not even tell whether existence of a Möbius number system implies existence of a sofic system for the same iterative system or not.

A large part the complexity of above problems seems to come not from the number systems themselves but from the fact that we don't properly understand how do large numbers of MTs compose (or, equivalently, how long sequences of matrices multiply). This suggests that maybe the way forward lies in studying the limits of products of matrices. Unfortunately, this area is full of hard questions, see for example [2].

Hard problems notwithstanding, we conclude on a positive note: Although there are numerous open questions about Möbius number systems, and some properties of these systems might turn out to be undecidable, current tools do allow us to deal with systems that are likely to be used elsewhere (for example, the continued fraction number system).

Acknowledgements

The text is based on our Master's thesis [5] written under the supervision of prof. Petr Kůrka (which, in turn, is an extension of [4]). We thank prof. Kůrka for support and stimulating discussions. The research was supported by the Czech Science Foundation research project GAČR 201/09/0854.

11. Appendix

The Appendix contains various proofs that we felt should be included in this paper, yet their length or technical nature would disturb the flow of the rest of the text. Note that these are all well known results; the proofs here are just for the sake of completeness and better understanding of the topic.

Let us first show that all the disc preserving Möbius transformations have a rather simple form.

Lemma 21. *A Möbius transformation F is disc preserving (i.e. $F(\mathbb{D}) = \mathbb{D}$) iff it is of the form*

$$F = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix},$$

where $|\alpha|^2 - |\beta|^2 = 1$.

Proof. Let F have the given form. We prove that then $F(\mathbb{D}) = \mathbb{D}$. First, consider $z = e^{i\phi}$. We have:

$$|\bar{\beta}e^{i\phi} + \bar{\alpha}| = \left| \overline{(\alpha e^{i\phi} + \beta)} e^{i\phi} \right| = |\alpha e^{i\phi} + \beta|.$$

And so

$$|F(e^{i\phi})| = \frac{|\alpha e^{i\phi} + \beta|}{|\bar{\beta}e^{i\phi} + \bar{\alpha}|} = 1.$$

Therefore $F(\mathbb{T}) \subseteq \mathbb{T}$. Because F is an MT, the image of \mathbb{T} must be a circle, so $F(\mathbb{T}) = \mathbb{T}$. The unit circle divides $\bar{\mathbb{C}}$ into two components: The inside (containing zero) and the outside (containing ∞). As F is a bijection on $\bar{\mathbb{C}}$, all we have to do to obtain $F(\mathbb{D}) = \mathbb{D}$ is prove that $F(0)$ lies inside \mathbb{D} . But this is simple: $F(0) = \frac{\beta}{\bar{\alpha}}$ and $|\beta| < |\alpha|$, so $|F(0)| < 1$.

On the other hand, consider any disc preserving MT $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $\det F = 1$. Because F is continuous, it must be $F(\mathbb{T}) = \mathbb{T}$. Therefore, for every ϕ , we must have $|ae^{i\phi} + b| = |ce^{i\phi} + d|$. A little thought gives us that if $a = 0$ then $d = 0$ and similarly $b = 0$ implies $c = 0$; in both cases we are done. Assume $a, b, c, d \neq 0$ and continue.

Choose ϕ so that the quantity $|ae^{i\phi} + b| = |ce^{i\phi} + d|$ is maximal. The maximal value of the function on the left side is $|a| + |b|$, on the right side $|c| + |d|$, thus $|a| + |b| = |c| + |d|$. Similarly, by choosing the minimal quantity, we obtain that $||a| - |b|| = ||c| - |d||$. Moreover, as ϕ is the same on the right and left, we also have $\arg a - \arg b = \arg c - \arg d$. These three equalities will be enough to complete the proof.

Assume for a moment that the equality $||a| - |b|| = ||c| - |d||$ actually means $|a| - |b| = |c| - |d|$. Then, together with $|a| + |b| = |c| + |d|$, we have $|a| = |c|$ and $|b| = |d|$, obtaining a matrix of the form $F = \begin{pmatrix} a & b \\ ae^{i\psi} & be^{i\psi} \end{pmatrix}$. But this matrix is singular, a contradiction.

Therefore, we must have $|a| - |b| = |d| - |c|$, which implies $|a| = |d|, |b| = |c|$ and, after a brief calculation,

$$F = \begin{pmatrix} a & b \\ \bar{b}e^{i\psi} & \bar{a}e^{i\psi} \end{pmatrix}$$

for a suitable ψ .

Now it remains to use the normalization formula $\det F = 1$ to see that ψ is either 0 or π . If $\psi = 0$, we are done. Otherwise, we would have $\det F = |b|^2 - |a|^2 = 1$, so $|b| > |a|$. But then $|F(0)| = \frac{|b|}{|a|} > 1$, so F would turn the disc inside out, a contradiction. \square

We now present a series of three lemmas concerning the classification of disc preserving MTs into elliptic, parabolic and hyperbolic transformations. We will use a bit of linear algebra machinery. We will understand each disc preserving Möbius transformation F both as a map and as the corresponding normalized matrix

$$F = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, |\alpha|^2 - |\beta|^2 = 1.$$

An important role in the following proofs belongs to the eigenvalues of the matrix F . However, these eigenvalues are not uniquely defined: The Möbius transformation F always has two corresponding normalized matrices $F, -F$, therefore it also has two different sets of eigenvalues. We deal with this problem by always fixing one matrix of F for the whole proof.

Lemma 22. *Let F be a disc preserving Möbius transformation. Fix a matrix of F such that $\det F = 1$. Then the following holds:*

1. *F is elliptic iff the eigenvalues of F are not real iff F has one fixed point inside and one fixed point outside of \mathbb{T} (the outside point might be ∞),*
2. *F is parabolic iff F has the single eigenvalue equal to 1 or -1 iff F has a single fixed point and it lies on \mathbb{T} ,*
3. *F is hyperbolic iff F has two different real eigenvalues iff F has two different fixed points, both lying on \mathbb{T} .*

Proof. We begin by providing a connection between $(\operatorname{Tr} F)^2$ and the eigenvalues of F . The characteristic polynomial of F is:

$$(\alpha - \lambda)(\bar{\alpha} - \lambda) - \beta\bar{\beta} = |\alpha|^2 - |\beta|^2 - \operatorname{Tr} F \cdot \lambda + \lambda^2 = 1 - \operatorname{Tr} F \cdot \lambda + \lambda^2.$$

We see that for $(\operatorname{Tr} F)^2 < 4$, F has two distinct complex conjugate eigenvalues, while if $(\operatorname{Tr} F)^2 > 4$, then F has two distinct real eigenvalues. Finally, if $(\operatorname{Tr} F)^2 = 4$, there is only one eigenvalue $\lambda = 1$ or $\lambda = -1$.

Observe that for all $z \in \mathbb{C}$ such that $F(z) \neq \infty$ we have:

$$F \cdot \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha z + \beta \\ \bar{\beta} z + \bar{\alpha} \end{pmatrix} = (\bar{\beta} z + \bar{\alpha}) \cdot \begin{pmatrix} F(z) \\ 1 \end{pmatrix}.$$

We see that $z \in \overline{\mathbb{C}}$ is a fixed point of F iff $(z, 1)^T$ is an eigenvector of F . Similarly, the point ∞ is fixed iff $(1, 0)^T$ is an eigenvector of F .

Let λ be an eigenvalue of F . If F is not the identity, then the eigenspace of λ must have dimension 1 (otherwise all the points of $\overline{\mathbb{C}}$ would be fixed points of F). Therefore, we have a one to one correspondence between the eigenvalues and fixed points of F .

Let $v = (v_1, v_2)^T$ be an eigenvector corresponding to the eigenvalue λ . Then:

$$\begin{aligned} F \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \begin{pmatrix} \alpha v_1 + \beta v_2 \\ \bar{\beta} v_1 + \bar{\alpha} v_2 \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ F \cdot \begin{pmatrix} \bar{v}_2 \\ \bar{v}_1 \end{pmatrix} &= \begin{pmatrix} \beta \bar{v}_1 + \alpha \bar{v}_2 \\ \bar{\alpha} \bar{v}_1 + \bar{\beta} \bar{v}_2 \end{pmatrix} = \begin{pmatrix} \overline{\beta v_1 + \alpha v_2} \\ \overline{\alpha v_1 + \beta v_2} \end{pmatrix} = \begin{pmatrix} \overline{\lambda v_2} \\ \overline{\lambda v_1} \end{pmatrix} = \bar{\lambda} \begin{pmatrix} \bar{v}_2 \\ \bar{v}_1 \end{pmatrix}. \end{aligned}$$

If λ is real, then the vectors $(v_1, v_2)^T$ and $(\bar{v}_2, \bar{v}_1)^T$ must be linearly dependent as they both belong to the same eigenspace. In particular, if $v_2 = 0$ we would have $v_1 = 0$, so we can assume that $v_1 = z, v_2 = 1$. But then the linear dependence is equivalent with

$$\det \begin{pmatrix} z & 1 \\ 1 & \bar{z} \end{pmatrix} = 0$$

which is equivalent with $|z| = 1$ and so $z \in \mathbb{T}$.

On the other hand, if λ is not real then $\bar{\lambda} \neq \lambda$ and the vectors $(v_1, v_2)^T$ and $(\bar{v}_2, \bar{v}_1)^T$ must be linearly independent as they are eigenvectors of different eigenvalues. Assume $z \in \mathbb{C}$ is a fixed point of F . Were $|z| = 1$ then the determinant argument above would give us a contradiction with linear independence. Therefore $z \notin \mathbb{T}$. But there is more: If z is a fixed point of F then so is $\frac{1}{\bar{z}}$, the image of z under circle inversion with respect to \mathbb{T} :

$$F \cdot \begin{pmatrix} \frac{1}{\bar{z}} \\ 1 \end{pmatrix} = \frac{1}{\bar{z}} F \cdot \begin{pmatrix} 1 \\ \bar{z} \end{pmatrix} = \frac{1}{\bar{z}} \bar{\lambda} \begin{pmatrix} 1 \\ \bar{z} \end{pmatrix} = \bar{\lambda} \begin{pmatrix} \frac{1}{\bar{z}} \\ 1 \end{pmatrix}.$$

Similarly, if ∞ is a fixed point then so is 0. Therefore, if F has an eigenvalue that is not real, then F has one fixed point outside \mathbb{T} and one inside \mathbb{T} (and we can even map one onto another using the circle inversion with respect to \mathbb{T}). This is precisely the case of elliptic F .

On the other hand, if F is hyperbolic, then F has two distinct real eigenvalues and, therefore, two distinct fixed points on \mathbb{T} .

If F is elliptic then there is a single real eigenvalue of F and so the Möbius transformation F has a single fixed point on \mathbb{T} . \square

Lemma 23. *Assume F is a hyperbolic transformation, x_1 and x_2 its fixed points. Then $F'(x_1) = \lambda_2/\lambda_1$ and $F'(x_2) = \lambda_1/\lambda_2$ where λ_1, λ_2 are the eigenvalues of F associated to x_1 and x_2 .*

Similarly, if F is a parabolic transformation and x is its fixed point then $F'(x) = 1$.

Proof. First observe that the ratio of λ_1 and λ_2 does not depend on the choice of the matrix for F so the claim is sensible.

Let us again fix a normalized matrix corresponding to F . Let J be the Jordan matrix similar to the matrix F , i.e. there exists an MT M such that $F = M \circ J \circ M^{-1}$. We can understand J as a Möbius transformation.

Let x be a fixed point of F . Then we have:

$$F'(x) = (M \circ J \circ M^{-1})'(x) = M'(J \circ M^{-1}(x)) \cdot J'(M^{-1}(x)) \cdot (M^{-1})'(x).$$

Because $F(x) = x$, we must have $J \circ M^{-1}(x) = M^{-1}(x)$, so:

$$F'(x) = M'(M^{-1}(x)) \cdot J'(M^{-1}(x)) \cdot (M^{-1})'(x),$$

and using the formula $M'(M^{-1}(x)) \cdot (M^{-1})'(x) = 1$ we get

$$F'(x) = J'(M^{-1}(x)).$$

However, $J'(M^{-1}(x))$ is easy to compute because $M^{-1}(x)$ is the fixed point of J corresponding to the same eigenvalue as x . The only problem is that we have to ensure $M^{-1}(x) \neq \infty$ to have the derivative well defined. We deal with this problem by carefully choosing our J .

Let us begin with the hyperbolic case. To calculate, x_1 choose the Jordan matrix $J = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}$ (note the switched order of λ_1, λ_2). Now $J(z) = \frac{\lambda_2}{\lambda_1}z$ and $M^{-1}(x_1) = 0$. Thus $F'(x_1) = J'(0) = \frac{\lambda_2}{\lambda_1}$. Similarly, we calculate $F'(x_2) = \frac{\lambda_1}{\lambda_2}$ from the Jordan form $J = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

In the parabolic case, we avoid problems with the point at ∞ , by taking a matrix similar to the usual Jordan form but with different interpretation as a Möbius transformation: $J = \begin{pmatrix} 1 & 0 \\ \mp 1 & 1 \end{pmatrix}$ (the sign in the lower left corner depends on the Jordan matrix for F ; there are two possibilities). Now $J(z) = \frac{z}{\mp z + 1}$ and $M^{-1}(x) = 0$, so $F'(x) = J'(0) = 1$ and we are done. \square

Lemma 24. *Let F be a parabolic or a hyperbolic transformation, let x be the (stable, for F hyperbolic) fixed point of F . Let $z \in \mathbb{C}$ (assume that z is not the unstable fixed point of F in the hyperbolic case). Then $\lim_{n \rightarrow \infty} F^n(z) = x$.*

Proof. Let us again fix a matrix of F such that $\det F = 1$. Denote by J the Jordan matrix similar to F . Then $F = M \cdot J \cdot M^{-1}$ and $F^n = M \cdot J^n \cdot M^{-1}$. We know that J^n has one of the two possible forms:

$$\begin{pmatrix} 1 & \pm n \\ 0 & 1 \end{pmatrix}, \text{ or } \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$$

where without loss of generality $|\lambda_1| > 1 > |\lambda_2|$. Let $z \in \overline{\mathbb{C}}$ (if F is hyperbolic, let z be different from the unstable fixed point of F) and consider the vector

$$F^n \begin{pmatrix} z \\ 1 \end{pmatrix} = M^{-1} \cdot J^n \cdot M \cdot \begin{pmatrix} z \\ 1 \end{pmatrix}.$$

It is easy to see that the first component of $J^n \cdot M \cdot \begin{pmatrix} z \\ 1 \end{pmatrix}$ tends to infinity, while the second is bounded. Therefore (understanding J and M as Möbius transformations), we have $\lim_{n \rightarrow \infty} J^n \circ M(z) = \infty$.

However, $M^{-1}(\infty) = x$ as $(x, 1)^T$ and $(1, 0)^T$ are eigenvectors of F and J belonging to the same eigenvalues. It follows that

$$\lim_{n \rightarrow \infty} M^{-1} \circ J^n \circ M(z) = M^{-1}(\infty) = x. \quad \square$$

We conclude the Appendix with a consequence of the Riesz representation theorem. (See [1, page 184] for details). In its statement, we are going to identify $C^*(\mathbb{T}, \mathbb{R})$ with the space of signed Radon measures on \mathbb{T} .

Lemma 25. *Let E be a measurable set on \mathbb{T} . Then the map $\alpha : \lambda \mapsto \lambda(E)$ from $C^*(\mathbb{T}, \mathbb{R})$ to \mathbb{R} is linear and continuous on $C^*(\mathbb{T}, \mathbb{R})$.*

Proof. Linearity of α is obvious. To obtain continuity, it is enough to show that $|\alpha(\lambda)|$ is bounded whenever $|\lambda|$ is bounded. By definition,

$$|\lambda| = \sup\{\lambda(f) : f \in C(\mathbb{T}, \mathbb{R}), |f| \leq 1\}.$$

We first observe that every λ can be written as $\lambda_1 - \lambda_2$ where λ_1, λ_2 are positive measures. Moreover, as shown in [1], we can choose λ_1, λ_2 so that

$$\begin{aligned} \lambda_1(f) &= \sup\{\lambda(g) : 0 \leq g \leq f\} \\ \lambda_2(f) &= \sup\{\lambda(g) : -f \leq g \leq 0\} \end{aligned}$$

for all $f \geq 0$.

Then we have

$$|\lambda| \geq \sup\{\lambda(f) : 0 \leq f \leq 1\} = \lambda_1(1) = \lambda_1(\mathbb{T}) \geq \lambda_1(E) \geq \lambda(E)$$

as well as

$$|\lambda| \geq \sup\{\lambda(f) : -1 \leq f \leq 0\} = \lambda_2(1) = \lambda_2(\mathbb{T}) \geq \lambda_2(E) \geq -\lambda(E).$$

Therefore, $|\lambda| \geq |\lambda(E)| = |\alpha(\lambda)|$ for every λ , proving the continuity of α . \square

References

- [1] Ash., R. B., Measure, Integration and Functional Analysis. Academic Press, London, 1972.
- [2] Blondel, V. D., Theys, J., Tsitsiklis, J. N., When is a pair of matrices stable? In: Unsolved Problems in Mathematical Systems and Control Theory (eds. V. D. Blondel, A. Megretski), pp. 304–308. Princeton University Press, 2004.

- [3] Katok, S., Fuchsian Groups. The University of Chicago Press, Chicago, London, 1992.
- [4] Kazda, A., Convergence in Möbius number systems. *Integers* 9 (2009), 261–279.
- [5] Kazda, A., Symbolic Representations of Compact Spaces. Master’s thesis, Charles University, Prague, 2009.
- [6] Marcus, B., Lind, D., Introduction to Symbolic Dynamics and Coding. Cambridge University Press, 1995.
- [7] Kazda, A., Kůrka, P., Möbius number systems based on interval covers. *Nonlinearity* 23 (2010), 1031–1046.
- [8] Kůrka, P., A symbolic representation of the real Möbius group. *Nonlinearity* 21 (2008), 613–623.
- [9] Kůrka, P., Iterative systems of real Möbius transformations. *Discrete and Continuous Dyn. Syst.* 25(2) (2009), 567–574.
- [10] Kůrka, P., Möbius number systems with sofic subshifts. *Nonlinearity* 22(2) (2009), 437–456.

Received by the editors April 1, 2010. Revised October 6, 2010.