LANGUAGES OF GENERAL INTERVAL EXCHANGE TRANSFORMATIONS

SÉBASTIEN FERENCZI, PASCAL HUBERT, AND LUCA Q. ZAMBONI

ABSTRACT. The languages generated by interval exchange transformations have been characterized by Ferenczi-Zamboni (2008) and Belov-Cernyatev (2010) under some extra conditions on the system. Lifting these conditions leads us to consider successively: natural codings of standard interval exchange transformations, natural codings of affine interval exchange transformations, grouped codings of affine interval exchange transformations, and natural codings of generalized interval exchange transformations. We show that these four classes of languages are strictly increasing, and give necessary and/or sufficient (but not all equally explicit) combinatorial criteria to describe each of them. These work also, mutatis mutandis, for interval exchanges with flips.

The present paper is the long-awaited sequel of two independent papers. One of them is [22], by two of the present authors, who characterize the *languages* of the most studied interval exchange transformations, namely those which are *standard*, have no *flips*, and satisfy a version of Keane's *i.d.o.c. condition*, a condition defined in [31] where it is shown to be stronger (indeed, strictly stronger) than minimality. More far-reaching is [5] by Belov and Chernyat'ev, where the characterization applies to all standard interval exchange transformations, including those with flips. Later, the evolution of the theory gave more and more importance to more general families of interval exchange transformations, raising a need to extend the results of [22] and [5]; as it turned out, a better understanding of the criterion in [22] makes this generalization quite natural.

Interval exchange transformations were originally introduced by Oseledec [43], following an idea of Arnold [1], see also [30]; the unit interval is partitioned into subintervals which are rearranged by piecewise translations. These *standard* exchange transformations have been successively generalized to:

- interval exchange transformations *with flips* [41], where some or all translations are replaced by affine maps of slope -1;
- *affine* interval exchange transformations, where the translations are replaced by affine maps of arbitrary nonzero slope (if some are negative we have a flipped affine interval exchange transformation), thus the Lebesgue measure is not preserved anymore; these appear in the precursor [2], the oldest published reference may be [33];
- *generalized* interval exchange transformations, where the translations are replaced by any continuous monotone nonconstant maps (if some are decreasing we have a flipped generalized interval exchange transformation), see [2] [37];
- *systems of piecewise isometries*, see [25], where the intervals are moved by isometries, but they or their images may be non-disjoint, thus the transformation may be non-injective; these include the *interval translation mappings* of [9].

Date: March 11, 2024.
2010 *Mathematics Subject Classification.* Primary 37B10; Secondary 68R15, 37E05.

These generalizations have seen a recent surge in activity (see [36] [37] [26] [27]) primarily centered on the conjugacy problem between these different classes of maps. In this context, standard and generalized interval exchange transformations are the extreme cases while affine interval exchange transformations constitute a fundamental middle step.

Interval exchange transformations are first return map of linear flows on translation surfaces. This geometrical point of view allows to use modern and powerful techniques, like $\mathrm{SL}(2,\mathbb{R})$ action on moduli spaces and Teichmüller flow, which have been pushed very far these last years (see the surveys [24], [38], [48], [51], [53] and references therein). Many results on interval exchange transformations are obtained in this way (see for instance [50], [4], [16]). Nevertheless it was Rauzy [45] who first suggested using interval exchange transformations as a possible framework for generalizing the well-known interaction between dynamical systems, arithmetics and word combinatorics of [40] and [16]. In these papers, the languages which arise as *natural codings* of minimal rotations by two intervals, aka *Sturmian* languages, are identified with the *uniformly recurrent* languages of *complexity* $n+1$. Rauzy asked which languages arise as natural codings of interval exchange transformations through their defining intervals (see Definition 7), the Sturmian case corresponding to an exchange of two intervals. Several partial answers to Rauzy's question were given for 3-interval exchange transformations in [19] [46] [20]. Then the two fairly general answers of [22] and [5] appeared independently for all numbers of intervals. The characterization in [5] is written in terms of the evolution of *Rauzy graphs*. The one in [22] makes use of two main ingredients: an *order condition* (see Definition 1) which dictates the way *bispecial* words can be *resolved* (see Section 1.1), and the absence of *connections* (see Definition 2).

In the present paper, we aim to study the languages of interval exchange transformations in the most general framework. We prove first that for every transformation which is called an interval exchange, the language satisfies an order condition (Proposition 8). This requires a new definition (Definition 1) in the flipped cases, allowing the orders to change according to the number of flipped letters in each bispecial word. The order condition is intrinsically linked to the one-dimensional nature of interval exchange transformations, and constitutes a strong constraint on the possible ways of resolving the bispecials. Thus those languages arising from interval exchange transformation form a small family, which becomes larger when all flips are allowed. The order condition fails in general for piecewise isometries (see Remark 3), and the characterization of the associated languages is completely unknown.

When the order condition is satisfied, to which extent do we need additional conditions? The following three counter-examples will be described in Section 6 below: they all satisfy an unflipped order condition.

- The *fake Sturmian* language of Example 3 is not a natural coding of a standard interval exchange transformation, but is the natural coding of an affine interval exchange transformation.
- The *skew Sturmian* language of Example 4 is not a natural coding of an affine interval exchange transformation, but is the coding of an affine interval exchange transformation, where the slope may take more than one value inside the coding intervals, we call it a *grouped coding* of an affine interval exchange transformation.
- The *episkew* language of Example 9 is not a natural or grouped coding of an affine interval exchange transformation, but is the natural coding of a generalized interval exchange transformation.

Thus all these four families of languages are different, and we can characterize some cases by simple word-combinatorial conditions. Under an order condition

- by Theorem 19, *all* languages correspond to generalized interval exchange transformations,
- by Theorem 13, *recurrent* languages correspond to standard interval exchange transformations,
- additionally, by Theorem 14, *aperiodic uniformly recurrent languages* correspond to minimal standard interval exchange transformations.

But for the affine case, the characterization of the languages is more complicated, and is very much linked to the extensive work of [33] [35] [11] [10] [15] [36], who build affine interval exchange transformations with wandering intervals semi-conjugated to a given standard interval exchange transformation. The main tool is the study of Birkhoff sums at special points. Thus our criteria to characterize their codings use Birkhoff sums as in these papers. What we get, under an order condition, are necessary or sufficient criteria for a language to correspond to an affine interval exchange transformation by a natural coding (Theorem 20 and Corollary 21). We also get a not very explicit necessary and sufficient criterion for a language to correspond to an affine interval exchange transformation by a grouped coding (Theorem 22).

Then Section 6 is devoted to the examples distinguishing our families of languages. A still open problem is to find an aperiodic grouped coding of an affine interval exchange transformation which is not a natural coding of an affine interval exchange transformation. It is quite possible that these do not exist, and in that case the problems of characterizing the natural and grouped codings of affine interval exchange transformations might get a natural solution (Conjecture 1 and Question 3).

Another problem is to build a language which is a natural coding of a generalized interval exchange transformation, but which is not a grouped coding of an affine interval exchange transformation. This reveals different behaviours for affine and generalized interval exchange transformations. In Example 9, the first property comes from the fact that, by a blow-up process due to Denjoy [18], it is possible to build a generalized interval exchange transformation T, with a wandering interval, semi-conjugate to a given standard interval exchange transformation T'. This is indeed the basis of the proof of our Theorem 19), but if T' is conjugate to a rotation, the Denjoy-Koksma inequality [29] prevents such T from being affine, even after subdividing the continuity intervals. We also prove that T fails to satisfy some regularity condition called class P in [29]. Another way to get such examples, allowing T' to be any non purely periodic interval exchange transformation (not only a rotation) is given in Theorem 24 using Rokhlin towers. By using the results of [39], we have also been able to build exmmples where T' is an interval exchange transformation naturally defined from any one of two famous translation surfaces, the *Eierlegende Wollmilch Sau* and the *Ornithorynque* (Section 6.5); but we could prove only that they are not natural codings of an affine interval exchange transformation.

All our results extend without further difficulty to flipped interval exchange transformation by using flipped order conditions (our counter-examples, however, are only written for the particular case where there are no flips).

1. Languages

1.1. **Usual definitions.** Let $\mathcal{A}$ be a finite set called the *alphabet*, its elements being *letters*. A *word* w of *length* $n = |w|$ is $a_1a_2 \cdots a_n$, with $a_i \in \mathcal{A}$. The *concatenation* of two words w and w'

is denoted by ww'.

By a language L over $\mathcal{A}$ we mean a *factorial extendable language*: a collection of sets $(L_n)_{n\geq 0}$ where the only element of L_0 is the *empty word*, and where each L_n for $n \geq 1$ consists of words of length n, such that for each $v \in L_n$ there exists $a, b \in \mathcal{A}$ with $av, vb \in L_{n+1}$, and each $v \in L_{n+1}$ can be written in the form $v = au = u'b$ with $a, b \in \mathcal{A}$ and $u, u' \in L_n$.

The *complexity function* $p : \mathbb{N} \to \mathbb{N}$ is defined by $p(n) = \#L_n$.

A word $v = v_1...v_r$ *occurs at place* i in a word $w = w_1...w_s$ or a (one-sided or two-sided) infinite sequence (w_n) if $v_1 = w_i$, ...$v_r = w_{i+r-1}$; we say that v is a *factor* of w; it is a *prefix* if it occurs at the initial place k of $w = w_k w_{k+1}...$ and a *suffix* if it occurs at place $l - r + 1$ of $w = ...w_{l-1}w_l$.

A *return word* of a word w is a nonempty word v such that there are exactly two occurrences of w in wv, one as a prefix and one as a suffix.

Let W be a family of words or (one- or two-sided) infinite sequences. Whenever the set L of all the factors of the words or sequences in W is a language (namely, is factorial and extendable), we say that L is *the language generated by* W and denote it by $L(W)$.

A language L is *recurrent* if for each $v \in L$ there exists a nonempty $w \in L$, such that vw ends with v.

A language L is *uniformly recurrent* if for each $v \in L$ there exists n such that v is a factor of each word $w \in L_n$.

A language L is *aperiodic* if for all nonempty words w in L, there exists n such that w^n is not in L.

The *Rauzy graph* G_n of a language L is a directed graph whose vertex set consists of all words of length n of L, with an edge from w to w' whenever $w = av$, $w' = vb$ for letters a and b, and the word avb is in L.

For a word w in L, we call *arrival set of* w and denote by $A(w)$ the set of all letters x such that xw is in L, and call *departure set of* w and denote by $D(w)$ the set of all letters x such that wx is in L.

A word w in L is called *right special*, resp. *left special* if $\#D(w) > 1$, resp. $\#A(w) > 1$. If $w \in L$ is both right special and left special, then w is called *bispecial*. If $\#L_1 > 1$, the empty word ε is bispecial, with $A(\varepsilon) = D(\varepsilon) = L_1$.

A bispecial word w in L is a *weak bispecial* if $\#\{awb \in L, a \in A(w), b \in D(w)\} < \#A(w) + \#D(w) - 1$.

A bispecial word w in L is a *neutral bispecial* if $\#\{awb \in L, a \in A(w), b \in D(w)\} = \#A(w) + \#D(w) - 1$.

A bispecial word w in L is a *strong bispecial* if $\#\{awb \in L, a \in A(w), b \in D(w)\} > \#A(w) + \#D(w) - 1$.

To *resolve* a bispecial word w is to find all words in L of the form awb for letters a and b.

The *symbolic dynamical system* associated to a language L is the two-sided shift S acting on the subset X_L of $\mathcal{A}^{\mathbb{Z}}$ consisting of all bi-infinite sequences such that $x_r \cdots x_{r+s-1} \in L_s$ for each r and s, defined by $(Sx)_n = x_{n+1}$ for all $n \in \mathbb{Z}$.

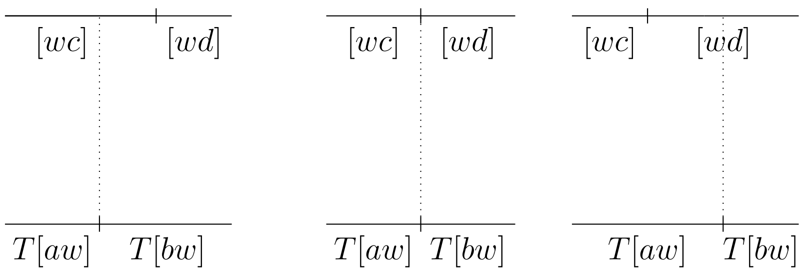

FIGURE 1. Order condition

For a word $w = w_0 \cdots w_t \in L$, the *cylinder* $[w]$ is the set $\{x \in X_L : x_0 = w_0, \dots, x_t = w_t\}$.

In many papers, (X_L, S) is the one-sided shift the subset X_L of $\mathcal{A}^{\mathbb{N}}$ consisting of all infinite sequences such that $x_r \cdots x_{r+s-1}$ is in L_s for every r, s. In the present paper, we shall use two-sided sequences $x \in X_L$, but also their *infinite suffixes* $(x_n, n \geq k)$ and *infinite prefixes* $(x_n, n \leq k)$.

An *orbit* of S is a set $\{S^n x, x \in X_L\}$.

1.2. **Languages with order conditions.** The key property we shalll use, for which we choose the general name of *order condition*, with some ad hoc variants, appears in a not completely explicit way in Theorem 1 of [22]; it is stated under its present form, though only in a particular case, in [17].

In hope of presenting it intuitively, we anticipate on Proposition 8 below, in a simple case. We take a bispecial word w in the natural coding of a standard interval exchange (Definitions 5 and 7), such that w can be preceded by two possible letters, a and b, and followed by two possible letters, c and d, and look at how it can be resolved. Then the cylinder $[w]$ is an interval, partitioned into two subintervals by the cylinders $[wc]$ and $[wd]$, and into another pair of subintervals by $T[aw]$ and $T[bw]$. Moreover, $[wc]$ is to the left of $[wd]$ whenever I_c is to the left of I_d, which we denote by $c <_D d$; $T[aw]$ is to the left of $T[bw]$ whenever TI_a is to the left of TI_b, which we denote by $a <_A b$. Then the one-dimensional nature of the system implies that there are only three possible ways to resolve w, as shown in Figure 1: the word bwc is in L iff the interval $[wc]$ is strictly longer than the interval $T[aw]$, and awd is in L iff the interval $T[aw]$ is strictly longer than the interval $[wc]$. Thus, if $a <_A b$ and $c <_D d$, awd and bwc cannot be both in L, which is equivalent to the property used in Definition 1 below. Note that this involves the *strict* orders, if $a \leq_A b$ and $c \leq_D d$, we may have awd and bwc both in L.

In this example the orders do not depend on w, but this will not hold when T has flips. Suppose that we are in the same situation, but w is the one letter word x and I_x is flipped, then $T[wa]$ is to the left of $T[wb]$ whenever TI_a is to the left of TI_b, but, as T is decreasing on I_x, $[wc]$ is to the left of $[wd)$ whenever I_c is to the right of I_d. We denote this last property as $c <_{D,w} d$, and the sequel of the reasoning works in the same way. In the cases $w = x$ where I_x is not flipped, or if $w = yz$ for letters such that both I_y and I_z are flipped, both orders we consider are the same as in the unflipped case. So in general we have to allow the D order to depend on w, and in the framework of languages we allow that also for the A order.

Definition 1. *A language L on an alphabet $\mathcal{A}$ satisfies a* local order condition *if, for each bispecial word w, there exist two strict total orders on $\mathcal{A}$, denoted by $<_{A,w}$ and $<_{D,w}$, such that whenever awc and bwd are in L with letters $a \neq b$ and $c \neq d$, then $a <_{A,w} b$ if and only if $c <_{D,w} d$.*

*A language L on an alphabet $\mathcal{A}$ satisfies an $\mathcal{F}$-*flipped order condition *for a (possibly empty) subset $\mathcal{F}$ of $\mathcal{A}$ if there exist two strict total orders on $\mathcal{A}$, denoted by $<_A$ and $<_D$, such that L satisfies a local order condition where*

- *the order $<_{A,w}$ is the same as $<_A$ for all w,*
- *$<_{D,w}$ is the same as $<_D$ when the number of occurrences in w of letters belonging to $\mathcal{F}$ is even,*
- *$<_{D,w}$ is the reverse order of $<_D$ when the number of occurrences in w of letters belonging to $\mathcal{F}$ is odd.*

In Definition 2 below, in the framework of languages, the notion of *connection* generalizes the middle case of Figure 1 above, where neither awd nor bwc is in L: in that case, for the corresponding interval exchange, $T^k\beta_i = \gamma_j$ for some i, j, $k > 0$, two singularities meet at the same orbit.

Definition 2. *If a language L on an alphabet $\mathcal{A}$ satisfies a local order condition, a bispecial word w has a* connection *if there are letters $a <_{A,w} a'$, consecutive in the order $<_{A,w}$, letters $b <_{D,w} b'$, consecutive in the order $<_{D,w}$, such that awb and $a'wb'$ are in L, and neither awb' nor $a'wb$ is in L.*

In this section, we state general combinatorial properties of languages satisfying some order conditions, or weaker properties, to be used in the next sections. Note that some of these use the same ideas as Section 2.2 of [21], though the present context is both more combinatorial and more general. Further properties with a more dynamical flavor will be studied at the beginning of Sections 4 and 5.

We define first a new property for bispecial words.

Definition 3. *In a language L, a* locally strong bispecial word *is a bispecial word w such that there exist nonempty subsets $A' \subset A(w)$, $D' \subset D(w)$ such that $\#\{awb \in L, a \in A', b \in D'\} > \#A' + \#D' - 1$.*

Lemma 1. *A language L which satisfies a local order condition contains no locally strong bispecial word, and thus no strong bispecial word.*

Proof

Let w, A', D' be as in Definition 2. If $\#A'$ or $\#D'$ is 1, then the result is immediate. Assume $\#A' = \#D' = 2$ and put $A' = \{a_1, a_2\}$, $D' = \{b_1, b_2\}$, $a_1 <_A a_2$, $b_1 <_D b_2$. If a_1wb_2 and a_2wb_1 both belong to L, this contradicts the local order condition. Now suppose we have proved the result for all A' and D' such that $\#A' \leq p$ and $\#D' \leq q$; take $A' = \{a_1 <_{A,w} \cdots <_{A,w} a_p <_{A,w} a_{p+1}\}$, $D' = \{b_1 <_{D,w} \cdots <_{D,w} b_q\}$; then, let r be the largest integer between 1 and q such that $a_pwb_r \in L$. By the local order condition, $a_iwb_j \in L$ for some $1 \leq i \leq p$ only if $j \leq r$, thus by the induction hypothesis there are at most $p + r - 1$ possible a_iwb_j, $1 \leq i \leq p$, $1 \leq j \leq q$. Again by the local order condition, if $a_{p+1}wb_j \in L$, $1 \leq j \leq q$, then $r \leq j \leq q$. Thus the number of possible awb belonging to L with $a \in A'$ and $b \in D'$ is at most $p + r - 1 + q + 1$ and we have extended the induction hypothesis to $p + 1$ and q. A similar reasoning extends it to p and $q + 1$, thus the result holds for all possible A' and D'. □

For languages where each word has at most two right (resp. left) extensions, the absence of strong bispecial words, the absence of locally strong bispecial words, and a local order condition are all equivalent. In the general case, it is easy to find bispecials which are locally strong but not strong (suppose for example that the possible xwy are awa, awb, bwa, bwb, cwc), and a local order condition is stricter than the absence of locally strong bispecials.

Example 1. *Suppose L is a language whose words of length 2 are ac, ad, ba, bc, cb, cc, da. Then the empty word is not a locally strong bispecial, yet L does not satisfy any local order condition.*

Indeed, assume to the contrary that ε satisfies the order condition with respect to $<_{A,\varepsilon}$ and $<_{D,\varepsilon}$. By possibly reversing both orders, we can assume $c <_{A,\varepsilon} a$. Since $\{cb, ac, cc, ad\} \subset L$ this implies $b <_{D,\varepsilon} c <_{D,\varepsilon} d$. This in turn implies $c <_{A,\varepsilon} b <_{A,\varepsilon} a$ (since $\{cb, bc, ad\} \subset L$) which in turn implies $c <_{D,\varepsilon} a <_{D,\varepsilon} d$ (since $\{cc, ba, ad\} \subset L$). Also $b <_{A,\varepsilon} d <_{A,\varepsilon} a$ because $\{bc, da, ad\} \subset L$. Finally, as ac and da are in L, we get $a <_{D,\varepsilon} c$ which is a contradiction.

We can then choose L_3 to be made with acc, ada, bac, bad, bcb, bcc, cba, cbc, ccb, dac, where each word has at most two left (resp. right) extensions and continue by resolving the bispecials so that they are all neutral. We get a language without locally strong bispecials but not satisfying any local order condition.

Lemma 2. *A language L which has no strong bispecial factor has a finite number of weak bispecial factors.*

Proof

Let $s(n) = p(n+1) - p(n)$. By Theorem 4.5.4 of [13], see also [12], if L has only weak or neutral bispecials $s(n+1) \leq s(n)$ for all n, and, if additionally L has a weak bispecial of length n, $s(n+1) < s(n)$. This cannot happen infinitely many times as $p(n)$ is a growing function. □

Another notion aiming at generalizing the properties of some natural codings of interval exchange transformations is a *dendric* language, which is defined in [7] (under the name of *tree set*): it turns out that a language is dendric iff it has neither locally strong bispecial words nor weak bispecial words, thus there is no inclusion relation between this family and the various order conditions. However, by Lemmas 1 and 2, a language satisfying a local order condition is *ultimately dendric*.

Lemma 3. *If a language L satisfies a local order condition, a word w is a weak bispecial iff it has a connection.*

Proof

Let $A(w)$, ordered by $<_{A,w}$, be $\{a_1, \dots, a_p\}$, and $D(w)$, ordered by $<_{D,w}$, be $\{b_1, \dots, b_q\}$. If there is no connection, then each $D(a_k w) \cap D(a_{k+1} w)$ has at least one element, and indeed exactly one element by Lemma 1: thus the number of possible awb, which is $\sum_{k=1}^{p} \#D(a_k w)$, is $q + p - 1$, and w is a neutral bispecial.

If there is a connection, then for some l $D(a_l w) \cap D(a_{l+1} w)$ is empty. By the local order condition, if b_j is the maximal element (for $<_{D,w}$) of $D(a_l w)$, or equivalently of $D(a_1 w) \cup \cdots D(a_l w)$, then b_{j+1} is the minimal element (for $<_{D,w}$) of $D(a_{l+1} w) \cup \cdots \cup D(a_p w)$. By Lemma 1 applied first to $\{a_1, \dots, a_l\}$ and $\{b_1, \dots, b_j\}$, then to $\{a_{l+1}, \dots, a_p\}$ and $\{b_{j+1}, \dots, b_q\}$, we get that the number of possible awb is at most $j + l - 1 + p - l + q - j - 1 = p + q - 2$, thus w is a weak bispecial. □

Corollary 4. *A language satisfying a local order condition has complexity $p(n) = kn + l$ for all n large enough and with $0 \leq k \leq \#\mathcal{A} - 1$. Moreover, $k = \#\mathcal{A} - 1$ if and only if L has no connection, and in that case $l = 1$.*

Proof

This comes from Lemmas 1, 2, 3, and Theorem 4.5.4 of [13]. □

The following result is well known and could be deduced from Section 3.3 of [12], or Theorem 4.5.4 of [13].

Lemma 5. *If a language L has no strong bispecial word, the left special words are the prefixes of a finite number of infinite suffixes of sequences of X_L, the right special words are the suffixes of a finite number of infinite prefixes of sequences of X_L.*

Proof

Suppose v is left special, and va and vb are left special for letters $a \neq b$, then both $A(va)$ and $A(vb)$ are strictly included in $A(v)$. Indeed, otherwise, the possible cvd in L contain at least, for example cva for all c in $A(v)$, $c'vb$ for at least two different c' in $A(v)$, and some c_ivd_i for each d_i in $D(v) \setminus \{a, b\}$, and thus v is a strong bispecial word.

Starting from each left special letter e, we extend it to the right by one letter at a time; each time two different extensions of the same word are left special, the arrival sets decrease strictly; thus, after a finite number of cases where this happens, each new left special extension has the previous one as a prefix, and all these are prefixes of a finite number of sequences $ex_1...x_n...$, which are infinite suffixes of sequences of X_L.

And a similar reasoning works for the right specials. □

Lemma 6. *Let L be a language satisfying a local order condition. If G_n is connected (as a nonoriented graph) but G_{n+1} is not connected, then some bispecial word in L_n has a connection.*

Proof

As G_n is connected, if all pairs of vertices of G_{n+1} which correspond to two consecutive edges of G_n belong to the same connected component of G_{n+1}, G_{n+1} is connected. Thus assume that for some $w \in L_n$ and letters a and b, aw and wb are not in the same connected component of G_{n+1}. Then, in particular awb is not in L. But as every word is extendable to the left and right, there exist a' and b' such that $a'wb$ and awb' are in L, thus $a \neq a'$, $b \neq b'$, and w is bispecial.

Suppose w has no connection. If a_1w is in some connected component U of G_{n+1}, so is wx for every x in $D(a_1w)$, hence also wx is in U for one element of $D(a_2w)$ and so on, hence all the a_iw and all wx for x in $D(w)$ are in U, and this contradicts the fact that G_{n+1} is not connected. □

Note that the terminology inherited from two different fields is somewhat counter-intuitive; a *connection*, named for geometrical reasons as explained before Definition 2, is necessary to *disconnect* the Rauzy graphs. But it is not a sufficient condition, see examples in [21].

An order condition on L allows us to define an order on the set X_L, which will be used in Section 4 below. It is well-defined because of the order condition.

Definition 4. *Suppose L satisfies an $\mathcal{F}$-flipped order condition. Let x and y be in X_L. We define an order relation by $x < y$ if one or several of the following assertions is satisfied:*

- $x_0 <_D y_0$,
- $x_i = y_i$ *for* $0 \leq i \leq k$, $k > 0$, $x_{k+1} <_D y_{k+1}$, *an even number of the* $x_i, 0 \leq i \leq k$, *is in* $\mathcal{F}$,
- $x_i = y_i$ *for* $0 \leq i \leq k$, $k > 0$, $x_{k+1} >_D y_{k+1}$, *an odd number of the* $x_i, 0 \leq i \leq k$, *is in* $\mathcal{F}$,
- $x_{-1} <_A y_{-1}$,

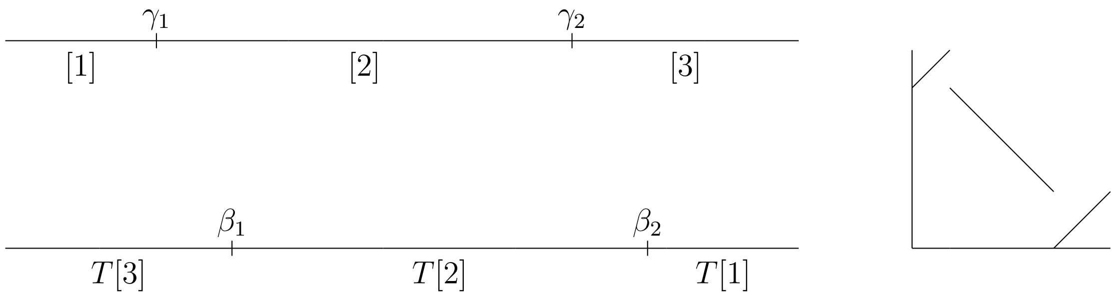

FIGURE 2. A standard 3-interval exchange with one flip

- $x_i = y_i$ *for* $k \leq i \leq -1$, $k < -1$, $x_{k-1} <_A y_{k-1}$, *an even number of the* $x_i, k \leq i \leq -1$, *is in* $\mathcal{F}$,
- $x_i = y_i$ *for* $k \leq i \leq -1$, $k < -1$, $x_{k-1} >_D y_{k-1}$, *an odd number of the* $x_i, k \leq i \leq -1$, *is in* $\mathcal{F}$.

Lemma 7. *The order on* X_L *is total, and every subset* Y *of* X_L *admits an upper bound and a lower bound in* X_L.

Proof
For each n we can find a word $y_m^{(n)}, -n \leq m \leq n$, such that $y^{(n)}$ is the largest, resp. smallest, element of the set $\{y_m, -n \leq m \leq n, y \in Y\}$ for the order defined on such words by Definition 4. Then $y_m^{(n+1)} = y_m^{(n)}$ for $-n \leq m \leq n$ and by definition of X_L there is a sequence y in X_L such that $y^m = y_m^{(n)}$ for $-n \leq m \leq n$, which will be the desired upper, resp. lower bound. □

2. INTERVAL EXCHANGE TRANSFORMATIONS

A generalized interval exchange transformation is basically a bijection of the unit interval to itself, which is continuous and monotone on a finite number of subintervals which partition the unit interval. However, such maps can be built, by using semi-open intervals of continuity, only when T is increasing on all these intervals. In the *flipped* cases, T will have to stay undefined on a finite set of points.

Definition 5. *Let* $\mathcal{A}$ *be a finite alphabet,* $\mathcal{F}$ *a possibly empty subset of* $\mathcal{A}$. *An* $\mathcal{F}$-flipped generalized interval exchange transformation *is a map* T *defined on a disjoint union of* open *intervals* I_e, $e \in \mathcal{A}$, *such that the union of their closures is* $[0, 1]$, *continuous and (strictly) increasing on each* I_e, $e \in \mathcal{F}^c$, *continuous and (strictly) decreasing on each* I_e, $e \in \mathcal{F}$, *and such that the* TI_e, $e \in \mathcal{A}$, *are disjoint open intervals and the union of their closures is* $[0, 1]$.

The I_e, *indexed in* $\mathcal{A}$, *are called the* defining intervals *of* T.

If the restriction of T *to each* I_e *is an affine map,* T *is an* $\mathcal{F}$-flipped affine *interval exchange transformation.*

If the restriction of T *to each* I_e *is an affine map of slope* ± 1, T *is an* $\mathcal{F}$-flipped standard *interval exchange transformation.*

The endpoints of the I_e, *resp.* TI_e, *excluding* 0 *and* 1, *will be denoted by* γ_i, *resp.* β_j, *for* i, *resp.* j, *taking* $\#\mathcal{A} - 1$ *values.*

In this definition, the defining intervals I_e are not necessarily the intervals of continuity of T, as it may happen that I_e and I_f are adjacent and sent to adjacent intervals with the same flips in the correct order, thus T may be extended to a continuous map on their union. Thus in the present paper an interval exchange transformation T is always supposed to be given *together with its defining intervals* as keeping the same T but changing the defining intervals would change the coding.

In all cases, T is undefined on a finite number of points, namely the γ_i, 0 and 1. In the unflipped cases, T can be extended to $[0, 1)$, resp. $(0, 1]$, by including in each I_e its left (resp. right) endpoint. Also in this case, as did Keane in Section 5 of [31] for unflipped standard interval exchange transformations and Arnoux in Chapter 1 of [2] (Definition 1.4 and Lemma 1.5) for unflipped generalized interval exchange transformations, by carefully doubling the endpoints and their orbits, it is possible to extend T to an homeomorphism on the union of the unit interval and a Cantor set.

Throughout this paper, all defining intervals will be nonempty open intervals. In the unflipped cases, all our results and proofs remain valid if we define T with semi-open intervals, open on the right (resp. left), or define T on Keane's set (mentioned above) and use closed intervals.

Definition 6. *Let I be the subset (with countable complement) of $(0, 1)$ where all T^n, $n \in \mathbb{N}$, are defined. Every statement concerning T^n will be tacitly assumed to be valid if it is valid on I.*

T is minimal *if every orbit is dense.*

A wandering interval *is an interval J for which T^nJ is disjoint from J for all $n > 0$.*

2.1. **Codings.**

Definition 7. *For a generalized $\mathcal{F}$-flipped interval exchange transformation T, coded by $\mathcal{A}$, its* natural coding *is the language $L(T)$ generated by all the* trajectories*, namely the sequences $(x_n, n \in \mathbb{Z}) \in \mathcal{A}^{\mathbb{Z}}$ where $x_n = e$ if T^nx falls into I_e, $e \in \mathcal{A}$.*

Thus we can look at the symbolic system associated to $L(T)$. Note that the set $X_{L(T)}$ is the closure in $\mathcal{A}^{\mathbb{Z}}$ of the set of trajectories, for the product topology defined by the discrete topology on $\mathcal{A}$.

Example 2. *A* Sturmian language *is the natural coding of the unflipped standard interval exchange transformation T sending $(0, 1-\alpha)$ to $(\alpha, 1)$ and $(1-\alpha, 1)$ to $(0, \alpha)$ for α irrational; T is conjugate to a rotation of angle α on the 1-torus.*

Remark 1. *The elements of $X_{L(T)}$ which are not actual trajectories of points under T are called* improper *trajectories. For each improper trajectory u in $X_{L(T)}$, either there exists an endpoint γ of a defining interval I_e, and an integer $l \leq 0$ such that u is the limit of the trajectories of x when x tends to $T^l\gamma$, either from the left or from the right, or there exists an endpoint β of a TI_e, and an integer $l \geq 0$ such that u is the limit of the trajectories of x when x tends to $T^l\beta$, either from the left or from the right.*

In the unflipped cases, if we define T on Keane's set (as mentioned after Definition 5), each element of $X_{L(T)}$ is an actual trajectory for T; if we use semi-open intervals, each element of $X_{L(T)}$ is an actual trajectory, either for T defined and coded with semi-open intervals $[a, b)$, or for T defined and coded with semi-open intervals $(a, b]$, in the same way as Sturmian trajectories in [40].

We shall also consider slightly more general codings, by merging into intervals $\tilde{I}_e$ some adjacent intervals I_e whose images by T are also adjacent, and flipped in the same way. This is equivalent to taking the natural coding of another interval exchange transformation $\tilde{T}$, but when T is affine, if we define $\tilde{T}$ by the intervals $\tilde{I}_e$ it will not necessarily be affine by our definition, as the slope is not constant on its defining intervals, see Example 4 below. Thus we define

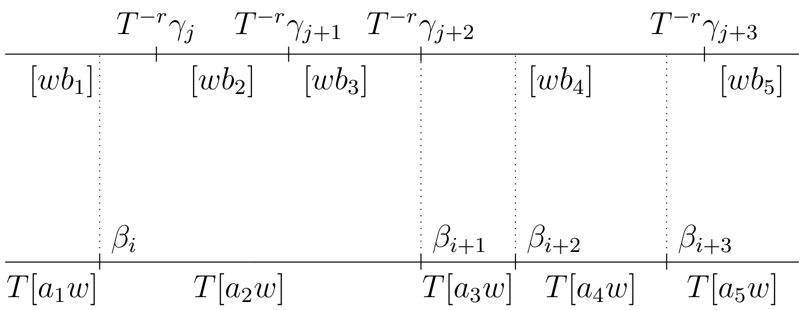

FIGURE 3. A general bispecial interval

Definition 8. *A language L is a* grouped coding *of an $\mathcal{F}$-flipped affine interval exchange transformation T if there exist intervals $\tilde{I}_e$, $e \in \tilde{\mathcal{A}}$ such that*

- *each $\tilde{I}_e$ is an open interval, and a disjoint union (plus intermediate endpoints) of defining intervals of T,*
- *T can be extended by continuity to a continuous monotone map on each $\tilde{I}_e$,*
- *L is the coding of T by the $\tilde{I}_e$, that is the language generated by the trajectories $(x_n, n \in \mathbb{Z}) \in \tilde{\mathcal{A}}^{\mathbb{Z}}$ where $x_n = e$ if $T^n x$ falls into $\tilde{I}_e$, $e \in \tilde{\mathcal{A}}$.*

2.2. **Orders.**

Definition 9. *A generalized interval exchange transformation defines two orders on $\mathcal{A}$:*

- *$e <_D f$ whenever the interval I_e is strictly to the left of the interval I_f,*
- *$e <_A f$ whenever the interval TI_e is strictly to the left of the interval TI_f.*

These orders correspond to the two permutations used by Kerckhoff [32] to define standard interval exchange transformations: the unit interval is partitioned into semi-open intervals which are numbered from 1 to k, ordered according to a permutation π_0 and then rearranged according to another permutation π_1; in more classical definitions, there is only one permutation π, which corresponds to π_1 while $\pi_0 = Id$; note that sometimes the orderings are by π_0^{-1} and π_1^{-1}.

Proposition 8. *Let T be a generalized $\mathcal{F}$-flipped interval exchange transformation, for $\mathcal{F}$ a subset of an alphabet $\mathcal{A}$. Then its language $L(T)$ satisfies an $\mathcal{F}$-flipped order condition.*

Proof
Let $<_A$ and $<_D$ be the orders from Definition 9.

Let $w = w_0 \cdots w_{r-1}$ be a bispecial word in $L(T)$. The cylinder $[w]$ is the interval $I_{w_0} \cap T^{-1}I_{w_1} \cap \cdots \cap T^{-r+1}I_{w_{r-1}}$; after deleting a finite number of points, this interval is partitioned into the $T[aw] = TI_a \cap [w]$, $a \in A(w)$. As $T[aw] \subset TI_a$ for all $a \in A(w)$, $T[aw]$ is (strictly) left of $T[a'w]$ whenever $a <_A a'$.

Similarly, after deleting a finite number of points, $[w]$ is also partitioned into the $[wb] = I_{w_0} \cap T^{-1}I_{w_1} \cap \cdots \cap T^{-r+1}I_{w_{r-1}} \cap T^{-r}I_b \subset T^{-1}I_{w_1} \cap \cdots \cap T^{-r+1}I_{w_{r-1}} \cap T^{-r}I_b$.

The order (from left to right) between two intervals $[b] = I_b$ and $[b'] = I_{b'}$ is the order $<_D$. The order between $TI_{w_{r-1}} \cap I_b \subset I_b$ and $TI_{w_{r-1}} \cap I_{b'} \subset I_{b'}$ is also the order $<_D$. The order between $I_{w_{r-1}} \cap T^{-1}I_b$ and $I_{w_{r-1}} \cap T^{-1}I_{b'}$ is the same order if T is increasing on $I_{w_{r-1}}$ or equivalently $w_{r-1} \in \mathcal{F}^c$, the opposite order otherwise. Similarly the order between $TI_{w_{r-2}} \cap I_{w_{r-1}} \cap T^{-1}I_b \subset$

$I_{w_{r-1}} \cap T^{-1}I_b$ and $TI_{w_{r-2}} \cap I_{w_{r-1}} \cap T^{-1}I_{b'} \subset I_{w_{r-1}} \cap T^{-1}I_{b'}$ is the same as the last one, thus the order between $I_{w_{r-2}} \cap T^{-1}I_{w_{r-1}} \cap T^{-2}I_b$ and $I_{w_{r-2}} \cap T^{-1}I_{w_{r-1}} \cap T^{-2}I_{b'}$ is the same order as the last one if T is increasing on $I_{w_{r-2}}$ or equivalently $w_{r-2} \in \mathcal{F}^c$, the opposite order otherwise. And so on, finally we get that the order between $[wb]$ and $[wb']$ is exactly the order $<_{D,w}$ defined in the second item of Definition 1.

Thus the intervals $T[aw]$ are ordered from left to right by the order $<_A$, while the $[wb]$ are ordered from left to right by the order $<_{D,w}$. The word awc exists whenever $T[aw] \cap [wc]$ is nonempty, similarly for bwd: as in Figure 3, we get that if awc and bwd exist, $a \neq c$, $b \neq d$, then $a <_A c$ iff $b <_{D,w} d$. Note that this is true also for the empty word, which is bispecial and for which no letter is in $\mathcal{F}$. Thus $L(T)$ does satisfy the $\mathcal{F}$-flipped order condition defined by $<_A$ and $<_D$. □

Remark 2. *By the same reasoning as in Proposition 8, we check that if* $(x_n, n \in \mathbb{Z})$ *in* $X_{L(T)}$ *is the (actual or improper) trajectory of some point* x*,* $(y_n, n \in \mathbb{Z})$ *in* $X_{L(T)}$ *is the (actual or improper) trajectory of some point* y*,* $(x_n, n \in \mathbb{Z}) \leq (y_n, n \in \mathbb{Z})$ *for the order of Definition 4 if and only if* $x \leq y$ *in the natural order on* $(0, 1)$*.*

Remark 3. *The language of an* interval translation mapping [9] *does not necessarily satisfy an* $\mathcal{F}$*-flipped order condition. It is possible that* TI_a *intersects* TI_b *for* $a \neq b$*, and, if* $TI_a \cap TI_b$ *intersects both* I_c *and* I_d*,* $c \neq d$*, then the empty word is a locally strong bispecial, which contradicts the local order condition by Lemma 1.*

Question 1. *How to characterize the natural codings of systems of piecewise isometries, resp. interval translation mappings ?*

Even for the latter family, almost nothing is known. Their complexity is the object of a question of Boshernitzan; it was proved to be linear in some particular cases by Cassaigne and Nicolas [13].

3. STANDARD INTERVAL EXCHANGE TRANSFORMATIONS

We continue the study of languages satisfying order conditions, or weaker properties, by looking at *recurrence*. Recurrence and uniform recurrence of a language are defined in Section 1.1 above. A different concept is the recurrence of infinite sequences, which is a property of the associated symbolic dynamical system.

Definition 10. *A bi-infinite sequence* x *in* $\mathcal{A}^{\mathbb{Z}}$*, or an infinite suffix of it, is* right recurrent *if any factor* w *of* x *is equal to* $x_{n_k}...x_{n_k+t}$ *for a sequence* n_k *tending to* $+\infty$*.*

A bi-infinite sequence x *in* $\mathcal{A}^{\mathbb{Z}}$*, or an infinite prefix of it, is* left recurrent *if any factor* w *of* x *is equal to* $x_{n_k}...x_{n_k+t}$ *for a sequence* n_k *tending to* $-\infty$*.*

A bi-infinite sequence x *in* $\mathcal{A}^{\mathbb{Z}}$ *is* recurrent *if it is both left and right recurrent.*

Lemma 9. *Suppose* L *has no strong bispecial word.*

If an infinite suffix of a sequence in X_L *is right recurrent, it generates a uniformly recurrent language.*

If an infinite prefix of a sequence in X_L *is left recurrent, it generates a uniformly recurrent language.*

If a sequence in X_L *is left or right recurrent, it is recurrent and generates a uniformly recurrent language.*

Proof

Let w be any word in L, we look at the possible return words of w, denoted by $v = v_1...v_n$. If $wv_1...v_k$ (or w if $k = 0$) is not right special, there is only one possible choice for v_{k+1}; if it is right special, there are at most $\#\mathcal{A}$ choices for v_{k+1}. In this last case, by Lemma 5 $wv_1...v_k$ (or w if $k = 0$) must be a suffix of one of K infinite prefixes of sequences in X_L. Thus if $v_1...v_n$ is a return word of w, there are at most K values of $0 \leq j < n$ such that $wv_1...v_j$ (or w if $j = 0$) is right special; otherwise, for some $0 \leq k < l < n$, $wv_1...v_k$ (or w if $k = 0$) and $wv_1....v_l$ are different suffixes of the same infinite prefix, thus there is an occurrence of w in $wv_1...v_l$ which is not the initial one, and thus an occurrence of w in $wv_1...v_n$ which is not the initial or final one, and this contradicts the definition of a return word in Section 1.1 above. Hence there are at most $K\#\mathcal{A}$ return words of w.

If an infinite suffix x is right recurrent, as any prefix of x occurs further to the right, the set of its factors is a language L'. For each factor w of x, we see it infinitely many times, with at most $K\#\mathcal{A}$ ways of going from one occurrence to the next one. Thus we see w in x infinitely many times with bounded gaps, thus w occurs in any long enough factor of x, and L' is uniformly recurrent. And the same works for left recurrence.

Now, suppose for example $x \in X_L$ is right recurrent. Then it is also left recurrent: otherwise there exists a factor w of x which does not occur in $(x_n, n \leq k)$; thus there exists arbitrarily long factors $v_{(m)}$ of x with no occurrence of w; but then any infinite suffix of x contains every $v_{(m)}$, and thus cannot generate an uniformly recurrent language. Hence any factor w occurs in x at infinitely many places to the right and to the left, and with bounded gaps, thus the language generated by x is uniformly recurrent. □

It will result from Proposition 12 that, under a local order condition, the recurrence of L implies the recurrence of every sequence in X_L; here we prove first a weaker statement.

Lemma 10. *Suppose L has no locally strong bispecial, and is recurrent. Then if $w'w^n$ is in L for all $n \in \mathbb{N}$, with $|w'| \leq |w|$, then w' is a suffix of w, and if $w^n w'$ is in L for all n, then w' is a prefix of w.*

Proof

Suppose first $w'w$ is a non-initial factor of some $w'w^m$; then either w' occurs as a suffix of w, or we have $w = w_1...w_l = w_{d+1}...w_l w_1...w_d$ for some $0 < d < l$; in the latter case, we get that $w = vv...v$ for a factor v whose length is the largest common divisor of d and l, and that w' is a factor of ww ending at index d or $l - d$, and this factor is also a suffix of w.

Suppose $w'w^n$ is in L for all n, and w' is not a suffix of w. By recurrence, for each $n \geq 1$ $w'w^n$ must occur after $w'w^n$. As it cannot occur as a non-initial factor of some $w'w^N$, then for each $n \geq 1$ there is a right special word which is a prefix of some $w'w^N$ and has $w'w^n$ as a prefix; for infinitely many p_n, it will be of the form $w'w^{p_n}u$, for a fixed prefix u of w, and followed by two fixed letters b and b'. If u' is the longest common suffix of w and w', we get four different words of the form $xu'w^{p_n}uy$, and thus a locally strong bispecial. And similarly for the second assumption. □

Lemma 11. *A finite or countable union of uniformly recurrent languages satisfies the* measure condition*: there exists an invariant probability measure μ on the symbolic system (X_L, S) associated to L such that $\mu[w] > 0$ for any $w \in L$.*

Proof
Assume L is a uniformly recurrent language and let μ be any invariant probability measure on the symbolic system (X_L, S) associated to L. Such a measure μ can be constructed as in [8]. Fix w in L. Then, for each positive integer N there is at least one word w' of length N with $\mu[w'] > 0$, and, as by uniform recurrence w is a factor of w' if N is large enough, $[w']$ is included in $\cup_{i=0}^{N} S^i[w]$ from which it follows that $\mu[w] > 0$.

If L is a finite or countable union of uniformly recurrent languages L_i, then each L_i satisfies the measure condition for some measure μ_i, and hence L will satisfy the measure condition for any average of the measures μ_i. □

Proposition 12. *Let L be a recurrent language satisfying a local order condition. Then L is a finite union of uniformly recurrent languages.*

Proof
The Rauzy graph G_1 has a finite number of connected components. The language L restricted to any connected component of any of its Rauzy graphs G_n still satisfies the local order condition. We apply repeatedly Lemma 6 to any of these languages: if G_{n+1} has more connected components than G_n, then there exists a connected component C_n of G_n such that the Rauzy graph C_{n+1} made with the edges of C_n is not connected. Thus there is an edge missing in C_{n+1} as in Lemma 3, and it cannot be elsewhere in G_{n+1}, thus this edge is missing in G_{n+1} and there is a weak bispecial word in L by the reasoning of Lemma 3. By Lemma 2 this happens a finite number of times, thus, for all n, G_n has at most N connected components, and L is a union of N languages for which all Rauzy graphs are connected.

Suppose there is a word w in L for which no left extension is left special. Then, before $w = w_0 \cdots w_l$, we can see only one sequence of letters, denoted by w_n, $n < 0$. As L is recurrent, w must occur as some $w_s...w_t$, $s < t < 0$, thus the sequence $(w_n, n < 0)$ begins with a periodic infinite prefix $\ldots vvv$. For any m large enough, this gives a loop in the Rauzy graph G_m; it is impossible both to enter this loop from any other point in G_n and to exit from it to any other point in G_n, otherwise this contradicts Lemma 10. Thus the language generated by $\ldots vvv$ corresponds to a full connected component of G_m for all m large enough, thus there are at most N such languages. Thus L is the union of at most N languages satisfying the local order condition, where each language either is generated by the v^n, $n \geq 0$, for a word v, or contains only words which can be extended to a left special word.

Let L' be one of these last languages, if they exist. Let G'_n be its Rauzy graphs. Because there is no strong bispecial word in L', by Lemma 5 all left special words in L' are prefixes of K infinite suffixes W_i; as every word of L' can be extended to a left special, all the words of L' are the factors of the W_i.

Let w be a word of L': w must be a suffix of infinitely many words of L', because every word of L' can be extended infinitely many times to the left; thus w occurs infinitely often in at least one W_i. Hence all the prefixes of W_i occur infinitely often in at least one W_j, and if $j \neq i$, then we can drop W_i and use only the other W_j to generate L'. Thus, after a renumbering of the W_i, L' is the union of the languages $L(W_i)$, $1 \leq i \leq q \leq K$, where the W_i are right recurrent infinite suffixes.

Then by Lemma 9 each $L(W_i)$ is uniformly recurrent, while any of language generated by the v^n, $n \geq 0$, is also uniformly recurrent. □

We can now prove our first main theorem, which might also be deduced from rewriting the criterion of [5].

Theorem 13. *For a language L on an alphabet $\mathcal{A}$ and $\mathcal{F} \subset \mathcal{A}$, the following are equivalent:*

- (i) *L satisfies an $\mathcal{F}$-flipped order condition and is recurrent;*
- (ii) *L is the language of a standard $\mathcal{F}$-flipped interval exchange transformation;*
- (iii) *L is the language of a generalized $\mathcal{F}$-flipped interval exchange transformation without wandering intervals.*

Proof

If L satisfies (i), by Proposition 12 and Lemma 11, L satisfies the measure condition of Lemma 11. We prove now that it implies (ii); this is indeed the same reasoning as in Proposition 4 of [22], modified to take into account flips and connections.

Let μ be as in the measure condition. Let $\mathcal{A} = L_1$, and let T be the standard $\mathcal{F}$-flipped interval exchange transformation defined by intervals I_e, $e \in \mathcal{A}$, of respective length $\mu[e]$ ordered from left to right by the order $<_D$, while the TI_e, $e \in \mathcal{A}$, of respective length $\mu[e]$, are ordered from left to right by the order $<_A$, and T has slope 1 on the I_e, $e \in \mathcal{F}^c$, and slope -1 on the I_e, $e \in \mathcal{F}$.

Let $L' = L(T)$, μ' the shift-invariant measure on $X_{L'}$ defined by the Lebesgue measure on $(0, 1)$.

We prove inductively on n that $w \in L_n$ if and only if $w \in L'_n$, and for these words $\mu[w] = \mu'[w]$. This is true for $n = 1$ by our choice of T. We denote by A_L, $A_{L'}$, D_L, $D_{L'}$ the arrival and departure sets in L and L'.

Let the hypothesis be proved for n, and take a word w in L of length $n - 1$ (possibly the empty one if $n = 1$). Because of the induction hypothesis, $A_L(w) = A_{L'}(w)$, $D_L(w) = D_{L'}(w)$, as they depend only on L_n and L'_n.

If w is not right special, $D_L(w)$ is made with a single letter a, and every word xw is always followed by a, thus xwa is in L if and only if xw is in L, with $\mu[xwa] = \mu[xw]$; because of our induction hypothesis, all this remains true with “prime” signs added. Thus our hypothesis is carried over to all awb, for a and b letters. And similarly if w is not left special.

Suppose now that w is bispecial in L (hence also in L').

Let $A_L(w) = \{a_1, \dots, a_p\}$, in the $<_A$ order, and $D_L(w) = \{b_1, \dots, b_q\}$ in the $<_{D,w}$ order defined in the second item of Definition 1. Then what we shall show is that the situation is analogous to Figure 3 above.

Starting from the left, $D_L(a_1w)$ must contain b_1, otherwise $D_L(a_1w)$ contains some $b >_D b_1$, while $A_L(wb_1)$ contains some $a >_A a_1$, a_1wb and awb_1 exist, this contradicts the order condition. Then a_1wb_1 is in L, and

- if $\mu[a_1w] \leq \mu[wb_1]$, then, as μ gives positive measure to every cylinder, $D_L(a_1w)$ is reduced to the element b_1 and thus $\mu[a_1wb_1] = \mu[a_1w]$,
- if $\mu[a_1w] > \mu[wb_1]$, then $D_L(a_1w)$ contains also another element than b_1, thus it must contains b_2, otherwise a_1wb and awb_2 exist for some $b >_{D,w} b_2$ and $a >_A a_1$; hence $A(wb_2)$ contains a_1, $A(wb_1)$ cannot contain $a >_A a_1$, thus is reduced to the element a_1 and $\mu[a_1wb_1] = \mu[wb_1]$.

Thus $\mu[a_1wb_1] = \min(\mu[a_1w], \mu[wb_1]) > 0$; similarly, because of the definition of L', a_1wb_1 is in L', and, using our induction hypothesis, we get

$$\mu'[a_1wb_1] = \min(\mu'[a_1w], \mu'[wb_1]) = \min(\mu[a_1w], \mu[wb_1]) = \mu[a_1wb_1].$$

Suppose we have proved that a_rwb_s is in L iff it is in L', and $\mu[a_rwb_s] = \mu'[a_rwb_s]$, for every $r \leq i$, and $s \leq j$; suppose that a_iwb_j is in L and L', with $i < p$ or $j < q$.

Suppose first that $\sum_{r=1}^{i} \mu[a_rw] > \sum_{s=1}^{j} \mu[wb_s]$; then $j < q$ and $\cup_{r=1}^{i} D_L(a_rw)$ contains strictly $\{b_1, \dots, b_j\}$, thus by the order condition $D_L(a_iw)$ contains b_{j+1}, thus a_iwb_{j+1} is in L while $a_{i+1}wb_j$ is not in L. And we have again two cases:

- if $\sum_{r=1}^{i} \mu[a_rw] > \sum_{s=1}^{j+1} \mu[wb_s]$, then by the order condition $A_L(wb_{j+1})$ is reduced to the element a_i, thus $\mu[a_iwb_{j+1}] = \mu[wb_{j+1}]$,
- if $\sum_{r=1}^{i} \mu[a_rw] < \sum_{s=1}^{j+1} \mu[wb_s]$, then by the order condition $D_L(a_iw)$ contains (at least) the two elements b_j and b_{j+1}, and either $\mu[a_iwb_{j+1}] = \sum_{r=1}^{i} \mu[a_rw] - \sum_{s=1}^{j} \mu[wb_s]$, or $\mu[a_iwb_{j+1}] = \sum_{s=1}^{j+1} \mu[wb_s] - \sum_{r=1}^{i} \mu[a_rw]$, according to which of these two quantities is positive.

But then $\sum_{r=1}^{i} \mu'[a_rw] > \sum_{s=1}^{j} \mu'[wb_s]$, and the same analysis applies to L', thus we get the same conclusions with "prime" signs added; as all estimates depend only on properties of words of length n, we use our induction hypothesis to get $\mu'[a_iwb_{j+1}] = \mu[a_iwb_{j+1}]$, and our induction hypothesis is carried to i and $j+1$.

In the opposite case where $\sum_{r=1}^{i} \mu[a_rw] < \sum_{s=1}^{j} \mu[wb_s]$, then $i < p$, a_iwb_{j+1} is not in L while $a_{i+1}wb_j$ is in L, and a similar reasoning applies, carrying our induction hypothesis to $i+1$ and j.

If $\sum_{r=1}^{i} \mu[a_rw] = \sum_{s=1}^{j} \mu[wb_s]$, then there is a connection. Then $i < p$, $j < q$, by the reasoning above $\mu[a_iwb_{j+1}] = \mu[a_{i+1}wb_j] = \mu'[a_iwb_{j+1}] = \mu'[a_{i+1}wb_j] = 0$ and the words a_iwb_{j+1} and $a_{i+1}wb_j$ are not in L and L'. This carries our induction hypothesis both to $i+1$ and j and to i and $j+1$.

In all cases our induction hypothesis is carried to all awb, for a and b letters, and is now proved for all words of length $n+1$.

(ii) implies (iii), as a standard interval exchange transformation cannot have a wandering interval as it preserves the Lebesgue measure.

(iii) implies the $\mathcal{F}$-flipped order condition by Proposition 8. It implies recurrence as if a word w of $L(T)$ is such that there is no $v \neq w$ beginning and ending with w, then the cylinder $[w]$ is a wandering interval. □

Theorem 14. *Let L be a language on an alphabet $\mathcal{A}$ and $\mathcal{F} \subset \mathcal{A}$. Then L is the language of a generalized $\mathcal{F}$-flipped minimal interval exchange transformation, or equivalently of a standard $\mathcal{F}$-flipped minimal interval exchange transformation, if and only if it is aperiodic, uniformly recurrent, and satisfies an $\mathcal{F}$-flipped order condition.*

Proof

For any language L, the minimality of the system (X_L, S) is equivalent to the uniform recurrence of L, see for example [44] ch 5.

If T is an $\mathcal{F}$-flipped interval exchange transformation, $L(T)$ satisfies the required order condition by Proposition 8. If T is minimal, T has no periodic point, as finite orbits cannot be dense, thus $L(T)$ is aperiodic. The minimality of T implies the minimality of $X_{L(T]}$ as arbitrarily long cylinders correspond to arbitrarily small intervals, thus $L(T)$ is uniformly recurrent.

In the other direction, we apply Theorem 13 to get a standard interval exchange transformation for which $L = L(T)$. (X_L, S) is minimal, thus T can be not minimal only if arbitrarily long cylinders do not correspond to arbitrarily small intervals. This happens only if an infinite trajectory $u = u_0 \cdots u_n \cdots$ corresponds to an interval J in $(0, 1)$. But either $T^k u = u$ for some $k > 0$, which is excluded by aperiodicity, or else for all $k > 0$ $T^k J$ is disjoint from J as they are in two disjoint cylinders; but this is excluded as T is a standard interval exchange transformation, preserving the Lebesgue measure. □

Both Theorem 13 and Theorem 14 should be compared with the characterizations in [5]. A particular case of Theorem 14 is proved in [22], Theorem 2, for unflipped standard interval exchange transformations. It uses a modified version of the *i.d.o.c. condition* introduced by Keane [31]; in the present context, this condition we use can be stated as: there is at least one point γ_i (of Definition 5), and there is no i, j, $k \geq 0$, such that $T^k\beta_i = \gamma_j$. Note that this condition depends on the defining intervals, and is not intrinsic to T as the original Keane's condition. The i.d.o.c. condition can be generalized to flipped standard interval exchange transformations, it is studied by Linero Bas and Soler López in [34]; note that among flipped standard interval exchange transformations, the minimal ones constitute a small [42] albeit non-empty [41] set.

Now we can restate our old theorem in a cleaner form.

Theorem 15. *A language L on at least two letters is the language of a standard unflipped interval exchange transformation satisfying the i.d.o.c. condition if and only if it satisfies an unflipped order condition, is aperiodic and uniformly recurrent, and has no connection.*

This can be deduced from the present Theorem 14 by noticing that our i.d.o.c. condition implies Keane's one and thus minimality, and that for a minimal standard interval exchange transformation, our i.d.o.c. condition is just the absence of connections in $L(T)$.

Note that minimal generalized interval exchange transformation cannot have wandering intervals; this can also be proved by using the above results: by minimality the existence of a wandering interval for T implies the existence of a wandering cylinder for the shift on $(X_{L(T)}, S)$, and this contradicts the measure condition for $L(T)$.

4. Generalized interval exchange transformations

We turn again to the study of languages satisfying an order condition or a weaker property: we are now interested in non-recurrent elements of X_L.

Lemma 16. *Suppose L has no strong bispecial word. Let x be in X_L. There exists m such that the infinite prefix $(x_n, n \leq -m)$ is left recurrent and the infinite suffix $(x_n, n \geq m)$ is right recurrent.*

Proof

Suppose for all $m \geq m_0$ the infinite suffix $(x_n, n \geq m)$ is not right recurrent. Then for such m there exists a word $w^{(m)}$ which occurs exactly once in $(x_n, n \geq m)$. As there is a finite number of letters, by going to a subsequence we can suppose all the $w^{(m)}$ begin by the same letter, and thus this letter occurs infinitely often in $(x_n, n \geq m_0)$. Thus the longest prefix of $w^{(m)}$ which occurs at least twice in $(x_n, n \geq m)$ is a nonempty $v^{(m)}$; then $v^{(m)}$ is right special, and, by applying Lemma 5 and going to a subsequence, we can suppose $v^{(m)}$ is a suffix of a fixed infinite prefix, and also that the letter following $v^{(m)}$ in $w^{(m)}$ is a fixed a. Now we can find $m < m'$ such that $v^{(m')}$ is not shorter than $v^{(m)}$, thus $v^{(m')}$ contains $v^{(m)}$, $v^{(m')}a$ contains $v^{(m)}a$, thus $v^{(m)}a$ occurs in $(x_n, n \geq m')$, which contradicts the assumption that $v^{(m)}$ is the longest possible. And similarly for the other assumption. □

Proposition 17. *Suppose L has no strong bispecial word. Then the non-recurrent elements of X_L constitute at most a finite number of orbits of S.*

Proof

If x is not recurrent, then some $(x_n, n \geq k)$ is not right recurrent. Let k be the largest of these m: it is finite by Lemma 16, and necessarily the infinite suffix $(x_n, n \geq k+1)$ has at least two left extensions. Thus $(x_n, n \geq k+1)$ is one of the finitely many infinite suffixes W_i of Lemma 5. If $(x_n, n \geq l)$ is one of the W_i for a sequence $l \to -\infty$, then it is the same W_i by going to a subsequence, thus x is periodic and hence recurrent. Thus there is a smallest $m = m_0$ such that $(x_n, n \geq m)$ is one of the W_i, thus for $m < m_0$ there is only one left extension for $(x_n, n \geq m)$, and the sequence x is known (up to shifts) if we know W_i. As there are finitely many W_i, this give finitely many orbits. □

Lemma 18. *Let L be a non-recurrent language satisfying an $\mathcal{F}$-flipped order condition. We denote by L', on $\mathcal{A}' \subset \mathcal{A}$, the language generated by all the recurrent sequences in X_L. L' is nonempty, recurrent, satisfies an $\mathcal{F}'$-flipped order condition for some $\mathcal{F}' \subset \mathcal{F}$, and the orders defined on X_L and $X_{L'}$ coincide on $X_{L'}$.*

Proof

L' is not empty as there are recurrent sequences in X_L: indeed, by Lemma 16, there are always right recurrent infinite suffixes $(x_n, n \geq k)$ and these can be extended to the left as for each $m \geq k$ the word $x_k ... x_m$ occurs somewhere to the right. Thus for some letter x_{k-1} the word $x_{k-1}x_k ... x_m$ occurs in $(x_n, n \geq k)$ for infinitely many m, hence for all $m \geq k$. Thus we get a right recurrent infinite suffix $(x_n, n \geq k-1)$, and by iterating the process we get a recurrent bi-infinite sequence $(x_n, n \in \mathbb{Z})$.

The other assertions are immediate as L' is recurrent by construction and satisfies a restriction of the order condition on L. □

Theorem 19. *A language L is a natural coding of an $\mathcal{F}$-flipped generalized interval exchange transformation iff L satisfies an $\mathcal{F}$-flipped order condition.*

Proof

The "only if" direction is Proposition 8. We suppose now L satisfies an $\mathcal{F}$-flipped order condition. We suppose also that L is not recurrent, otherwise we can conclude by Theorem 13. Let L' be as in Lemma 18.

Let z be in $X_L \setminus X_{L'}$; by Lemma 7 if the set $Z^- = \{t \in X_{L'}; t \leq z\}$ and $Z^+ = \{t \in X_{L'}; t \geq z\}$ are nonempty, Z^+ has a unique lower bound $z^+ \in X_{L'}$ and Z^- has a unique upper bound $z^- \in X_{L'}$; in that case, $z^+ > z^-$, otherwise z would be in $X_{L'}$. If $\{t \in X_{L'}; t \leq z\}$ is empty, we define $z^- = z^+$ to be the minimum of $X_{L'}$, and if $\{t \in X_{L'}; y \geq z\}$ is empty, we define $z^+ = z^-$ to be the maximum of $X_{L'}$.

We apply Theorem 13 to find a standard interval exchange transformation T', coded by $\mathcal{A}'$, such that $L' = L(T')$. T' is not necessarily unique, we choose one. By Remark 1, every sequence in $X_{L'}$ is either the trajectory of a point x under T', or an improper trajectory associated to a point x. x is not unique for periodic sequences, but for a given sequence the set of possible x is an interval. For z in $X_{L'} \setminus X_L$, if $z^- < z^+$, there is no t in $X_{L'}$ between them, this is possible only if z^+ and z^- are two improper trajectories associated to the same point x; then x is unique as it is the rightmost point of the interval associated to z^- and the leftmost point of the interval associated to z^+, and we call it $p(z)$. If $\{t \in X_{L'}; t \leq z\}$ is empty, we define $p(z)$ to be the leftmost point in the domain of T', namely 0, and if $\{t \in X_{L'}; t \geq z\}$ is empty, we define $p(z)$ to be the rightmost point in the domain of T', namely 1. In all cases, for a given z, for $|m|$ large enough $(S^m z)^- = (S^m z)^+$, which implies $p(S^{m+1}z) = T'p(S^m z)$.

By Proposition 17 and its proof, there are finitely many orbits in $X_L \setminus X_{L'}$; we describe them as $S^n z_{(i)}, n \in \mathbb{Z}, 1 \leq i \leq K$, where $z_{(i)}$ is chosen so that $(z_{(i)m}, m \leq 0)$ is left recurrent but $(z_{(i)m}, m \leq 1)$ is not left recurrent. We can do the same reasoning in both directions, thus $z_{(i)} = y_{(i)} w_{(i)} y'_{(i)}$ for an infinite prefix $y_{(i)}$ of a sequence in $X_{L'}$, an infinite suffix $y'_{(i)}$ of a sequence in $X_{L'}$, a (possibly empty) finite word $w_{(i)}$ in L. We know also that $y_{(i)} = (z_{(i)m}, m \leq 0)$, but $y'_{(i)}$ is not necessarily the first left recurrent infinite suffix of $z_{(i)}$ as the latter might begin at a negative coordinate.

We look at the possible $S^n z_{(i)}$ such that $p(S^n z_{(i)})$ is a given point: this is a possibly infinite subset of X_L, we show that it has at most finitely many accumulation points for the usual topology: indeed, any accumulation point is a limit of points in the same orbit; if $p(S^n z_{(i)}) = p(S^{n'} z_{(i)})$ for $n \neq n'$, then $y'_{(i)}$ is periodic, equal to $vvv...$, $y_{(i)}$ is periodic, equal to $...v'v'v'$, and on this orbit there are at most finitely many accumulation points, which are the iterates by S of $...vvv...$ and $...v'v'v'...$

Thus for any given x in $p(X_L \setminus X_{L'})$, all the $S^n z_{(i)}$ such that $p(S^n z_{(i)}) = x$ can be ordered (by the order on $X_{L'}$) in a sequence $\zeta_n(x)$, $n \in H(x)$, $H(x)$ being a finite or countable set. We now blow up x to a closed interval $\hat{J}(x)$ which is divided into adjacent open intervals $J(\zeta_n(x))$, ordered from left to right by the growing order on $\zeta_n(x)$, where the length of $J(\zeta_n(x))$ is $2^{-|m|}$ if $\zeta_n(x) = S^m z_{(i)}$.

Namely, for x in $p(X_L \setminus X_{L'})$, let $l(x)$ be the sum on $n \in H(x)$ of the lengths of $J(\zeta_n(x))$; $l(x)$ is finite; for x in $(0,1)$ and not in $p(X_L \setminus X_{L'})$, we put $l(x) = 0$. For all x in $(0,1)$, we define $\phi(x) = \sum_{y \in p(X_L \setminus X_{L'}), y < x} l(y)$, this last sum being also finite. Let $M = 1 + \phi(1) + l(1)$. Then we send each point x in $(0,1)$ to the interval (possibly reduced to a point) $\hat{J}(x) = [x + \phi(x), x + \phi(x) + l(x)]$. For $l(x) \neq 0$ we divide $\hat{J}(x)$ into $J_{\zeta_n(x)}$ as described above; the image of $(0,1)$ by this Denjoy-type blow-up [18] [29] is $[0, M]$. On this interval we define T by sending in an affine way $J(S^m z_{(i)})$ onto $J(S^{m+1} z_{(i)})$ for all m and i, with a negative slope whenever $z_{(i)m}$ is in $\mathcal{F}$. If both x and $T'x$ are not in $p(X_L \setminus X_{L'})$, we define $T(x + \phi(x)) = T'x + \phi(T'x)$. This leaves T or T^{-1} undefined on some points corresponding to x where T' or T'^{-1} is undefined, on the endpoints of the $J(S^m z_{(i)})$ and finitely many images by T'^n of these endpoints. but also to x (resp. $T'x$) not in $p(X_L \setminus X_{L'})$ but such that $T'x$ (resp. x) is in $p(X_L \setminus X_{L'})$. The latter case will happen only for finitely many points, and then the trajectory under T' of either x or T'^x is some improper trajectory $(S^{m+1} z_{(i)})^+$ or $(S^{m+1} z_{(i)})^-$. Finally, we put in the interval I_e, $e \in \mathcal{A}$, all points $x + \phi(x)$ for x not in $p(X_L \setminus X_{L'})$ such that x is in the defining interval I'_e of T', and all the $J(S^m z_{(i)})$ for which $e = z_{(i)m}$. By the above considerations, all trajectories for T' remain as actual or improper trajectories for T.

Because we have respected the order, T is indeed a generalized interval exchange transformation with the required orders and $\mathcal{F}$, and its language is indeed L, as the sequences in X_L are the actual or improper trajectories for T. □

Remark 4. *The blow-up we made in the above proof sends points x to intervals $[x+\phi(x), x+\phi(x)+l(x)]$, while* [11] [10] [15] [36] *use intervals $[\phi(x), \phi(x)+l(x)]$ instead. The latter kind of blow-up is needed to build affine T, but two different points x and x' may be sent to the same point, though this will not happen if T' is minimal as then between two different points there is always some $p(z)$. In the former case the blow-up is always injective but does not produce affine interval exchange transformations. In the proof of Theorem 20 below we shall use a mixture of these two.*

We shall see how this proof works on Examples 4 of Section 6.2 and 9 of Section 6.4.

5. Affine interval exchange transformations

Theorem 20. *If L is a natural coding of an $\mathcal{F}$-flipped affine interval exchange transformation for which the absolute value of the slope is $\exp\theta_e$ on the defining interval I_e, then L satisfies an $\mathcal{F}$-flipped order condition and for each non recurrent sequence z in X_L, $\sum_{n\geq 0}\exp\sum_{j=0}^{n}\theta_{z_j} < +\infty$, and $\sum_{n>0}\exp -\sum_{j=-n}^{-1}\theta_{z_j} < +\infty$.*

If L satisfies an $\mathcal{F}$-flipped order condition and there exist real numbers $\theta_e, e\in\mathcal{A}$, such that for each non recurrent sequence z in L, $\sum_{n\geq 0}\exp\sum_{j=0}^{n}\theta_{z_j} < +\infty$, and $\sum_{n>0}\exp -\sum_{j=-n}^{-1}\theta_{z_j} < +\infty$, then L is a group coding of an $\mathcal{F}$-flipped affine interval exchange transformation.

Proof

Let z be a non recurrent sequence. Let L' be as in the proof of of Theorem 19. As z is not recurrent, z cannot have all its words in the same uniformly recurrent component (of Proposition 12) of L', and thus not all its words in L'. Thus we can find a word $w = w_1...w_k$ which is not in L' and appears only once in z. By Lemma 16 it will appear only finitely many times in each of the other non recurrent sequences, and by Lemma 17 we can extend it so that it appears only in z and its shifts. Thus the cylinder $[w]$ is disjoint from all its iterates by T, and for the Lebesgue measure μ both $\sum_{n\geq 0}\mu(T^n[w])$ and $\sum_{n\geq 0}\mu(T^{-n}[w])$ must be finite. Suppose $w = z_k...z_{k+k'}$; then T is of slope $\exp\theta_{z_k}$ on $[w]$, $\exp\theta_{z_{k+1}}$ on $[Tw]$, $\exp\theta_{z_{k-1}}$ on $T^{-1}[w]$ and so on, thus $\mu(T[w]) = \exp\theta_{z_k}\mu[w]$, $\mu(T^2[w]) = \exp\theta_{z_{k+1}}\mu(T[w])$, $\mu([w]) = \exp\theta_{z_{k-1}}\mu(T^{-1}[w])$ and so on. Our conditions come from replacing k by 0, which amounts to multiplying by a constant.

We start now from L and build a generalized interval exchange transformation T, as in the proof of Theorem 19, but with two modifications.

First, the length we assign to the interval $J(S^m z_{(i)})$ is 1 for $m = 0$, then is defined such that $J(S^m z_{(i)})$ is sent to $J(S^{m+1}z_{(i)})$ by an affine map of slope $\exp\theta_{z_{(i)m}}$ if $z_{(i)m}$ is in $\mathcal{F}^c$, $-\exp\theta_{z_{(i)m}}$ if $z_{(i)m}$ is in $\mathcal{F}$.

Second, when we make the blow-up of $(0,1)$ to the domain of T, we replace the map $x\to\phi(x)+x$ by the map $x\to\phi(x)+1_K(x)x$ for some union of intervals K. To find K, we apply the proof of Proposition 12 to the recurrent language L' : for some fixed N large enough, its Rauzy graph G_N splits into a finite number of connected components, corresponding to languages which either come from one purely periodic orbit or are aperiodic and uniformly recurrent. This implies that there exist a finite number of T'-invariant sets K_j, which are finite unions of cylinders of length N, and such that, on each K_j, either T' is minimal, or every point is periodic (note that this reproves Corollary 2.9 of [2], and extends it to flipped interval exchange transformations).

Now, for each i, we look at those points $p(S^m z_{(i)}), |m| \geq m_i$, where m_i is chosen such that for these points $p(S^{m+1} z_{(i)}) = T' p(S^m z_{(i)})$, and let K^c be the union of those K_j which contain at least one $p(S^{m_i} z_{(i)})$ or $p(S^{-m_i} z_{(i)})$ in their closure. If $x \neq x'$, our construction ensures that x and x' are blown-up into different points, unless maybe if x and x' belong to the same connected component K' of the same K_i. In the latter case either minimality implies that there is at least one $J(S^m z_{(i)})$ between their images, in which case again x and x' are blown-up into different points, or periodicity implies that x and x' have the same trajectory, and this trajectory is not lost in the blow-up as it remains as the improper trajectory of some endpoint of K'. Thus again all trajectories for T' remain as actual or improper trajectories for T.

The map T built in this way is a generalized $\mathcal{F}$-flipped interval exchange transformation with defining intervals I_e. By the same proof as in [11] [10] [15] [36], if the set K in $(0, 1)$ is blown up to $\hat{K}$, T is affine of slope $\exp \theta_e$ (resp. $-\exp \theta_e$ if e is in $\mathcal{F}$) on each connected component of $I_e \cup \hat{K}^c$. Each connected component of $I_e \cup \hat{K}$ is the union of finitely many $J(S^m z_{(i)})$, on which T is affine of slope $\exp \theta_e$ (resp. $-\exp \theta_e$ if e is in $\mathcal{F}$) and finitely many intervals on which T is affine of slope 1 (resp. -1 if e is in $\mathcal{F}$). Thus by dividing the intervals further we can make T affine with constant slope on its defining intervals, and L is a grouped coding of an affine interval exchange transformation. ☐

Corollary 21. *If L satisfies an $\mathcal{F}$-flipped order condition and X_L contains some non-recurrent sequences, a sufficient condition for L to be a natural coding of an $\mathcal{F}$-flipped affine interval exchange transformation is that*

- *there exist real numbers $\theta_e, e \in \mathcal{A}$, such that for each non recurrent sequence z in L, $\sum_{n \geq 0} \exp \sum_{j=0}^{n} \theta_{z_j} < +\infty$, and $\sum_{n > 0} \exp - \sum_{j=-n}^{-1} \theta_{z_j} < +\infty$,*
- *each factor of a recurrent sequence in X_L is also a factor of a non-recurrent sequence in X_L.*

Proof

With the notations of the proof of Theorem 20, the last condition implies that in the closure of every K_j there is at least one $p(S^{m_i} z_{(i)})$ or $p(S^{-m_i} z_{(i)})$, thus K is empty and there is no subinterval of I_e on which T is affine of slope $1 \neq \exp \theta_e$ or $-1 \neq -\exp \theta_e$, which ensures that L is the natural coding of T. ☐

This criterion will be further discussed in Section 6.3 below.

We can get a criterion for L to be a grouped coding of an affine interval exchange transformation by essentially rephrasing Theorem 20.

Definition 11. *Suppose L satisfies an $\mathcal{F}$-flipped order condition. A language $\hat{L}$ is a* splitting *of L if its letters are e_i, $e \in \mathcal{A}$, $1 \leq i \leq k_e$, forming the set $\hat{\mathcal{A}}$, such that $L = \phi(\hat{L})$ where we define $\phi(e_i) = e$ for all e and i, and extend ϕ to a morphism for the concatenation on words, and $\hat{L}$ satisfies an $\hat{\mathcal{F}}$-flipped order condition, where*

- *$\hat{\mathcal{F}}$ is made of all the e_i such that e is in $\mathcal{F}$,*
- *if e is in $\mathcal{F}^c$, e_1, ... e_{k_e} are consecutive and ordered in the same way by the orders $<_A$ and $<_D$ of $\hat{L}$,*

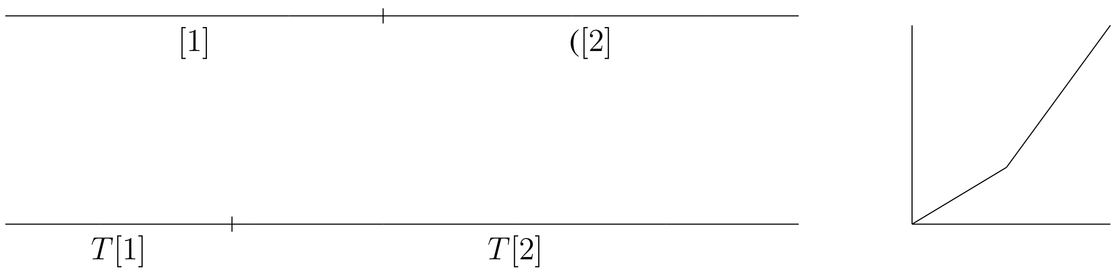

FIGURE 4. Fake Sturmian

- *if e is in $\mathcal{F}$, e_1, ... , e_{k_e} are consecutive and ordered in opposite ways by the orders $<_A$ and $<_D$ of $\hat{L}$,*
- *if $e <_A e'$, resp. $e <_D e'$ in L, then $e_i <_A e'_j$, resp. $e_i <_D e'_j$, in L' for all i and j,*

Theorem 22. *L is a group coding of an $\mathcal{F}$-flipped affine interval exchange transformation if and only L satisfies an $\mathcal{F}$-flipped order condition and there exists a splitting $\hat{L}$ of L, and real numbers $\theta_e, e \in \hat{\mathcal{A}}$, such that for each non recurrent sequence z in $X_{\hat{L}}$, $\sum_{n\geq 0} \exp \sum_{j=0}^{n} \theta_{z_j} < +\infty$, and $\sum_{n>0} \exp - \sum_{j=-n}^{-1} \theta_{z_j} < +\infty$.*

Proof

In one direction, we suppose L is a grouped coding of T affine, define $\hat{L}$ to be the natural coding of T and apply Theorem 20. In the other direction, we start from $\hat{L}$ and use Theorem 20 to build an affine interval exchange transformation T such that $\hat{L}$ is a grouped coding of T. But then $L = \phi(\hat{L})$ is a further grouped coding of T. □

6. COUNTER-EXAMPLES AND QUESTIONS

6.1. Affine with natural coding.

Example 3. *Let L be generated by the bi-infinite sequence $\ldots 111222\ldots$. Note that it is of complexity $n + 1$ but not uniformly recurrent, and in the founding paper* [40] *it is not included in Sturmian languages, we can call it a* fake Sturmian *language. It satisfies the unflipped order condition with $1 <_D 2$, $1 <_A 2$, but (unsurprisingly as it is not recurrent) is not the language of a standard interval exchange transformation: that could only be the identity on two disjoint open intervals I_1 and I_2, and the only possible words are 1^n and 2^m. However, L is the natural coding of an affine interval exchange transformation: L_2 is the language of length 2 of any affine 2-interval exchange transformation, with the same orders, such that TI_1 is strictly longer than I_1, and, as L is determined by L_2 because there is no bispecial word except the empty one, L is indeed the natural coding of any of these affine interval exchange transformations.*

Example 3 can also be dealt with as in Theorem 20, with T' being the identity acting on I_1 and I_2.

Because of Theorem 20 and Corollary 21, if we want to build nontrivial examples of affine interval exchange transformation with a non recurrent language L as a natural coding, we can just use [11] [10] [15] [36]. Starting from a minimal standard interval exchange transformation T', we find a point $x_\star$ and a map θ constant on the defining intervals such that the Birkhoff sums

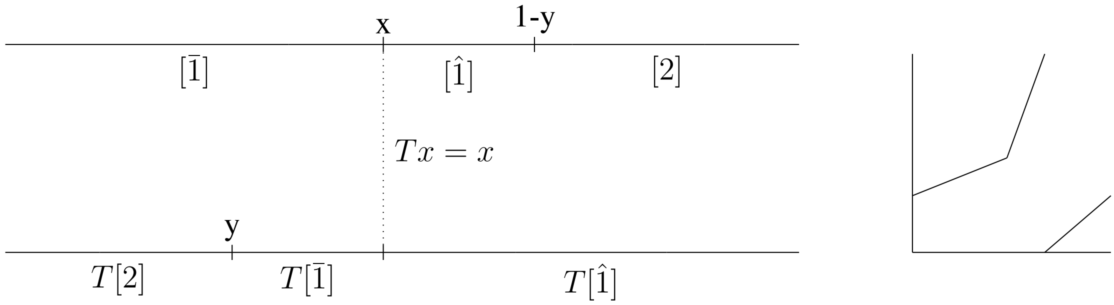

FIGURE 5. Skew Sturmian

$\sum_{j=0}^{n}\theta(T'^j x_\star)$ and $\sum_{j=-1}^{-n}\theta(T'^j x_\star)$ behave as we need. Then we make a blow-up of all points in the orbit of $x_\star$, and get an affine T with a non-recurrent trajectory using an extra letter as in Example 9. Note however that, because of the order condition, $x_\star$ must be an endpoint of a defining interval of T'; if $x_\star$ is not already such an endpoint, we can make it be one with a modification of T', by splitting the interval I'_e containing $x_\star$ into $I'_e \cap (0, x_\star)$ and $I'_e \cap (x_\star, 1)$. We can also get examples by blowing up two half-orbits, of two points $x_\star$ and $x'_\star$ so that the Birkhoff sums $\sum_{j=0}^{n}\theta(T'^j x_\star)$ and $\sum_{j=-n}^{-1}\theta(T'^j x'_\star)$ behave as we need, provided we then modify T' so that $x_\star$ and $x'_\star$ are endpoints of defining intervals. This is always possible if T' is minimal but non uniquely ergodic. In that case we can find two invariant measures ν and ν' and a vector θ so that $\sum \theta_e \nu(I_e) < 0$, $\sum \theta_e \nu'(I_e) > 0$, and then choose $x_\star$ generic for ν, $x_\star$ generic for ν'. But if T' is uniquely ergodic, then we need $\sum \theta_e \nu(I_e) = 0$ for the invariant measure ν, and, by [28], for any given θ the set of x such that $\sum_{j=0}^{n}\theta(T'^j) \to -\infty$ or $\sum_{j=-1}^{-n}\theta(T'^j x) \to +\infty$ is of measure 0.

6.2. **Affine with grouped coding.**

Example 4. *Let L the language generated by the bi-infinite sequence $\dots 1112111\dots$, which is a* skew Sturmian *language as defined in* [40]. *It satisfies the unflipped order condition with $1 <_D 2$, $2 <_A 1$, but is not the natural coding of any affine interval exchange transformation T. Indeed, the sequence $\dots 1111\dots$ in X_L would define a fixed point x for T, in the interior of* $[1]$, *and, if $0 < y < x$ is the right endpoint of* $T[2]$, *T would have to send $(0, x)$ to (y, x) and $(x, 1-y)$ to $(x, 1)$, thus having a slope < 1 on a subinterval of* $[1]$ *and a slope > 1 on another subinterval of* $[1]$. *However, if $\tilde{L}$ is the language generated by the bi-infinite sequence $\dots \hat{1}\hat{1}\hat{1}2\bar{1}\bar{1}\bar{1}\dots$, as in Example 3 $\tilde{L}$ is the natural coding of any affine interval exchange transformation T sending $I_{\bar{1}} = (0, x)$ to (y, x), $I_{\hat{1}} = (x, 1-y)$ to $(x, 1)$, $I_2 = (1-y, 1)$ to $(0, y)$, with $0 < y < x < 1-y$. If we now code T by the intervals $\tilde{I}_1 = I_{\bar{1}} \cup I_{\hat{1}}$ and $\tilde{I}_2 = I_2$, we see that L is indeed a grouped coding of an affine interval exchange transformation as in Definition 8.*

We show how the proof of Theorem 19 works on Example 4. L' is the language generated by the single sequence $y = ...111...$, and thus T' is the identity acting on $I_1 = (0, 1)$. In the single non recurrent orbit ...11211..., z is the point for which $z_0 = 2$, and we have $Sz < S^2z < S^3z < ... < y < z < ...S^{-3}z < S^{-2}z < S^{-1}z < z$, with y being the limit of the $S^n z$ when $n \to \pm\infty$. Thus $p(S^n z) = 1$ for all $n \leq 0$, $p(S^n z) = 0$ for all $n \geq 1$; our construction gives an affine interval exchange transformation T sending $(0, 1)$ to $(1/2, 1)$, $(1, 2)$ to itself, $(2, 3)$ to $(2, 4)$, $(3, 4)$ to $(0, 1/2)$, while the cylinder $[1]$ is the interval $(0, 3)$ and the cylinder $[2]$ is the interval $(3, 4)$. L is the natural coding of T if we consider it as a generalized interval exchange transformation, but is only a grouped coding of an affine interval exchange transformation, as remarked above. Note that the fact that T is affine is a (conceivably) accidental result of our choices of the lengths of the added intervals, see the proof of Theorem 20, and that the present map T is slightly different from the one of Example 4. The latter could be obtained from the former by collapsing the interval $(1, 2)$, as in the second blow-up of Remark 4: this is possible though T' is not minimal because all the points in $(1, 2)$ have the same trajectory, ...1111..., and this will remain, as an improper trajectory, after the collapsing.

6.3. **Natural versus grouped.**

Conjecture 1. *The conditions in Theorem 20 are necessary and sufficient for L to be a natural coding of an $\mathcal{F}$-flipped affine interval exchange transformation.*

Indeed, the last condition of Corollary 21 is satisfied if T' is minimal, but also by Example 3 above. But this condition is *not necessary*, for different reasons shown in the following three examples, most significantly Example 7.

Example 5. *Let L be the language generated by the two infinite sequences ...111222.... and ...3333.... Then K is a union of defining intervals of T', so the grouped coding of Theorem 20 turns out to be a natural coding.*

Example 6. *Let L be generated by ...121312131213... and ..114114114115115115... Then K is nonempty and not a union of defining intervals, but an allowed value of θ_1 is 0, and 1 is the only e for which I'_e contains strictly a connected component of K, thus again the proof of Theorem 20 provides a natural coding.*

Example 7. *Let L be generated by ...332332331331.... and ...413241324132...; then either $\theta_1 \neq 0$ or $\theta_2 \neq 0$, hence the proof of Theorem 20 would yield an interval exchange transformation with two different slopes, 1 and $\exp \theta_e$, on I_e for $e = 1$ or $e = 2$, thus L is not the natural coding of any affine interval exchange transformation built by the proof of Theorem 20. However, L is the natural coding of another affine interval exchange transformation, namely any member of the family defined by the orders $1 <_D 2 <_D 3 <_D 4$ and $4 <_A 3 <_A 1 <_A 2$, slopes a on I_1, $1/a$ on I_2, 1 on I_3 and I_4, and lengths $|I_1| = l + r$, $|I_2| = a(l + r)$, $|I_3| = al + 2(a + 1)r$, $|I_4| = l$, for any given $0 < a < 1$, $l > 0$ and $r > 0$.*

The only example we know of a grouped coding which is not a natural coding is Example 4, which does not satisfy the condition on Birkhoff sums. Whatever the value of θ_1 and θ_2, either the $\exp n\theta_1$, $n \geq 0$, or the $\exp n\theta_1$, $n < 0$, must be bounded away from 0 (while for the language $\tilde{L}$ of Example 4 we can take $\theta_1 < 0$ and $\theta_3 > 0$).

Question 2. *Does there exist an aperiodic language which is a grouped coding of an affine interval exchange transformation, but not a natural coding of any affine interval exchange transformation?*

Questions 1 and 2 suggest what we dare not call a conjecture.

Question 3. *Is it true that L is a group coding of an $\mathcal{F}$-flipped affine interval exchange transformation if and only if L satisfies an $\mathcal{F}$-flipped order condition and there exist real numbers $\theta_e, e \in \mathcal{A}$, such that the two following conditions hold?*

- *For each non recurrent sequence z in L which is not ultimately periodic to the left, $\sum_{n \geq 0} \exp \sum_{j=0}^{n} \theta_{z_j} < +\infty$.*

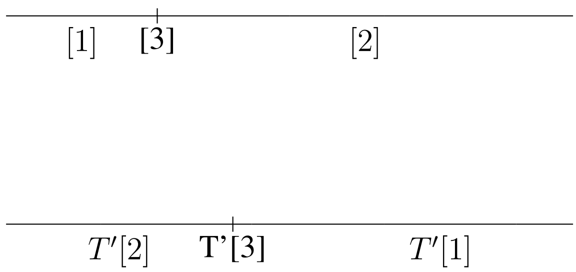

FIGURE 6. Episkew before blow-up

- *For each non recurrent sequence z in L which is not ultimately periodic to the right,* $\sum_{n>0} \exp - \sum_{j=-n}^{-1} \theta_{z_j} < +\infty$.

Remark 5. *As in the last example above, the same language can be a coding of two fairly different affine interval exchange transformations. In particular, Sturmian languages are languages of standard interval exchange transformations, thus also of affine interval exchange transformations without wandering intervals. It is known since* [40] *that if we want to generate a given Sturmian language L by a standard 2-interval exchange transformation, the parameter $\mu[I_2]$ has two possible values, or one up to reversal of the orientation. If we want the interval exchange transformation to be affine, we have two parameters, for example its slopes a and b on I_1 and I_2, and we do not know which amount of freedom we have in this case.*

Question 4. *For a given L, what can be said of the set of (a, b) for which L is the language of the corresponding affine interval exchange transformation?*

About this question, it is worth mentioning a surprising example due to M. Shannon (unpublished):

Example 8. *One can build (by geometrical methods) a grouped coding of an affine interval exchange transformation T which is a Sturmian language, associated to a rotation of angle α, but while α is in $\mathbb{Q}(\sqrt{3})$, all the parameters defining T are in $\mathbb{Q}(\sqrt{5})$.*

6.4. **Generalized.**

Example 9. *Let L' be the Sturmian language which is the natural coding of the unflipped standard interval exchange transformation T' sending $I_1 = (0, 1-\alpha)$ to $(\alpha, 1)$ and $I_2 = (1-\alpha, 1)$ to $(0, \alpha)$ for an irrational $\alpha < 1/2$. Let $y_n = i$ whenever $T^n\alpha$ is in I_i, $n \geq 0$, and $y'_n = i$ whenever $T^n(1-2\alpha)$ is in I_i, $n \leq 0$; when $\alpha = \frac{3-\sqrt{5}}{2}$, y is the so-called Fibonacci sequence on 1 and 2, and y' is y written backwards. Let L be the language generated by the infinite sequence $...y'_{-2}y'_{-1}y'_0 3 y_0 y_1 y_2...$ It satisfies the unflipped order condition with $1 <_D 3 <_D 2$, $2 <_A 3 <_A 1$ (note that no other unflipped order is possible, because of the way the empty bispecial is resolved, nor is any flipped order because of the way the bispecial 0 is resolved). By* [6], *we can call L an* episkew *language*

By Theorem 19 L is the natural coding of a generalized interval exchange transformation, but it is not the natural or grouped coding of any affine interval exchange transformation: this will be a straightforward consequence of either one of two independent results we show below, Theorems 23 and Theorem 24.

Example 9 is a typical case to see how the proof of Theorem 19 works. Then L' and T' are as described in its definition. Then the only not right recurrent orbit in X_L is the one of $y'3y$ defined above. There is one z, the element of this orbit such that $z_0 = 3$. Because of the order condition, the defining interval I_3 of T must be between I_1 and I_2, and TI_3 between TI_2 and TI_1; this implies that $p(z) = 1 - \alpha$ and $p(Tz) = \alpha$. Thus to get T we define T' on the extra point z and then blow up its orbit, replacing each one of its points by small enough intervals.

In Theorem 23 below, we show that the episkew language, and others sharing its properties, cannot be codings of generalized interval exchange trarnsformations satisfying some conditions of regularity. We shall use here a set of conditions, the *class P*, defined by Herman in his *Thèse d'État* (1976), first published in [29].

Definition 12. *A transformation T is of class P if, except possibly on a countable set of points, DT exists and $DT = h$ where h is a function with bounded variation, and $|h|$ is bounded from below by a strictly positive number.*

These conditions constitute a natural set of sufficient conditions for a function to satisfy the Denjoy-Koksma inequality, stated and proved in [29] chapter VI, and it is exactly for that purpose that we use them in the following Theorem. In the same chapter Herman proves also that a functions of class P with irrational rotation number satisfies the Denjoy theorem on C^0 conjugacy with a rotation, and gives two families of examples of such functions, namely C^1 diffeomorphisms with Df of bounded variation, and continuous and piecewise linear (with finitely many pieces) homeomorphisms.

Theorem 23. *Let L be non recurrent, and a natural coding of an unflipped generalized interval exchange transformation T. Suppose the language L' of Lemma 18 is aperiodic, uniformly recurrent, and its arrival and departure order are conjugate by a circular permutation. Then T cannot be of class P.*

Proof

A sequence in X_L can be a trajectory for T (actual or improper) either of a single point or of all points in an interval; the intervals which correspond to a single trajectory include those coded by non-recurrent trajectories, but possibly others; as intervals corresponding to different trajectories are all disjoint, there are at most countably many of them; let E_n, $n \in \mathbb{Z}$, be these intervals. We build another interval exchange transformation $\hat{T}$ on the interval $(0, 1)$ by making a deblow-up: this will be done in two different ways corresponding to the inverses of the two blow-ups in Remark 4. If $\sum_{n\in\mathbb{Z}} \mu(E_n) < 1$, μ being the Lebsegue measure, we just shrink each E_n (including its endpoints) to a single point, by sending x to $x - \mu(]0, x[\cup \cap_{n\in\mathbb{Z}} E_n)$, sending $(0, 1)$ to a smaller interval, which we rescale to get $(0, 1)$. If $\sum_{n\in\mathbb{Z}} \mu(E_n) = 1$, we use the semi-conjugacy of [11] [10] [15] [36]. Let x_n be the left end of E_n; the condition on the lengths of E_n implies the x_n are dense in $(0, 1)$; Then we send to x_n all points in E_n, while a point x which is not im any E_n is sent to the unique (by density of the x_n) point x' such that $\sum_{x_n<x'} \mu(E_n) = x$. In both cases all intervals coded by non-recurrent trajectories are shrunk to single points, and $\hat{T}$ has no wandering interval. Then the natural coding of $\hat{T}$ is the language L', as every recurrent trajectory of T remains as an actual or improper trajectory of $\hat{T}$, while non-recurrent trajectories of T are not trajectories of $\hat{T}$.

By construction, each trajectory in $X_{L'}$ is the (actual or improper) trajectory under $\hat{T}$ of one point in $(0, 1)$, and also the (actual or improper) trajectory under T' of one point in $(0, 1)$. By Remark 2 the order on $X_{L'}$, whether we consider trajectories of T' or $\hat{T}$, corresponds to the natural order on $(0, 1)$, thus we can make a bijective, continuous and increasing map ψ from $(0, 1)$ to $(0, 1)$

such that $\psi \circ \hat{T} = T' \circ \psi$.

Suppose T is of class P. Let J be a wandering interval for T; then we have $\mu(T^n J) = \mu(J) \exp \sum_{j=0}^{n-1} \theta_j$ where $\exp \theta_j$ is the average value of DT on $T^j J$. If h is the function defining class P, we define a function θ, by $\theta(x) = \log h(x)$ when x is not in any E_n, while, for any x in E_n, $\theta(x)$ is the logarithm of the average value of h on E_n. Then the measures of $T^n J$ are given by for $n > 0$, $\mu(T^n J) = \mu(J) \exp \sum_{j=0}^{n-1} \theta(T^j x)$ for some point x in J, while for $n < 0$, $\mu(T^n J) = \mu(J) \exp - \sum_{j=n}^{-1} \theta(T^j x')$ for some point x' in J. As J is a wandering interval, $\mu(T^n J)$ tends to 0 when n goes to $\pm\infty$.

The deblow-up sends Birkhoff sums of θ under T to Birkhoff sums of a function $\hat{\theta}$ under $\hat{T}$, thus by the hypothesis these are Birkhoff sums of a measurable function under an irrational rotation, and, as in the case of a natural coding, because of the unique ergodicity, we must have $\int \hat{\theta} d\mu = 0$ for the Lebesgue measure. The properties of h imply that the function θ has bounded variation, thus so does $\hat{\theta}$; but then by the Denjoy-Koksma inequality [29] the Birkhoff sums are bounded on a subsequence corresponding to the denominators of the partial quotients of the angle, and we get a contradiction. □

The above theorem applies to Example 9 or any example built in the same way from any Sturmian word, but also, in contrast with Theorem 24 or Corollary 26 below, we may equip T' with extra points γ_i which are not discontinuities, giving extra possibilities for building L. All these give counter-examples of languages which are natural codings of a generalized interval exchange transformation, but not natural codings of any generalized interval exchange transformation of class P, and thus, in particular, not grouped codings of any affine interval exchange transformation.

Theorem 24. *Let L' be a natural coding of a non purely periodic unflipped standard interval exchange transformation. Let $w_n = av_nb$, $a \in \mathcal{A}$, $b \in \mathcal{A}$, be an infinite sequence of bispecial words in L'. Let u be the infinite prefix in $X_{L'}$ ending with w_n for all n, and v the infinite suffix beginning with w_n for all n. Let ω be a symbol which is not a letter of L', and L be the language generated by the union of all words in L' and the bi-infinite word $u\omega v$.*

Then L is a natural coding of a generalized interval exchange transformation, but not a grouped coding of any affine interval exchange transformation.

Proof

We build a generalized interval exchange transformation with natural coding L as in the proof of Theorem 19, by adding a point coded by ω at position γ, where $T'^{-|w_n|}\gamma$ is in $[w_n]$ for all n, and its image at position β, where β is in $[w_n]$ for all n, and making a Denjoy-Koksma blow-up.

Suppose L is a grouped coding of an affine interval exchange transformation T.Let $\exp \theta(x)$ be the slope of T at point x, θ is piecewise constant with $K \geq 0$ jumps. As $[\omega]$ is a wandering interval, we can write the $T^j[\omega]$ as a (Rokhlin) tower, in which each discontinuity of θ appears at most once. Thus $[\omega]$ is partitioned into at most $K+1$ disjoint intervals, on which for any $n > 0$ $\sum_{j=-n}^{-1} \theta(T^j x)$ and $\sum_{j=0}^{n} \theta(T^j x)$ are constant; we choose one of these subintervals, which is also a wandering interval, and call Q_{-n} and Q_n the respective value of these sums on it; when $n \to +\infty$, $Q_{-n} \to +\infty$ and $Q_n \to -\infty$.

We look now at the induced (or first return) map of T on the interval $[w_n]$, for fixed n. By the standard reasoning of [31], this interval is partitioned into R subintervals $J_{i,n}$, which are the bases of R disjoint Rokhlin towers made with disjoint intervals $J_{i,n}$, ... $T^{h_{i,n}-1}J_{i,n}$, and then $T^{h_{i,n}}J_{i,n}$ is in $[w_n]$. The levels $T^l J_{i,n}$ are in the defining interval corresponding to the $l+1$-th letter of the word w_n, and their union for fixed l is an interval for $l \leq h'_n - 1$, where $h'_n = \min_i h_{i,n}$. Among these towers, there is one for which $h_{i_0,n} = |w_n| + 1$ and $T^{h_{i_0,n}-1}[w_n]$ is in $[\omega]$. All the others are coded by orbits of T', and $h'_n = \to +\infty$ when $n \to +\infty$, otherwise the w_n would be in a periodic orbit, which is impossible as they are bispecial.

As ω is always followed by w_n, there is a subinterval of $[w_n]$ on which $\sum_{j=-1}^{l-1} \theta(T^j x) = Q_l$ for all $l \leq h'_n$.

For fixed n, each discontinuity of θ appears at most once in the union of towers of bases $J_{i,n}$, and all these towers are adjacent from levels 0 to $h'_n - 1$. Thus, for two different points x and x' in $[w_n]$ the values of $\theta(T^j x)$ and $\theta(T^j x')$ are different for at most K distinct values of j between 0 and $h'_n - 1$. Thus, taking also into account the replacement of $T^{-1}x$ by $T^l x$, there exists a constant M such that for all n, all $l < h'_n$, all x in $[w_n]$, $Q_l - M < \sum_{j=0}^{l} \theta(T^j x) < Q_l + M$.

As ω is always preceded by w_n, there is a subinterval of $J_{i_0,n}$ on which $\sum_{j=0}^{h_{i_0,n}-2} \theta(T^j x) = Q_{-h_{i_0,n}+1}$. This implies that $Q_{-h_{i_0,n}+1}$ is between $Q_{h_{i_0,n}-2} - M$ and $Q_{h_{i_0,n}-2} + M$ for all n, which is a contradiction because of their behaviour when $n \to +\infty$. □

This gives us many counter-examples, of natural codings of a generalized interval exchange transformation which are not grouped codings of any affine interval exchange transformation, including Example 9 for completely different reasons as the ones in Theorem 23. But for a given T' a priori we cannot use all possible positions to add an interval $[\omega]$, nor can we create new points γ_i as possible positions.

6.5. **Geometric.** In an intermediate case between the rotations of Theorem 23 and the general systems of Theorem 24, we give counter-examples which are not natural codings of an affine interval exchange transformation, starting from some famous T', in which we can use all possible positions γ_i to add an interval, but not create new points γ_i as possible positions.

We construct examples with nice properties arising from exotic translation surfaces. We refer to [24] for an introduction on translation surfaces and their moduli spaces. The *Eierlegende Wollmilch Sau* and the *Ornithorynque* are square-tiled surfaces with remarquable properties. The Eierlegende Wollmilch Sau is described in Figure 7. In both cases, the Veech group of the translation surface is equal to $SL(2, \mathbb{Z})$. We consider ϕ a pseudo-Anosov homeomorphism in the affine group. When ϕ fixes a separatrix, using Veech's zippered rectangles method (see [49] or [51]), one can construct a self-similar interval exchange transformation S acting on a transversal of the expanding foliation of ϕ. We will call such interval exchange transformation a self-similar EW-interval exchange transformation (resp.self-similar Or-interval exchange transformation).

Theorem 25. *Let S be a self-similar EW-interval exchange transformation (resp. self-similar Or-interval exchange transformation) and S' an affine interval exchange transformation semi-conjugate to S, then S' is topologically conjugate to S. In particular, S' has no wandering interval.*

Proof

We make the proof when the surface is the Eierlegende Wollmilch Sau, denoted by EW. It is exactly similar for the Ornithorynque.

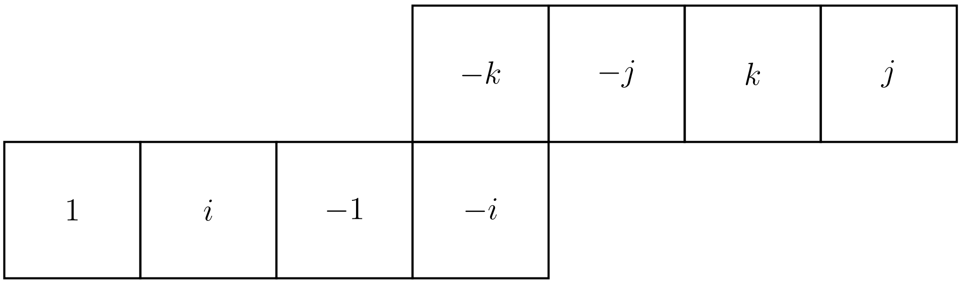

FIGURE 7. The Eierlegende Wollmilch Sau EW.
The gluings of the squares follow the action of the quaternion group. Multiplication by i defines the square on the right, by j the square on top.

The interval exchange S is a 9-interval exchange transformation defined on an interval I. The intervals of continuity of T are denoted by $I_1, \ldots, I_9$. We consider the function f from I to $\mathbb{R}$ with value at x, the logarithm of the slope of the affine interval exchange transformation T' at x. The function f is constant on the intervals of continuity of T and is orthogonal to the lengths vector of the intervals of S (see [11]). In other words, f can be also considered as a vector belonging to a co-dimension one subset of $\mathbb{R}^9$. It is enough to prove that, for every x there exist two sequences of positive integers (n_i) and (m_i) and a constant C such that

$$\sum_{j=0}^{n_i} |f(S^j(x)| < C \text{ and } \sum_{j=-m_i}^{0} |f(S^j(x)| < C. \tag{1}$$

This trivially implies that

$$\sum_n \exp(\sum_{j=0}^{n} f(S^j(x))$$

and

$$\sum_n \exp(\sum_{j=-n}^{0} f(S^j(x))$$

are divergent series which contradicts the existence of wandering intervals and prove the topological conjugacy between S and S' (see [11] or [15]).

To prove equation (1), we use the autosimilarity, and consider the sequence $I^{(p)}$ of intervals on which the induced map $S^{(p)}$ is similar to S. Let $C_1^{p)}, \cdots, C_9^{p)}$ be the Rokhlin towers with base the continuity intervals of $I^{(p)}$ and $r_1^{(p)}, \ldots r_9^{p)}$ their heights. For each $1 \leq k \leq 9$,

$$C_k^{(p)} = \bigcup_{i=0}^{r_k^{(p)}-1} T^i(I_k^{(p)})$$

where $I_1^{(p)}, \ldots, I_9^{(p)}$ form a partition of $I^{(p)}$. We assume there exists a constant K, such that, for every p and every $k \in \{1, \ldots, 9\}$, for every y in the base of the tower $C_k^{(p)}$

$$|\sum_{i=0}^{r_k^{(p)}} f(S^i y)| < K. \tag{2}$$

The inequality (2) proves equation (1) for the basis of a Rokhlin tower taking $n_p = r_k^{(p)}$.

Since $S^{(p)}$ is minimal (because $S^{(p)}$ is conjugate to S which is a minimal interval exchange transformation), there exists $\kappa > 0$ such that, for each point $x \in I_k^{(p)}$, there exists $j < \kappa$, with $(S^{(p)})^j(x)$ belongs to $I_k^{(p)}$. In other words, return times to continuity intervals of $S^{(p)}$ are bounded by a universal constant depending only on S.

Let now x be in $S^i(I_k^{(p)})$. The itinerary of x after time $r_k^{(p)}$ is a succession of towers $C_{1(x)}^{(p)}, \dots, C_{j(x)}^{(p)}$. By the previous remark, there exists $0 < j(x) < \kappa$ such that $j(x) = k$. Let

$$n_p = r_k^{(p)} - i + r_{1(x)}^{(p)} + \cdots r_{j(x)-1}^{(p)} + i$$

It means that we follow the itinerary of x until it comes back to its initial step, the step i of the Rokhlin tower k. Thus the Birkhoff sum of f from time 0 to time $n_p - 1$ is the sum along the towers $C_k^{(p)}, C_{1(x)}^{(p)}, \dots, C_{j(x)}^{(p)}$. Using bound (2), we have:

$$\sum_{j=0}^{n_p-1} |f(T^j(x)| < K\kappa,$$

and thus equation (1).

The proof of equation (2) comes from a classical argument that can be found, in a more general setting, in Zorich's work, for which we now give now give some geometric background. The EW is a surface of genus 3 with 4 singularities. We recall that the real dimension of the relative homology with real coefficients $H_1(EW, \Sigma, \mathbb{R})$ is 9. That's related to the fact that S is a 9-interval exchange transformation. The subspace, $H_1^{(0)}(EW, \Sigma, \mathbb{R})$, made of homology classes with zero holonomy has dimension 7. Since this surface is a covering of the torus, the homology of the surface splits over $\mathbb{Q}$ into the standard part that comes from the homology of the torus and the space $H_1^{(0)}(EW, \Sigma, \mathbb{R})$. The action of the affine group on $H_1^{(0)}(EW, \Sigma, \mathbb{R})$ is an action by a finite group (see [39] for precise definitions and a very detailed combinatorial approach to this problem). let ϕ be the pseudo-Anosov homeomorphism from which S is built.The result of Matheus and Yoccoz implies that ϕ acts on $H_1^{(0)}(EW, \Sigma, \mathbb{R})$ as an element of finite order.

In plain terms, the action of ϕ on the relative homology is given by a matrix B (a loop in the Rauzy diagram). From the work of Zorich [52] and Forni [23], we know that the action of ϕ^* is responsible for the deviations of ergodic sums at special times of T. By construction f is orthogonal to the Perron-Frobenius eigenvector of B. Thus f is contained in the space generated by the contracting eigendirection of B^t and its central space. We can forget the projection of f on the contracting direction since it does not change the study of Birkhoff sums (see [15] for instance). Thus, up to taking some power of ϕ and considering the projection of f on the central space, we can assume that $B^t f = f$.

Now, equation (2) comes from the fact that $\sum_{i=0}^{r_k^{(p)}} f(S^i y)| = |\langle (B^t)^p f, e_k \rangle| = |\langle f, e_p \rangle|$, and this is smaller than a constant K. This concludes the proof of Theorem 25. After this work was completed we noticed that a similar inequality to (2) can be found in [47]. □

Remark 6. *The same results hold for every EW-interval exchange transformation (resp. Or-interval exchange transformation) as soon as the slope of the corresponding linear flow is irrational. The proof goes along the same lines but is a little bit more technical.*

Corollary 26. *Let L be a non-recurrent language, satisfying an unflipped order condition, such that the language L' of Lemma 18 is $L(S)$ for any EW-interval exchange transformation or Or-interval exchange transformation S in Theorem 25 or Remark 6. Then L is not a natural coding of any affine interval exchange transformation.*

Proof

By Lemma 16, any non-recurrent orbit of X_L coincides with orbits of $X_{L'}$ on an infinite prefix and an infinite suffix. Thus the estimates made in the proof of Theorem 25 on the Birkhoff sums of S contradict the criterion in Theorem 20. □

REFERENCES

[1] V.I. ARNOLD: Small denominators and problems of stability of motion in classical and celestial mechanics, *Usp. Math. Nauk.* 18 (1963), p. 91–192, (in Russian) translated in *Russian Math. Surveys* 18 (1963), p. 86–194.

[2] P. ARNOUX: Un invariant pour les échanges d'intervalles et les flots sur les surfaces, (in French) *Thèse de 3e cycle*, Reims, 1981.

[3] P. ARNOUX, G. RAUZY: Représentation géométrique de suites de complexité $2n + 1$, (in French) *Bull. Soc. Math. France* 119 (1991), p. 199–215.

[4] A. AVILA, G. FORNI: Weak mixing for interval exchange transformations and translation flows. *Ann. of Math.* (2) 165 (2007), p. 637–664.

[5] A.Ya. KANEL-BELOV, A.L. CHERNYAT'EV: Describing the set of words generated by interval exchange transformations, *Comm. Algebra 38* (2010), p. 2588–2605.

[6] J. BERSTEL: Sturmian and episturmian words (a survey of some recent results), in Algebraic informatics, *Lecture Notes in Comput. Sci.* 4728, p. 23–47, Springer, Berlin, 2007.

[7] V. BERTHÉ, C. DE FELICE, F. DOLCE, J. LEROY, D. PERRIN, C. REUTENAUER, G. RINDONE: Acyclic, connected and tree sets, *Monatsh. Math.* 176 (2015), p. 521–550.

[8] M. BOSHERNITZAN: A unique ergodicity of minimal symbolic flows with standard block growth, *J. Analyse Math.* 44 (1984/85), p. 77–96.

[9] M. BOSHERNITZAN, I. KORNFELD: Interval translation mappings, *Ergodic Theory Dynam. Systems* 15 (1995), p. 821–832.

[10] X. BRESSAUD, P. HUBERT, A. MAASS: Persistence of wandering intervals in self-similar affine interval exchange transformations. *Ergodic Theory Dynam. Systems* 30 (2010), p. 665–686.

[11] R. CAMELIER, C. GUTTIEREZ: Affine interval exchange transformations with wandering intervals, *Ergodic Theory Dynam. Systems* 17 (1997), p. 1315–1338.

[12] J. CASSAIGNE: Complexité et facteurs spéciaux, (in French) Journées Montoises (Mons, 1994), *Bull. Belg. Math. Soc. Simon Stevin* 4 (1997), p. 67–88.

[13] J. CASSAIGNE, F. NICOLAS: Factor complexity, *Combinatorics, automata and number theory*, p. 163–247, Encyclopedia Math. Appl., 135, Cambridge Univ. Press, Cambridge, 2010.

[14] J. CHAIKA: Every ergodic transformation is disjoint from almost every interval exchange transformation. *Ann. of Math.* (2) 175 (2012), p. 237–253.

[15] M. COBO: Piece-wise affine maps conjugate to interval exchanges. *Ergodic Theory Dynam. Systems* 22 (2002), p. 375–407.

[16] E.M. COVEN, G.A. HEDLUND: Sequences with minimal block growth, *Math. Systems Theory* 7 (1973), p. 138–153.

[17] A. DE LUCA, M. EDSON, L.Q. ZAMBONI: Extremal values of semi-regular continuants and codings of interval exchange transformations, *Mathematika* 69 (2023), p. 432–457.

[18] A. DENJOY: Sur les courbes définies par les équations differentielles à la surface du tore. *J. Math. Pures Appl.* 11 (1932), p. 333–375.

[19] G. DIDIER: Échanges de trois intervalles et suites sturmiennes, *J. Théor. Nombres Bordeaux,* 9 (1997), p. 463–478.

[20] S. FERENCZI, C. HOLTON, L.Q. ZAMBONI: The structure of three–interval exchange transformations II: a combinatorial description of the trajectories, *J. Anal. Math.* 89 (2003), p. 239–276.

[21] S. FERENCZI, P. HUBERT: Minimality and unique ergodicity of Veech 1969 type interval exchange transformations, arXiv:2103.09018.

[22] S. FERENCZI, L.Q. ZAMBONI: Languages of k-interval exchange transformations, *Bull. Lond. Math. Soc.* 40 (2008), p. 705–714.

[23] G. FORNI: Deviation of ergodic averages for area-preserving flows on surfaces of higher genus, Ann. of Math. (2) 155 (2002), p. 1–103.

[24] G. FORNI, C. MATHEUS: Introduction to Teichmüller theory and its applications to dynamics of interval exchange transformations, flows on surfaces and billiards. *J. Mod. Dyn.* 8 (2014), p. 271–436.

[25] D. GABORIAU, G. LEVITT, F. PAULIN: Pseudogroups of isometries of R and Rips' theorem on free actions on R-trees, *Israël J. Math.* 87 (1994), p. 403–428.

[26] S. GHAZOUANI: Local rigidity for periodic generalised interval exchange transformations, *Invent. Math.* 226 (2021), p. 467–520.

[27] S. GHAZOUANI, C. ULCIGRAI: A priori bounds for GIETS, affine shadows and rigidity of foliations in genus two, arXiv:2106.03529.

[28] G. HALÁSZ: Remarks on the remainder in Birkhoff's ergodic theorem, *Acta Math. Acad. Sci. Hungar.* 28 (1976), p. 389–395.

[29] M.-R. HERMAN: Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations, (French), *Inst. Hautes Études Sci. Publ. Math.* 49 (1979), p. 5–233.

[30] A.B. KATOK, A.M. STEPIN: Approximations in ergodic theory, *Usp. Math. Nauk.* 22 (1967), p. 81–106 (in Russian), translated in *Russian Math. Surveys* 22 (1967), p. 76–102.

[31] M.S. KEANE: Interval exchange transformations, *Math. Zeitsch.* 141 (1975), p. 25–31.

[32] S. KERCKHOFF: Simplicial systems for interval exchange maps and measured foliations, *Ergodic Theory Dynam. Systems* 5 (1985), p. 257–271.

[33] G. LEVITT: La décomposition dynamique et la différentiabilité des feuilletages des surfaces, (French) [Dynamic decomposition and differentiability of surface foliations] *Ann. Inst. Fourier (Grenoble)* 37 (1987), p. 85–116.

[34] A. LINERO BAS, G. SOLER LÓPEZ: Minimal interval exchange transformations with flips, *Ergodic Theory Dynam. Systems* 38 (2018), p. 3101–3144.

[35] I. LIOUSSE: Dynamique générique des feuilletages transversalement affines des surfaces, (French) [Generic dynamics of transversely affine foliations of surfaces] Bull. Soc. Math. France 123 (1995), p. 493–516.

[36] S. MARMI, P. MOUSSA, J.-C. YOCCOZ: Affine interval exchange maps with a wandering interval. Proc. Lond. Math. Soc. (3) 100 (2010), p. 639–669.

[37] S. MARMI, P. MOUSSA, J.-C. YOCCOZ: Linearization of generalized interval exchange maps, *Ann. of Math. (2)* 176 (2012), p. 1583–1646.

[38] H. MASUR, S. TABACHNIKOV: Rational billiards and flat structures. *Handbook of dynamical systems*, Vol. 1A, 1015–1089, North-Holland, Amsterdam, (2002).

[39] C. MATHEUS, J.-C. YOCCOZ: The action of the affine diffeomorphisms on the relative homology group of certain exceptionally symmetric origamis, *J. Mod. Dyn.* 4 (2010), p. 453–486.

[40] M. MORSE, G.A. HEDLUND: Symbolic dynamics II. Sturmian trajectories, *Amer. J. Math.* 62 (1940), p. 1–42.

[41] A. NOGUEIRA: Nonorientable recurrence of flows and interval exchange transformations. *J. Differential Equations* 70 (1987), p. 153–166.

[42] A. NOGUEIRA: Almost all interval exchange transformations with flips are nonergodic, *Ergodic Theory Dynam. Systems* 9 (1989), p. 515–525.

[43] V.I. OSELEDEC: The spectrum of ergodic automorphisms, (in Russian) *Dokl. Akad. Nauk. SSSR* 168 (1966), p. 1009–1011.

[44] N. PYTHEAS FOGG: Substitutions in dynamics, arithmetics and combinatorics, edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel, *Lecture Notes in Mathematics* 1794. Springer-Verlag, Berlin, 2002. xviii+402 pp. ISBN: 3-540-44141-7.

[45] G. RAUZY: Échanges d'intervalles et transformations induites, (in French), *Acta Arith.* 34 (1979), p. 315–328.

[46] M.-L. SANTINI-BOUCHARD: Échanges de trois intervalles et suites minimales, (in French), *Theoret. Comput. Sci* 174 (1997), p. 171-191.

[47] F. TRUJILLO, C. ULCIGRAI: Affine IETs with a singular conjugacy to an IET, preprint 2022.

[48] M. VIANA: Ergodic theory of interval exchange maps. *Rev. Mat. Complut.* 19 (2006), p. 7–100.

[49] W. VEECH: Gauss measures for transformations on the space of interval exchange maps, *Ann. of Math.* (2) 115 (1982), p. 201–242.

[50] W. VEECH: The Metric Theory of interval exchange transformations, *American Journal of Mathematics* 106 (1984), p. 1331–1422.

[51] J. C. YOCCOZ: Interval exchange maps and translation surfaces. Homogeneous flows, moduli spaces and arithmetic, p. 1–69, *Clay Math. Proc., 10, Amer. Math. Soc., Providence, RI,* 2010.

[52] A. ZORICH: Deviation for interval exchange transformations, *Ergodic Theory Dynam. Systems* 17 (1997), p. 1477–1499.

[53] A. ZORICH: Flat surfaces, *Frontiers in number theory, physics, and geometry.* I, p. 437–583, Springer, Berlin, (2006).

AIX MARSEILLE UNIVERSITÉ, CNRS, CENTRALE MARSEILLE, INSTITUT DE MATHÉMATIQUES DE MARSEILLE, I2M - UMR 7373, 13453 MARSEILLE, FRANCE
Email address: ssferenczi@gmail.com

AIX MARSEILLE UNIVERSITÉ, CNRS, CENTRALE MARSEILLE, INSTITUT DE MATHÉMATIQUES DE MARSEILLE, I2M - UMR 7373, 13453 MARSEILLE, FRANCE
Email address: pascal.hubert@univ-amu.fr

UNIVERSITE CLAUDE BERNARD LYON 1, ICJ UMR5208, CNRS, ECOLE CENTRALE DE LYON, INSA LYON, UNIVERSITÉ JEAN MONNET,, 69622 VILLEURBANNE, FRANCE
Email address: zamboni@math.univ-lyon1.fr