Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 22–25

A PROFINITE APPROACH TO COMPLETE BIFIX DECODINGS OF RECURRENT LANGUAGES

ALFREDO COSTA

Abstract: We approach the study of complete bifix decodings of (uniformly) recurrent languages with the help of the free profinite monoid. We show that the complete bifix decoding of a uniformly recurrent language F by an F-charged rational complete bifix code is uniformly recurrent. An analogous result is obtained for recurrent languages. As an application, we show that the Schützenberger group of an irreducible symbolic dynamical system is an invariant of eventual conjugacy.

Keywords: Code, complete bifix decoding, free profinite monoid, profinite group, symbolic dynamics, subshift, eventual conjugacy.
Math. Subject Classification (2020): 20M05, 68Q70, 94B35, 20E18, 37B10.

1. Introduction

In this paper we see symbolic dynamical systems, codes, and free profinite monoids in interplay. The following introductory paragraphs provide a first picture of how these fields come together in this work. Several details are deferred to later sections or to the bibliography. We indicate the books [LM95, Fog02] as references for symbolic dynamics; the book [BPR10] for codes; and the book [ACKP20b] for an introduction to free profinite monoids and their connections with codes and symbolic dynamics.

When dealing with a subshift $\mathcal{S} \subseteq A^{\mathbb{Z}}$, we are often led to decode $\mathcal{S}$ into a new symbolic dynamical system $\mathcal{S}' \subseteq X^{\mathbb{Z}}$. This decoding is quite often better understood in terms of sets of finite words. From that viewpoint, we decode the language $F \subseteq A^*$ of blocks of elements of $\mathcal{S}$ into a language $F' \subseteq X^*$. An important example is the following: for a positive integer n, and for $Z = A^n$, let $F' = F \cap Z^*$; then F' is the language of blocks of a subshift, called the *n-th higher power* of $\mathcal{S}$, over the finite alphabet $X = F \cap Z$. Two subshifts are said to be *eventually conjugate* if for all sufficiently large n their n-th higher powers are conjugate. This notion received great attention

Received August 24, 2022.
Work carried out in part at City College of New York, CUNY, whose hospitality is gratefully acknowledged, with the support of the FCT sabbatical scholarship SFRH/BSAB/150401/2019, and it was partially supported by the Centre for Mathematics of the University of Coimbra - UIDB/00324/2020, funded by the Portuguese Government through FCT/MCTES.

since introduced by Williams, who conjectured that for finite type subshifts eventual conjugacy coincides with conjugacy [Wil73]. This conjecture was later disproved by Kim and Roush [KR92, KR99].

We now look at the following key notion in the algebraic theory of codes: a code Z over the alphabet A is *complete* if every element of A^* is the prefix of some element of Z^* and the suffix of some element of Z^*. The set $Z = A^n$ is an example of a complete bifix code. The 2012 seminal paper [BDFP$^+$12] prompted a research line giving center stage to the following operation: take a recurrent subset F of A^* (that is, F is the language of blocks of an irreducible subshift of $A^{\mathbb{Z}}$), and a complete bifix code Z, and investigate the properties of the intersection $X = F \cap Z$. For example, in [BDFP$^+$12] it is shown that X is finite if F is uniformly recurrent (that is, if F is the language of blocks of a minimal subshift of $A^{\mathbb{Z}}$). Since Z is a code, an element of $F \cap Z^*$ can be seen uniquely as an element of the free monoid over the finite alphabet $X = F \cap Z$, thus making $F \cap Z^*$ the language of a symbolic dynamical system over the alphabet X. The language $F \cap Z^*$ is then said to be a *complete bifix decoding* of F. In general, the hypothesis that F is uniformly recurrent does not imply that $F \cap Z^*$ is recurrent (over the alphabet X), even in the case where $Z = A^n$. One of the main results of [BDFD$^+$15c] states that if F is a uniformly recurrent dendric language (a class of languages that includes the famous Sturmian languages and whose precise definition we recall in Section 3), then every complete bifix decoding of F is also a uniformly recurrent dendric language. It turns out that in the proof of that result, one of the most "hairy points" (to borrow the words used by Perrin in his DLT 2018 survey [Per18]) is showing that the complete bifix decoding of a uniformly recurrent dendric language is indeed uniformly recurrent.

The application of the free profinite monoid within the line of research inaugurated by Berstel et al. in the aforementioned paper [BDFP$^+$12] was first made in [KP17, ACKP20a], in the study of the group of a complete bifix code. That "profinite" approach is centered around the notion of *F-charged code*, a notion whose precise definition involves a maximal subgroup of the free profinite monoid depending on the recurrent language F. In this paper we use that approach to significantly extend the property that all complete bifix decodings of uniformly recurrent dendric languages are also uniformly recurrent (the aforesaid "hairy point", for which this paper gives therefore a new proof). More concretely: we show that the decoding of any uniformly recurrent language F by an F-charged rational complete bifix code is always uniformly recurrent (Theorem 4.5). A similar general result concerning the preservation of recurrence is obtained in Theorem 4.6.

We now say some more words about the maximal subgroup of the free profinite monoid briefly mentioned in the previous paragraph. Let $\mathcal{S}$ be an irreducible subshift of $A^{\mathbb{Z}}$. The *Schützenberger group of* $\mathcal{S}$, denoted $G(\mathcal{S})$, is a profinite group, introduced by Almeida in [Alm05b], naturally located as a maximal subgroup of the free profinite monoid over A, with the elements of $G(\mathcal{S})$ being certain limits of blocks of elements of $\mathcal{S}$. The invariance of $G(\mathcal{S})$ under conjugacy was first proved in [Cos06]. As an application of our main result about complete bifix decodings of recurrent languages, applied to the special case of bifix codes of the form $Z = A^n$, we show that $G(\mathcal{S})$ is an invariant of eventual conjugacy (Theorem 10.10). In [CS21] it is shown that $G(\mathcal{S})$ is invariant under *flow equivalence*, a relation coarser than conjugacy. It's known that eventual conjugacy of irreducible subshifts of finite type implies flow equivalence (this is a corollary of the classification of irreducible subshifts of finite type up to flow equivalence [Fra84, PS75, BF77], as the reader may check when reading the book of Lind and Marcus [LM95], namely its Sections 3.4 and 13.6, or when reading Boyle's expository article [Boy02]). It is an open question as to whether this implication holds for all subshifts of finite type.* For the whole class of irreducible subshifts, it is unclear when the implication holds.

Since their introduction, the groups of the form $G(\mathcal{S})$ contributed significantly to the study of the structure of free profinite monoids. In [AV06] it is shown that $G(\mathcal{S})$ is a free procyclic group if $\mathcal{S}$ is periodic. The main result of [CS11] states that if $\mathcal{S}$ is a nonperiodic irreducible sofic subshift, then $G(\mathcal{S})$ is a free profinite group with rank $\aleph_0$; interestingly, the invariance under conjugacy of $G(\mathcal{S})$ was used in the proof. The case of minimal nonperiodic subshifts has a much more rich landscape, far from being fully understood. When $\mathcal{S}$ is minimal, then $G(\mathcal{S})$ may be a free profinite group or not. For example, $G(\mathcal{S})$ is a free profinite group of rank the size of the alphabet when $\mathcal{S}$ is dendric [AC16], but if $\mathcal{S}$ is the Thüe-Morse subshift then $G(\mathcal{S})$ is not free [AC13]. See [GO21, GO22] for criteria helping to decide if $G(\mathcal{S})$ is free or not, when $\mathcal{S}$ is defined by a primitive substitution φ. These criteria are associated to finite algebraic invariants of $G(\mathcal{S})$ that may be easily computed

*The implication for general subshifts of finite type was stated without proof as a "difficult result" in Boyle's 1997 Temuco lectures [Boy00, Section 3.7.9]. The author withdraws the claim now [Boy22], as he has not been able to recover or produce a proof.

from the matrix associated to φ (cf. comments at the end of [GO22, Section 4.4]). We now know that these algebraic invariants are invariants of eventual conjugacy of the dynamical system $\mathcal{S}$.

Finally, we mention that our proof of the invariance under eventual conjugacy of $G(\mathcal{S})$ allows, with minimal adaptations, a generalization involving relatively free profinite monoids (Theorem 10.12).

2. Subshifts, free profinite monoids, and connections between them

Along this paper, all alphabets are assumed to be nonempty finite sets.

2.1. Subshifts. Consider an alphabet A. A subset F of the free monoid A^* is said to be *factorial* if it is closed under taking factors, and is said to be *prolongable* if for all $u \in F$ there are $a, b \in A$ such that au and ub belong to F.

Let $x = (x_i)_{i\in\mathbb{Z}} \in A^{\mathbb{Z}}$. A word of the form $x_i \cdots x_{i+n}$, with $i \in \mathbb{Z}$ and $n \in \mathbb{N}$, is a *block* of x. We also allow the empty word to be a block of x. We denote by $\mathcal{B}(x)$ the set of blocks of x. Consider in $A^{\mathbb{Z}}$ the product topology, with A having the discrete topology. Since we are assuming A finite, the space $A^{\mathbb{Z}}$ is compact, by Tychonoff's theorem. (The Hausdorff property is being included in the definition of compact space.) Consider the shift map $\sigma_A \colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$, given by $\sigma_A((x_i)_{i\in\mathbb{Z}}) = (x_{i+1})_{i\in\mathbb{Z}}$. A *symbolic dynamical system*, or *subshift*, of $A^{\mathbb{Z}}$ is a nonempty closed subspace $\mathcal{S}$ of $A^{\mathbb{Z}}$ such that $\sigma_A(\mathcal{S}) = \mathcal{S}$. Given $F \subseteq A^*$, let $\mathcal{S}_F = \{x \in A^{\mathbb{Z}} \mid \mathcal{B}(x) \subseteq F\}$. The mapping $F \mapsto \mathcal{S}_F$ is a bijection from the set of nonempty factorial prolongable languages of A^* to the set of subshifts of $A^{\mathbb{Z}}$, whose inverse maps a subshift $\mathcal{S} \subseteq A^{\mathbb{Z}}$ to the language $\mathcal{B}(\mathcal{S}) = \bigcup_{x\in\mathcal{S}} \mathcal{B}(x)$. In view of this, we say that a nonempty factorial prolongable language is a *subshift language*.

Let F be a subshift language of A^*. One says that F is *recurrent* (over the alphabet A) when for every $u, v \in F$ there is $w \in F$ such that $uwv \in F$. Also, F is *uniformly recurrent* (over A) when, for every $u \in F$, there is an integer N such that u is a factor of every word of F with length at least N. Every uniformly recurrent language is recurrent. These definitions have dynamical meaning: a subshift $\mathcal{S}$ has a dense positive forward orbit (one says that such a subshift is *irreducible*) if and only if $\mathcal{B}(\mathcal{S})$ is recurrent; and $\mathcal{S}$ is a minimal subshift for the inclusion (equivalently, every orbit in $\mathcal{S}$ is dense in $\mathcal{S}$) if and only if $\mathcal{B}(\mathcal{S})$ is uniformly recurrent.

We refer to the books [LM95, Fog02] for information about subshifts. The first book is more focused on subshifts whose language of blocks is rational, the *sofic* subshifts (which include the *finite type subshifts* already mentioned at the introduction of this paper), while in the second book we find extensive information about a special class of minimal subshifts, the *primitive substitutive* subshifts. The two books complement each other, as minimal subshifts and irreducible sofic subshifts have in common only the periodic subshifts. Primitive substitutive subshifts may be defined as follows. Let φ be an endomorphism of A^*, where the finite set A has at least two elements. If there is a positive integer n such that for every $a, b \in A$, the letter b is a factor of $\varphi^n(a)$, then φ is called a *primitive substitution*; in that case the set F_φ of factors of words of the form $\varphi^k(a)$, with $k \in \mathbb{N}$ and $a \in A$, is indeed a uniformly recurrent language.

2.2. Profinite monoids and their Green's relations. For books about profinite monoids see e.g. [RS09], [ACKP20b], and [RZ10] for profinite groups. For the sake of fixing terminology and notation, we recall some basic definitions and results. We begin with the Green quasi-orders in a monoid M: for every $x, y \in M$, we have

- $x \leq_{\mathcal{J}_M} y$ if and only if y is a factor of x (that is, if and only if $x \in MyM$);
- $x \leq_{\mathcal{R}_M} y$ if and only if y is a prefix of x (that is, if and only if $x \in yM$);
- $x \leq_{\mathcal{L}_M} y$ if and only if y is a suffix of x (that is, if and only if $x \in My$).

The corresponding induced Green's equivalence relations are $\mathcal{J}_M$, $\mathcal{R}_M$ and $\mathcal{L}_M$. We also consider the intersection $\mathcal{H}_M = \mathcal{R}_M \cap \mathcal{L}_M$. The subscript M in $\mathcal{K}_M$ may be dropped, for $\mathcal{K} \in \{\mathcal{J}, \mathcal{R}, \mathcal{L}, \mathcal{H}\}$. An $\mathcal{H}_M$-class H contains an idempotent if and only if it is a subgroup of M (that is, a subsemigroup of M with a group structure), and in fact every subgroup of M is contained in an $\mathcal{H}_M$-class, for which reason such $\mathcal{H}_M$-classes are referred to as the *maximal subgroups* of M.

A *compact monoid* is a monoid endowed with a compact topology for which the multiplication is continuous. If M is a compact monoid, then M is a *stable monoid*, meaning that $\leq_{\mathcal{R}_M} \cap \mathcal{J}_M = \mathcal{R}_M$ and $\leq_{\mathcal{L}_M} \cap \mathcal{J}_M = \mathcal{L}_M$. In a compact monoid M, all Green's relations are closed, and so in particular maximal subgroups are closed subgroups.

A *profinite monoid* is an inverse limit of finite monoids in the category of compact monoids (we view finite monoids as compact monoids). All closed subgroups of a profinite monoid are profinite groups, that is, inverse limits of finite groups.

Let M be a profinite monoid. For every $x \in M$ and $k \in \mathbb{Z}$, the sequence $(x^{n!+k})_{n\geq|k|}$ converges to an element of M denoted $x^{\omega+k}$, and taking $k = 0$ we get the idempotent $x^\omega = \lim x^{n!}$. An element x in a monoid M is *regular* if $x \in xMx$. For $\mathcal{K} \in \{\mathcal{J}, \mathcal{R}, \mathcal{L}\}$, a $\mathcal{K}_M$-class contains a regular element if and only if it contains only regular elements, if and only if it contains some idempotent. A $\mathcal{K}_M$-class with regular elements is called a *regular $\mathcal{K}_M$-class*. All maximal subgroups contained in a regular $\mathcal{J}_M$-class J are isomorphic profinite groups, and so they can be identified as the same profinite group, referred to as the *Schützenberger group* of J.

2.3. Free profinite monoids. In the following lines, we mostly follow the notation and approach of [ACKP20b].

If u, v are distinct elements of the free monoid A^*, then there is some homomorphism $\varphi\colon A^* \to N$ onto a finite monoid N satisfying $\varphi(u) \neq \varphi(v)$. Let $r(u, v)$ be the smallest possible cardinal for such a monoid N. Consider the completion $\widehat{A^*}$ of A^* under the metric d defined, when $u \neq v$, by $d(u, v) = 2^{-r(u,v)}$. The free monoid A^* embeds as a topological subspace of $\widehat{A^*}$. Moreover, the monoid structure of A^* extends uniquely to a monoid structure on $\widehat{A^*}$, making $\widehat{A^*}$ a compact monoid. In fact, $\widehat{A^*}$ is a profinite monoid, known as the *free profinite monoid generated by* A because of the following universal property: for every map $\varphi\colon A \to M$ into a profinite monoid M, there is a unique continuous homomorphism $\hat{\varphi}\colon \widehat{A^*} \to M$ extending φ.

We view the elements of $\widehat{A^*}$ as generalizations of words, for which reason each element of $\widehat{A^*}$ is called a *pseudoword* over A. The elements of A^* are isolated points in $\widehat{A^*}$, and they are the only isolated points there, as A^* is dense in $\widehat{A^*}$. Each element of $\widehat{A^*} \setminus A^*$ is said to be an *infinite pseudoword*, while each word in A^* is a *finite pseudoword*. The set $\widehat{A^*} \setminus A^*$ is an ideal of $\widehat{A^*}$.

We may also consider profinite semigroups, *mutatis mutandis*. It turns out that the *free profinite semigroup* generated by A, denoted $\widehat{A^+}$, is (isomorphic) to $\widehat{A^*} \setminus \{\varepsilon\}$, where ε is the empty word.

The next theorem gives a hint for why free profinite monoids are important.

Theorem 2.1 (cf. [Alm95, Theorem 3.6.1]). *A language $L \subseteq A^*$ is rational if and only if its topological closure $\overline{L}$ in $\widehat{A^*}$ is open, if and only if $L = K \cap A^*$ for some clopen subset K of $\widehat{A^*}$.*

We next refer to some properties of pseudowords necessary along the paper. We begin with a cancellation property going back to [Alm95, Exercise 10.2.10]. A recent proof can be found in [ACKP20b, Solution to Exercise 4.20].

Proposition 2.2. *Let N be a positive integer. If $u, v \in A^*$ have length N and $x, y \in \widehat{A^*}$ are such that $xu = yv$ or $ux = vy$, then $x = y$ and $u = v$.*

Therefore, for every $N \in \mathbb{N}$, each infinite pseudoword x of $\widehat{A^*}$ has a unique prefix and a unique suffix of length N. In particular, taking $N = 1$, every pseudoword which is not the empty word has a unique "first letter" and a unique "last letter".

For the sake of simplicity, we may write $\mathcal{K}_A$ instead of $\mathcal{K}_{\widehat{A^*}}$, whenever $\mathcal{K}$ is one of Green's relations $\mathcal{R}, \mathcal{L}, \mathcal{J}, \mathcal{H}$.

Corollary 2.3. *Let u be a regular element of $\widehat{A^*}$. If the factorization $u = vs$ is such that $s \in A^*$, then $u \mathrel{\mathcal{R}_A} v$.*

Proof: Let $z \in \widehat{A^*}$ be such that $u = uzu$. From $u = vs = uzvs$ and Proposition 2.2 we get $v = uzv$, thus $u \mathrel{\mathcal{R}_A} v$. ∎

Corollary 2.4. *Suppose that yx is an idempotent in $\widehat{A^*}$, with $y \in A^*$. Then xy is an idempotent of $\widehat{A^*}$ which is $\mathcal{J}_A$-equivalent to yx.*

Proof: From $yx = yxyx$ and Proposition 2.2 we get $x = xyx$. Multiplying by y on the left of both sides of $x = xyx$, we conclude that xy is idempotent. ∎

Remark 2.5. If x, y are arbitrary elements of a monoid M such that xy is idempotent, then $(yx)^2$ is idempotent, but in general yx may not be idempotent.

Extending what we do for subsets of A^*, we say that a subset K of $\widehat{A^*}$ is *factorial* if it contains the factors of elements of K; that it is *prolongable* if for every $u \in K$ there are $a, b \in A$ such that au and $ub \in K$; and that K is *recurrent* if it is factorial, strictly contains the set $\{\varepsilon\}$, and $u, v \in K$ implies the existence of some $w \in \widehat{A^*}$ such that $uwv \in K$ (note that K is prolongable if K is recurrent). For a proof of the next proposition, see [ACKP20b, Proposition 5.6.1].

Proposition 2.6. *Let F be a subset of A^*. If F is factorial/prolongable/recurrent subset of A^*, then $\overline{F}$ is a factorial/prolongable/recurrent subset of $\widehat{A^*}$, respectively.*

Let K be a nonempty closed subset of $\widehat{A^*}$. Since the relation $\leq_{\mathcal{J}_A}$ is topologically closed in $\widehat{A^*} \times \widehat{A^*}$ and $\widehat{A^*}$ is a compact space, every $\leq_{\mathcal{J}_A}$-chain of elements of K has a lower bound in K, and so, by Zorn's lemma, the set K contains $\leq_{\mathcal{J}_A}$-minimal elements (that is, elements of K that are minimal for the restriction of $\leq_{\mathcal{J}_A}$ to K).

Remark 2.7. For every nonempty closed subset K of $\widehat{A^*}$, every element of K is a factor of a $\leq_{\mathcal{J}_A}$-minimal element of K: indeed, if $u \in K$, then the set $K_u = \{v \in K \mid v \leq_{\mathcal{J}_A} u\}$ is itself nonempty and closed, and so it contains $\leq_{\mathcal{J}_A}$-minimal elements, which are clearly $\leq_{\mathcal{J}_A}$-minimal elements of K.

For a nonempty factorial language $F \subseteq A^*$, we denote by $J_A(F)$ the set of $\leq_{\mathcal{J}_A}$-minimal elements of $\overline{F} \subseteq \widehat{A^*}$. Note that $J_A(F)$ is contained in $\overline{F}$ and is a union of $\mathcal{J}_A$-classes, since $\overline{F}$ is factorial (cf. Proposition 2.6). If the alphabets A, B satisfy $A \subseteq B$, then we see $\widehat{A^*}$ as a closed submonoid of $\widehat{B^*}$, and under that perspective we have $J_A(F) = J_B(F)$. Sometimes it will be convenient to take the smallest alphabet A for which F is a language of A^*. For a language $F \subseteq B^*$ containing nonempty words, we refer to the subset $A \subseteq B$ of letters that are factors of some element of F as the *alphabet of* F.

If the language F is recurrent over the alphabet A, then $J_A(F)$ is a regular $\mathcal{J}_A$-class [ACKP20b, Exercise 5.25]. If F is uniformly recurrent, then $J_A(F)$ is actually $\leq_{\mathcal{J}}$-maximal among regular $\mathcal{J}$-classes, as seen next. An infinite pseudoword u in $\widehat{A^*}$ is a *$\mathcal{J}$-maximal infinite pseudoword* if $u <_{\mathcal{J}_A} v$ implies $v \in A^*$.

Theorem 2.8. *If $F \subseteq A^*$ is uniformly recurrent, then $\overline{F} = F \cup J_A(F)$. Moreover, the mapping $F \mapsto J_A(F)$ is a bijection from the set of uniformly recurrent languages of A^* onto the set of $\mathcal{J}_A$-classes of $\mathcal{J}_A$-maximal infinite pseudowords of $\widehat{A^*}$.*

Theorem 2.8 is from [Alm05a]. A proof can be found in [ACKP20b, Propositions 5.6.12 and 5.6.13].

3. Codes

A *code* over the alphabet A is a nonempty subset X of A^+ that freely generates a submonoid of A^*. For example, we have the *prefix codes* (respectively, *suffix codes*), that is, nonempty subsets X of A^+ with no two elements u, v of X such that u is a prefix (respectively, suffix) of v. A *bifix code* is a code both prefix and suffix.

For dealing with the topological closure in $\widehat{A^*}$ of a rational code $X \subseteq A^+$, the following proposition is of great help.

Proposition 3.1 ([ACKP20a, Proposition 2.21])**.** *Let X be a rational code contained in A^+. In what follows, u, v, w are arbitrary pseudowords in $\widehat{A^*}$. If X is a code, then the implication*

$$u, vw, uv, w \in \overline{X^*} \Longrightarrow v \in \overline{X^*} \tag{3.1}$$

holds. Moreover, if X is a prefix code, then

$$u, uv \in \overline{X^*} \Longrightarrow v \in \overline{X^*} \tag{3.2}$$

also holds.

3.1. Complete codes. Consider a language $F \subseteq A^*$. A subset X of A^+ is *right F-complete* if every element of F is a prefix of an element of X^*. A prefix code X contained in F is an *F-maximal prefix code* if whenever Y is a prefix code with $X \subseteq Y \subseteq F$, one has $Y = X$. Combining the statements of [BDFP$^+$12, Propositions 3.3.1 and 3.3.2]) we get the following.

Proposition 3.2. *Let $F \subseteq A^*$ be a factorial language and let X be a prefix code contained in F. Then X is an F-maximal prefix code if and only if it is right F-complete.*

Example 3.3. The *Fibonnaci language* over $\{a, b\}$ is the uniformly recurrent language $F = F_\varphi$ induced by the primitive substitution φ (the *Fibonnaci substitution*) given by $\varphi(a) = ab$ and $\varphi(b) = a$. The prefix code $X = \{a, ba\}$ is right F-complete.

Of course, one has dual definitions of left F-complete set, F-maximal suffix code, and the corresponding counterpart of Proposition 3.2.

A code X contained in F is *F-complete* when it is both right F-complete and left F-complete. A bifix code X contained in the language $F \subseteq A^*$ is an *F-maximal bifix code* if whenever Y is a bifix code with $X \subseteq Y \subseteq F$, one has $Y = X$. It turns out that if F is recurrent, then the rational F-maximal bifix codes are precisely the rational F-complete bifix codes, as seen next.

Theorem 3.4. *Let F be a recurrent language, and let X be a rational bifix code contained in F. Then the following conditions are equivalent:*

(1) X is an F-maximal bifix code;
(2) X is left F-complete;
(3) X is right F-complete;
(4) X is F-complete.

Theorem 3.4 is from the article [BDFP$^+$12, Theorem 4.2.2]. In that paper it was assumed more generally that the bifix code X is *F-thin*, a property that is satisfied when X is rational (cf. [ACKP20a, Proposition 2.8]).

When it is clear which alphabet A we are talking about, when $F = A^*$ we may drop the "F-" in the previous definitions, writing simply "maximal prefix code", "complete code", etc..

A *group code* is a code Z with alphabet A for which the syntactic monoid of Z^* is a finite group.† Every group code is a rational maximal bifix code, see for example [BDFP$^+$12, Proposition 6.1.5].

Example 3.5. For every positive integer n, the language $A^n \subseteq A^+$ is a group code.

There is a close connection between complete and F-complete codes.

Theorem 3.6 (cf. [BDFP$^+$12, Theorems 4.2.11 and 4.4.3])**.** *Consider a recurrent language $F \subseteq A^*$. If Z is a rational complete bifix code of A^*, then $Z \cap F$ is an F-complete bifix code. Moreover, if F is uniformly recurrent, then $Z \cap F$ is finite.*

Remark 3.7. Assume that $F \subseteq A^*$ is recurrent. By [BPR10, Theorem 6.6.1], every rational bifix code $X \subseteq A^+$ is contained in a rational complete bifix code $Z \subseteq A^+$. For such Z, the bifix code $Z \cap F$ is F-complete by Theorem 3.6, thus an F-maximal bifix code by Theorem 3.4. Therefore, if X is a rational F-complete bifix code, then $X = Z \cap F$ for some rational complete bifix code Z. This fact is a sort of converse to Theorem 3.6.

For $L \subseteq A^*$, let η_L be the syntactic homomorphism from A^* onto the syntactic monoid $M(L)$ of L. If L is rational, then, as $M(L)$ is then finite, we may consider the unique continuous homomorphism $\hat{\eta}_L \colon \widehat{A^*} \to M(L)$ extending η_L. In this paper we take advantage of the following fact.

†The definition of group code in the book [BPR10] is more relaxed. We are following the definition of group code employed in the seminal paper [BDFP$^+$12]; the same definition is used in other sources that we cite, such as [ACKP20a, ACKP20b].

Proposition 3.8. *Let $F \subseteq A^*$. Suppose that X is a rational F-complete bifix code. If w is an element of $\overline{F} \subseteq \widehat{A^*}$ such that $\hat{\eta}_{X^*}(w)$ is idempotent, then $w \in \overline{X^*}$.*

A proof of Proposition 3.8 can be found in [ACKP20b, Solution to Exercise 8.16].[‡]

3.2. Charged codes. Consider a recurrent language $F \subseteq A^*$. Since $J_A(F)$ is a regular $\mathcal{J}$-class of $\widehat{A^*}$, it contains maximal subgroups, which are all isomorphic profinite groups. Let $Z \subseteq A^+$ be a rational complete bifix code. We say that Z is *F-charged* if for every maximal subgroup $K \subseteq J_A(F)$ the image $\hat{\eta}_{Z^*}(K)$ is a maximal subgroup contained in the minimum ideal of $M(Z^*)$ (the minimum ideal is the $\mathcal{J}$-class containing all $\leq_{\mathcal{J}}$-minimal elements). We also say that the rational code $X \subseteq A^+$ is *F-complete F-charged* if it is F-complete and there is an F-charged bifix code $Z \subseteq A^+$ such that $X = Z \cap F$.

As a motivation for our main results, we exhibit until the end of this section examples of F-charged group codes.

We recall some basic notions of profinite group theory, available in [RZ10]. Let H be a *formation* of finite groups, i.e. a class of finite groups closed under taking quotients and finite subdirect products, and containing some nontrivial group. A *pro*-H group is an inverse limit of groups from H, with onto connecting morphisms. The class of A-generated pro-H groups has a free object $\widehat{F}_{\mathsf{H}}(A)$, the *free pro-$\mathsf{H}$ group generated by A*. If H has nontrivial groups, then A embeds as a generating subset of the profinite group $\widehat{F}_{\mathsf{H}}(A)$. Denote p_{H} the continuous onto homomorphism $\widehat{A^*} \to \widehat{F}_{\mathsf{H}}(A)$ that fixes the elements of A.

Let us say that a recurrent language F with alphabet A is *H-charging* if for every maximal subgroup K of $J_A(F)$ (equivalently, for some maximal subgroup K of $J_A(F)$) the equality $p_{\mathsf{H}}(K) = \widehat{F}_{\mathsf{H}}(A)$ holds. Let us also say that a code Z is an *H-code* if its syntactic monoid $M(Z^*)$ belongs to H.

Proposition 3.9. *Let $F \subseteq A^*$ be an H-charging recurrent language. Then every H-code $Z \subseteq A^+$ is F-charged.*

Proof: By the universal property of $\widehat{F}_{\mathsf{H}}(A)$, the hypothesis $M(Z^*) \in \mathsf{H}$ implies that $\hat{\eta}_{Z^*} \colon \widehat{A^*} \to M(Z^*)$ factorizes as $\hat{\eta}_{Z^*} = \bar{\eta}_{Z^*} \circ p_{\mathsf{H}}$, for some continuous

[‡]In the statement of [ACKP20b, Exercise 8.16] it is only implicit that the code X is rational.

onto homomorphism $\bar{\eta}_{Z^*}\colon \widehat{F}_{\mathsf{H}}(A) \to M(Z^*)$. Therefore, if K is a maximal subgroup of $\widehat{A^*}$ contained in $J_A(F)$, then $\hat{\eta}_{Z^*}(K) = \bar{\eta}_{Z^*}(\widehat{F}_{\mathsf{H}}(A)) = M(Z^*)$. ∎

We denote by G the formation of all finite groups. The pro-G groups are the profinite groups, and the G-codes are the group codes.

Example 3.10. The *dendric* languages, also known as *tree* languages, were introduced in [BDFD⁺15b]. Their definition is as follows. The *extension graph* of a word w in a language $F \subseteq A^*$ is the bipartite graph $G_F(w)$ where vertices are partitioned into disjoint copies $1 \otimes L_F(w)$ and $R_F(w) \otimes 1$ of the sets $L_F(w) = \{a \in A \mid aw \in F\}$ and $R_F(w) = \{a \in A \mid wa \in F\}$, and where an edge from $1 \otimes a$ to $b \otimes 1$ is a pair $(a, b) \in A \times A$ such that $awb \in F$. The set F is *dendric* (respectively, *connected*) if $G_F(w)$ is a tree (respectively, connected graph) for every $w \in F$. In particular, every dendric language is connected. It turns out that every uniformly recurrent connected language is G-charging, a property shown in [ACKP20a, Theorem 2.19], basically repeating the proof in [AC16] for the dendric case.

For exhibiting another meaningful class of charged codes, it is convenient to introduce some more material about substitutions. Let φ be a primitive substitution over the alphabet A. Let us say that φ is *stable* if there is $k \in \mathbb{N}$ such that, for all $a, b \in A$, all factors of length two of $\varphi^k(ab)$ belong to F_φ (equivalently, if x is the last letter of $\varphi^k(a)$ and y is the first letter of $\varphi^k(b)$, then $xy \in F_\varphi$).

Example 3.11. The Fibonnaci substitution (Example 3.3) is stable, since we have $\varphi^2(\{a, b\}^2) \subseteq F_\varphi$.

We proceed to give in the next lemma a "profinite" characterization of stable primitive substitutions, which is the characterization we need. We denote by $\hat{\varphi}$ the unique continuous endomorphism of $\widehat{A^*}$ extending the endomorphism $\varphi\colon A^* \to A^*$. The monoid of continuous endomorphisms of a finitely generated profinite monoid is itself a profinite monoid for the pointwise topology (see [ACKP20b, Sections 3.12 and 3.15] for information about this result first obtained by Hunter [Hun83]). In particular, the idempotent continuous idempotent endomorphism $\hat{\varphi}^\omega\colon \widehat{A^*} \to \widehat{A^*}$ in the following lemma makes sense.

Lemma 3.12. *A primitive substitution $\varphi\colon A^* \to A^*$ is stable if and only if $\hat{\varphi}^\omega(A^+)$ is contained in $J(F_\varphi)$.*

Proof: Suppose φ is stable. Let $u \in A^+$. We show by induction on the length of u that $\hat{\varphi}^\omega(u) \in J(F_\varphi)$. The base case $u \in A$ holds for every primitive substitution, not necessarily stable ([Alm05a, Theorem 3.7], cf. [ACKP20b, Section 5.6]).

Proceeding with the inductive step, assume that $\hat{\varphi}^\omega(u) \in J(F_\varphi)$. Let $b \in A$. To show that $\hat{\varphi}^\omega(ub) \in J(F_\varphi)$, we use the following criterion: an infinite pseudoword $w \in \widehat{A^*}$ belongs to $J(F_\varphi)$ if and only if every finite factor of w belongs to F_φ ([Alm05a, Corollary 2.8], cf. [ACKP20b, Section 5.6.9]). Take a finite factor v of $\hat{\varphi}^\omega(ub)$. Then v is a finite factor of $\hat{\varphi}^\omega(u)$, or a finite factor of $\hat{\varphi}^\omega(b)$, or, if a is the last letter of u, the concatenation of a finite suffix of $\hat{\varphi}^\omega(a)$ with a finite prefix of $\hat{\varphi}^\omega(b)$, as words are placed in products of pseudowords as in the case of products of finite words (cf. [ACKP20b, Example 4.4.19]).

We claim that $v \in F_\varphi$. If $v \in \hat{\varphi}^\omega(u)$ or $v \in \hat{\varphi}^\omega(a)$, then, by the induction hypothesis, v is a factor of an element of $J(F_\varphi)$. Since the set $\overline{F_\varphi}$ is factorial and $J(F_\varphi) \subseteq \overline{F_\varphi}$ (cf. Proposition 2.6), we immediately obtain $v \in F_\varphi$.

It remains to look at the case where v is the concatenation of a finite suffix of $\hat{\varphi}^\omega(a)$ with a finite prefix of $\hat{\varphi}^\omega(b)$. Since φ is stable, there is $k \in \mathbb{N}$ such that if x is the last letter of $\varphi^k(a)$ and y is the first letter of $\varphi^k(b)$, then $xy \in F_\varphi$. Note that $\varphi^{\omega-k}(x)$ is an infinite suffix of $\varphi^{\omega-k}(\varphi^k(a)) = \varphi^\omega(a)$, and likewise $\varphi^{\omega-k}(y)$ is a infinite prefix of $\varphi^\omega(b)$. Therefore, v is a factor of $\varphi^{\omega-k}(xy)$. As $\varphi(F_\varphi) \subseteq F_\varphi$ and since $xy \in F_\varphi$, we know that $\varphi^{\omega-k}(xy) \in \overline{F_\varphi}$, and so $v \in F_\varphi$, as $\overline{F_\varphi}$ is factorial.

Hence, in all cases, $v \in F_\varphi$. We have thus shown that all finite factors of $\varphi^\omega(ub)$ belong to F_φ, which we already mentioned is the same as having $\varphi^\omega(ub) \in J(F_\varphi)$. This concludes the inductive proof of the "only if" part of the statement.

Conversely, suppose $\hat{\varphi}^\omega(A^+) \subseteq J(F_\varphi)$. Take $a, b \in A$. Let x and y respectively be the last letter of $\varphi^\omega(a)$ and the first letter of $\varphi^\omega(b)$. Since $\overline{F_\varphi}$ is factorial and xy is a factor of $\varphi^\omega(ab)$, it follows from our hypothesis that $xy \in F_\varphi$. Because $\varphi^{n!}(a) \to \hat{\varphi}^\omega(a)$ and $\widehat{A^*}x$ is clopen, x is a suffix of $\varphi^{n!}(a)$ for all sufficiently large n. Similarly, y is a prefix of $\varphi^{n!}(b)$ for all sufficiently large n. As $xy \in F_\varphi$, this shows that every factor of length two of $\varphi^{n!}(ab)$ belongs to F_φ for all sufficiently large n and all $a, b \in A$. This establishes that φ is stable. ∎

A substitution φ over the alphabet A is said to be *proper* if there are $b, c \in A$ such that $\varphi(A) \subseteq bA^* \cap A^*c$. If φ is a proper primitive substitution, then $\hat{\varphi}^\omega(A^+)$ is contained in a maximal subgroup of $J(F_\varphi)$ by [AV06, Proposition 5.3] (and in fact it is a maximal subgroup if φ is nonperiodic, by [AC13, Lemma 6.3]). Therefore, every proper primitive substitution is stable, according to Lemma 3.12.

For a formation H of finite groups, denote $\hat{\varphi}_{\mathsf{H}}$ the unique continuous endomorphism of $\widehat{F}_{\mathsf{H}}(A)$ such that $\hat{\varphi}_{\mathsf{H}}(a) = p_{\mathsf{H}}(\varphi(a))$ for every $a \in A$. We say that φ is H-*invertible* if $\hat{\varphi}_{\mathsf{H}}$ is an automorphism, a condition that is equivalent to have $(\hat{\varphi}_{\mathsf{H}})^\omega$ equal to the identity on $\widehat{F}_{\mathsf{H}}(A)$ (cf. [ACKP20b, Proposition 3.7.4]).

Proposition 3.13. *Let H be a formation of finite groups. Suppose that φ is an H-invertible stable primitive substitution. Then F_φ is H-charging.*

Proof: First notice that, for every $u \in \widehat{A^+}$, the projection $p_{\mathsf{H}}(u^\omega)$ is idempotent and thus is the neutral element of the group $\widehat{F}_{\mathsf{H}}(A)$. Hence, we have $p_{\mathsf{H}}(\widehat{A^+}) = \widehat{F}_{\mathsf{H}}(A)$.

By Lemma 3.12, the set $T = \hat{\varphi}^\omega(\widehat{A^+})$ is a closed subsemigroup of $\widehat{A^*}$ that is contained in the regular $\mathcal{J}_A$-class $J(F_\varphi)$. Therefore, T is a completely simple profinite semigroup (i.e., $\mathcal{J}_T$ is the universal relation on T).

Since the equality $p_{\mathsf{H}} \circ \hat{\varphi} = \hat{\varphi}_{\mathsf{H}} \circ p_{\mathsf{H}}$ holds, so does $p_{\mathsf{H}} \circ \hat{\varphi}^k = \hat{\varphi}_{\mathsf{H}}^k \circ p_{\mathsf{H}}$ for every $k \in \mathbb{N}$, whence $p_{\mathsf{H}} \circ \hat{\varphi}^\omega = \hat{\varphi}_{\mathsf{H}}^\omega \circ p_{\mathsf{H}}$. It follows that

$$p_{\mathsf{H}}(T) = p_{\mathsf{H}}(\hat{\varphi}^\omega(\widehat{A^+})) = (\hat{\varphi}_{\mathsf{H}})^\omega(p_{\mathsf{H}}(\widehat{A^+})) = (\hat{\varphi}_{\mathsf{H}})^\omega(\widehat{F}_{\mathsf{H}}(A)) = \widehat{F}_{\mathsf{H}}(A),$$

where the last equality holds because φ is H-invertible. Since $p_{\mathsf{H}}(T) = \widehat{F}_{\mathsf{H}}(A)$ and T is completely simple, there is a closed subgroup N contained in T such that $p_{\mathsf{H}}(N) = \widehat{F}_{\mathsf{H}}(A)$ (cf. [RS09, Lemma 4.6.10]). Therefore, if K is the maximal subgroup of $J(F_\varphi)$ containing N, then $p_{\mathsf{H}}(K) = \widehat{F}_{\mathsf{H}}(A)$. ∎

The subgroup of $\widehat{F}_{\mathsf{G}}(A)$ generated by A is the A-generated free group $FG(A)$. It is well known that a substitution $\varphi\colon A^* \to A^*$ is G-invertible if and only if the endomorphism of $FG(A)$ extending φ has an inverse, if and only if $FG(A)$ is generated by $\varphi(A)$ [ACKP20b, Proposition 4.6.8]. We use this in the next example.

Example 3.14. Consider the alphabet $A = \{0, 1, 2\}$, and let φ be the proper (and thus stable) primitive substitution φ over A given by

$$\varphi(0) = 012, \quad \varphi(1) = 0122, \quad \varphi(2) = 0121012.$$

Then φ is G-invertible. Therefore, by Propositions 3.9 and 3.13, every group code over the alphabet A is F_φ-charged.

The language F_φ is not connected: for example, the extension graph of the letter 1 is disconnected. Indeed, $F_\varphi \cap A1A = \{012, 210\}$, since for every $a \in A$ the letter 1 is a factor of $\varphi(a)$ which is neither a prefix nor a suffix of $\varphi(a)$, and so the elements of $F_\varphi \cap A1A$ must be factors of some of the words $\varphi(0)$, $\varphi(1)$, $\varphi(2)$.

The matrix M_φ associated to a substitution $\varphi\colon A^* \to A^*$ is the matrix $A \times A$ where each entry (a, b) is the number of occurrences of a in $\varphi(b)$. For a set π of prime numbers, the formation of finite nilpotent π-groups is denoted $\mathsf{G}_{\mathsf{nil},\pi}$. For the next example we take advantage of the fact that φ is $\mathsf{G}_{\mathsf{nil},\pi}$-invertible if and only if $\det M_\varphi \not\equiv 0 \pmod p$ for every $p \in \pi$ (see the proof of [Alm02, Corollary 5.3]).

Example 3.15. Let φ be the proper primitive substitution over $A = \{0, 1\}$ such that $\varphi(0) = 01$ and $\varphi(1) = 0001$. As $\det M_\varphi = -2$, if π is the set of odd primes, then φ is $\mathsf{G}_{\mathsf{nil},\pi}$-invertible. Hence, by Propositions 3.9 and 3.13, every group code $Z \subseteq A^+$ such that the finite group $M(Z^*)$ is nilpotent of odd order is F_φ-charged.

In contrast, the group code $Z = A^2$ is not F_φ-charged: indeed, $M(Z^*) = \mathbb{Z}/2\mathbb{Z}$, and there is a maximal subgroup K contained in the image of $\hat\varphi$ by [AC13, Lemma 6.3], thus $\hat\eta_{Z^*}(\hat\varphi(\widehat{A^*})) = 0 = \hat\eta_{Z^*}(K)$ as $\varphi(0), \varphi(1)$ have even length.

Several other examples of F-charged complete bifix codes Z are given in [ACKP20a], in which Z may not be a group code, or F is recurrent but not uniformly recurrent.

4. Decoding of languages

From hereon, X is a finite code contained in A^+. A *coding morphism* for X is an injective homomorphism $\beta_X\colon B^* \to A^*$ such that $\beta_X(B) = X$, for some alphabet B.

Example 4.1. Take $A = \{a, b\}$, $X = A^2$, and consider the coding morphism $\beta_X : B^* \to A^*$ given by $\beta_X(z) = ab$ and $\beta_X(t) = ba$, where $B = \{z, t\}$. Let F be the set of factors of $(ab)^*$. Then $\beta_X^{-1}(F) = z^* \cup t^*$.

The study of coding morphisms reduces to the case where $B = X$ and β_X is the inclusion $X^* \to A^*$, since such assumption reflects a mere relabeling of letters. If β_X is indeed the inclusion $X^* \to A^*$, then $\beta_X^{-1}(F)$ is the intersection $F \cap X^*$. The set $F \cap X^*$ is the *decoding* of F by X. If X is an F-complete bifix code, then $F \cap X^*$ is said to be a *complete bifix decoding*.

Take a factorial language $F \subseteq A^*$ and a code $Z \subseteq A^+$. Set $X = F \cap Z$. Then the equality

$$F \cap Z^* = F \cap X^*$$

holds. Therefore, by Theorem 3.6 and Remark 3.7, if F is uniformly recurrent, then a set is a complete bifix decoding of F by a finite F-complete bifix code if and only if it is of the form $F \cap Z^*$ for some rational complete bifix code Z. In symbolic dynamics, the following example of a complete bifix decoding process is of great importance (cf. [LM95, Section 1.4]).

Example 4.2. Let $\mathcal{S}$ be a subshift of $A^{\mathbb{Z}}$. Consider a positive integer n. Then $Z = A^n$ is a complete bifix code (actually, a group code). Let $F = \mathcal{B}(\mathcal{S})$. The *$n$-th higher power* of $\mathcal{S}$ is the subshift $\mathcal{S}^n$ of $X^{\mathbb{Z}}$ defined by the equality $\mathcal{B}(\mathcal{S}^n) = F \cap Z^*$.

Margolis, Sapir and Weil obtained the following key result [MSW98].

Theorem 4.3. *Let X be a finite code contained in A^+. The unique extension of a coding morphism $\beta_X : B^* \to A^*$ to a continuous homomorphism $\bar{\beta}_X : \widehat{B^*} \to \widehat{A^*}$ is an injective mapping.*

Under the assumption that β_X is the inclusion, Theorem 4.3 says that, for every finite code $X \subseteq A^+$, the free profinite monoid over X is identified with the closed submonoid of $\widehat{A^*}$ generated by X^*, that is, the equality

$$\widehat{X^*} = \overline{X^*}$$

holds whenever X is a finite code. The reader should bear in mind this equality along the paper. Notice also that $\widehat{X^*}$ is a clopen subset of $\widehat{A^*}$, by Theorem 2.1. Therefore, the equality

$$\overline{F \cap X^*} = \overline{F} \cap \widehat{X^*}$$

holds, for every subset F of A^*. (The topological closure in $\widehat{X^*}$ of a subset L of $\widehat{X^*}$ coincides with the topological closure of L in $\widehat{A^*}$, and for that reason the notation $\overline{L}$ is not ambiguous.)

Clearly, for every finite code $X \subseteq A^+$, if F is a factorial language over the alphabet A, then $F \cap X^*$ is a factorial langage over the alphabet X; and if F is prolongable over A and X is F-complete, then $F \cap X^*$ is prolongable over X. Therefore, if F is a subshift language over A and X is F-complete, then $F \cap X^*$ is a subshift language over X.

Example 4.4. Here is an example where $F \cap X^*$ is not be prolongable over X: for the Fibonacci language F and the right F-complete prefix code $X = \{a, ba\}$, we have $a^2 \in F \cap X^*$, but there is no $x \in X$ such that $xa^2 \in F$.

A complete bifix decoding of a uniformly recurrent language may not be recurrent, as seen in Example 4.1.

If X is a finite F-complete F-charged code, then we say that $F \cap X^*$ is a *charged complete bifix decoding.* In contrast with Example 4.1, we have the following theorem, one of the main results of this paper.

Theorem 4.5. *Every charged complete bifix decoding of a uniformly recurrent language is uniformly recurrent.*

We defer to Section 8 the proof of Theorem 4.5. We also defer (to Section 9) the proof of the following analog of Theorem 4.5.

Theorem 4.6. *Every charged complete bifix decoding of a recurrent language is recurrent.*

We let ℓ_n^A denote the unique continuous homomorphism $\widehat{A^*} \to \mathbb{Z}/n\mathbb{Z}$ such that $\ell_n^A(a)$ is the class of 1 modulo n for every letter $a \in A$. For an irreducible subshift $\mathcal{S} \subseteq A^{\mathbb{Z}}$, we let $J(\mathcal{S})$ denote the $\mathcal{J}_A$-class $\mathcal{J}_A(\mathcal{B}(\mathcal{S}))$.

Corollary 4.7. *Let $\mathcal{S}$ be an irreducible subshift of $A^{\mathbb{Z}}$. Let n be a positive integer, and let $X = \mathcal{B}(S) \cap A^n$. Suppose that for some maximal subgroup H of $J(\mathcal{S})$ we have $\ell_n^A(H) = \mathbb{Z}/n\mathbb{Z}$. Then the subshift $\mathcal{S}^n \subseteq X^{\mathbb{Z}}$ is irreducible. If moreover $\mathcal{S}$ is minimal, then $\mathcal{S}^n$ is a minimal subshift of $X^{\mathbb{Z}}$.*

Proof: Consider the group code $Z = A^n$. The syntactic homomorphism of Z^* is precisely the restriction $\ell_n^A|_{A^*}\colon A^* \to \mathbb{Z}/n\mathbb{Z}$. Since $\ell_n^A(H) = \mathbb{Z}/n\mathbb{Z}$ for a maximal subgroup $H \subseteq J(\mathcal{S})$, the complete bifix code Z is $\mathcal{B}(\mathcal{S})$-charged. Hence $\mathcal{B}(\mathcal{S}) \cap X^*$ is a recurrent language over the alphabet X, by Theorem 4.6, that is, the subshift $\mathcal{S}^n \subseteq X^{\mathbb{Z}}$ is irreducible. If moreover $\mathcal{S}$ is minimal, then so is $\mathcal{S}^n$ by Theorem 4.5. ■

5. Finite factors of idempotents of $J_A(F) \cap \widehat{X^*}$

For each pseudoword $v \in \widehat{A^*}$, let $\mathrm{Fin}_A(v)$ be the subset of A^* consisting of the finite factors in $\widehat{A^*}$ of v. Clearly, if v is a factor of w, then $\mathrm{Fin}_A(v) \subseteq \mathrm{Fin}_A(w)$, thus $\mathrm{Fin}_A(u) = \mathrm{Fin}_A(v)$ whenever $u \mathrel{\mathcal{J}_A} v$.

Lemma 5.1. *If e is an infinite idempotent of $\widehat{A^*}$, then $\mathrm{Fin}_A(e)$ is a subshift language of A^*.*

Proof: It is trivial that $\mathrm{Fin}_A(e)$ is a nonempty factorial language. Let $u \in \mathrm{Fin}_A(e)$. Then we have $e = xuy$ for some $x, y \in \widehat{A^*}$. Since e is idempotent, we may suppose that $x = ex$ and $y = ye$. Therefore, x has a last letter a, and y has a first letter b, and $aub \in \mathrm{Fin}_A(e)$, showing that $\mathrm{Fin}_A(e)$ is prolongable. ■

Denote $\mathrm{Pref}_A(v)$ the set of prefixes of the pseudoword $v \in \widehat{A^*}$ belonging to A^*.

Lemma 5.2 (cf. [ACKP20b, Exercise 5.28])**.** *Let F be a recurrent language over the alphabet A. If $v \in F$, then there is an idempotent $e \in J_A(F)$ such that $v \in \mathrm{Pref}_A(e)$.*

For a subshift language $F \subseteq A^*$, we denote by $E_A(F)$ the set of idempotents in $J_A(F)$, and by $E_X(F)$ the set of idempotents in $J_A(F) \cap \widehat{X^*}$.

Proposition 5.3. *Let $F \subseteq A^*$ be a recurrent language, and let X be a finite code. Then we have the inclusions*

$$F \cap X^* \supseteq \bigcup_{e \in E_X(F)} \mathrm{Fin}_X(e) \supseteq \bigcup_{e \in E_X(F)} \mathrm{Pref}_X(e). \tag{5.1}$$

Moreover, if X is a finite F-complete bifix code, then $F \cap X^*$ is a subshift language, and the equalities

$$E_X(F) = E_A(F)$$

and

$$F \cap X^* = \bigcup_{e \in E_A(F)} \mathrm{Fin}_X(e) = \bigcup_{e \in E_A(F)} \mathrm{Pref}_X(e)$$

hold.

Proof: For $e \in E_X(F)$, it is trivial that $\mathrm{Pref}_X(e) \subseteq \mathrm{Fin}_X(e)$. Since $\overline{F}$ is a factorial subset of $\widehat{A^*}$ (cf. Proposition 2.6), we have $\mathrm{Fin}_X(e) \subseteq F \cap X^*$, yielding (5.1).

Suppose that X is a finite F-complete bifix code. By Proposition 3.8, every idempotent in $\overline{F}$ belongs to $\overline{X^*}$, and so $E_A(F) = E_X(F)$ as $J_A(F) \subseteq \overline{F}$. By (5.1), it remains to establish $F \cap X^* \subseteq \bigcup_{e\in E_A(F)} \mathrm{Pref}_X(e)$. Let $u \in F \cap X^*$. Then there is an idempotent $e \in E_A(F)$ with $e = uw$ for some $w \in \widehat{A^*}$, by Lemma 5.2. Since $E_A(F) = E_X(F) \subseteq \overline{X^*}$, it then follows from implication (3.2) in Proposition 3.1 that $w \in \overline{X^*}$. Hence the factorization $e = uw$ yields $u \in \mathrm{Pref}_X(e)$. ■

Proposition 5.3 motivates the consideration of the subset F_X of $F \cap X^*$ given by

$$F_X = \bigcup_{e\in E_X(F)} \mathrm{Fin}_X(e),$$

where $X \subseteq A^+$ is a finite code and $F \subseteq A^*$ is a recurrent language. In what follows, $||X||$ is the length of the element of X with greatest length.

Proposition 5.4. *Suppose that F is a recurrent language over the alphabet A and that X is a finite code contained in F. Then, the union*

$$F_X = \bigcup_{e\in E_X(F)} \mathrm{Fin}_X(e)$$

is a finitary union: the set $\{\mathrm{Fin}_X(e) \mid e \in E_X(F)\}$ has at most $||X||$ elements.

For the proof of this proposition, we use the following fact.

Lemma 5.5. *Let X be a finite subset of A^*. Let $\rho = uv$ be a factorization of a pseudoword ρ in $\overline{X^*}$. Then we have $u = \mu x$, $v = y\nu$ for some pseudowords $\mu, \nu \in \overline{X^*}$, and some $x, y \in A^*$ such that $xy \in X$.*

Proof: Let $(u_n)_n$ and $(v_n)_n$ be sequences of elements of A^* such that $u_n \to u$ and $v_n \to v$. Since X^* is rational, the set $\overline{X^*}$ is an open neighborhood of $\rho = uv$. Hence, by taking subsequences, we may as well suppose that $u_n v_n \in X^*$ for all n. We then have factorizations $u_n = \mu_n x_n$ and $v_n = y_n \nu_n$ with $\mu_n, \nu_n \in X^*$ and $x_n y_n \in X$, for every n. Let (μ, x, y, ν) be an accumulation point of the sequence (μ_n, x_n, y_n, ν_n). Then $u = \mu x$, $v = y\nu$ and, since X is finite, x, y are words such that $xy \in X$. ■

Proof of Proposition 5.4: Given any infinite idempotent f of $\widehat{A^*}$, and any nonnegative integer n, there is a unique prefix y of length n of f, and for such y, there is a unique suffix f' of f with $f = yf'$, according to Proposition 2.2. In what follows, we denote $\sigma^n(f)$ the pseudoword $f'y$, which is idempotent by Corollary 2.4.

Fix $f \in E_X(F)$. Let e be an arbitrary element of $E_X(F)$. Take $p, q \in \widehat{A^*}$ such that $e = pfq$. Since f is idempotent, we may take p and q such that $p = pf$ and $q = fq$, so that $e = pq$. Because $e \in \overline{X^*}$, applying Lemma 5.5 we obtain $p = p'x$ and $q = yq'$ for some $p', q' \in \overline{X^*}$ and words x, y such that $xy \in X$. We claim that we may suppose that $x \neq \varepsilon$: indeed, since p' is an infinite pseudoword, we have in fact $p' = p''x'$ for some $p'' \in \overline{X^*}$ and $x' \in X$, so that if $x = \varepsilon$ (and thus $y \in X$) we could then replace x by x', p' by p'', y by ε, and q' by q, showing the claim. Throughout the proof, we suppose that $x \neq \varepsilon$.

From $fq = q = yq'$ we obtain a factorization $f = yf'$ for some (unique) infinite pseudoword f', by Proposition 2.2. From $yf'q = fq = yq'$ we get $f'q = q'$ by cancellation of the finite prefix y, again by Proposition 2.2. Let $n = |y|$. Note that

$$\sigma^n(f) = f'y.$$

Since $x \neq \varepsilon$, we have $n < ||X||$. Note also that

$$p'xy \cdot f'y = pfy = py = p'xy \in \overline{X^*} \qquad \text{and} \qquad f'y \cdot q' = f'q = q' \in \overline{X^*}.$$

By the implication (3.1) in Proposition 3.1, with $u = p'xy$, $v = f'y$ and $w = q'$, we conclude that $f'y \in \overline{X^*}$, that is $f'y \in E_X(F)$. Hence, in the factorization

$$e = py \cdot f'y \cdot q' \tag{5.2}$$

all factors $py = p'xy$, $f'y$ and q' belong to $\overline{X^*}$, whence

$$e \leq_{\mathcal{J}_X} f'y. \tag{5.3}$$

By (5.2) and Corollary 2.4, the pseudoword $f'' = (f'y \cdot q' \cdot py)^2$ is idempotent (cf. Remark 2.5). Moreover, we have $e \mathrel{\mathcal{J}_A} f'' \mathrel{\mathcal{J}_A} f'' \cdot f'y$ and $f'' \cdot f'y \in f'y \cdot \widehat{A^*} \cdot f'y$. Since the monoid $\widehat{A^*}$ is stable, we conclude that the pseudoword $f'' \cdot f'y$ belongs to the maximal subgroup of $J_A(F)$ containing the idempotent $f'y$. It follows that $f'y = (f'' \cdot f'y)^\omega = (f'y \cdot q' \cdot py \cdot f'y)^\omega$. Then, the next chain of equalities holds:

$$\begin{aligned} f'y &= f'y \cdot q' \cdot (py \cdot f'y \cdot q') \cdot py \cdot (f'y \cdot q' \cdot py \cdot f'y)^{\omega-2} \\ &= f'y \cdot q' \cdot e \cdot py \cdot (f'y \cdot q' \cdot py \cdot f'y)^{\omega-2}. \end{aligned}$$

In the former expression, all factors between consecutive dots belong to $\overline{X^*}$, and so, combining with (5.3), we conclude that

$$\sigma^n(f) = f'y \ \mathcal{J}_X \ e.$$

This shows that $\mathrm{Fin}_X(e) = \mathrm{Fin}_X(\sigma^n(f))$. Therefore, the equality

$$F_X = \bigcup_{\substack{n \leq ||X|| \\ \sigma^n(f) \in E_X(F)}} \mathrm{Fin}_X(\sigma^n(f)),$$

holds, and the proposition is established. ■

The set F_X is always a factorial and prolongable language over the alphabet X (cf. Lemma 5.1), but F_X may be empty since $E_X(F)$ may be empty.

Example 5.6. Let F be the Fibonacci language and let X be the bifix code $\{aa, bb\}$. Then $F \cap X^*$ is finite, as seen in [BDFD$^+$15a, Example 3.12]. Because $\mathrm{Fin}_X(e)$ is infinite whenever $e \in E_X(F)$ (cf. Lemma 5.1), the inclusion of F_X on the finite set $F \cap X^*$ implies that $E_X(F) = \emptyset$ and $F_X = \emptyset$.

In contrast with Example 5.6, we have the following.

Proposition 5.7. *Let $F \subseteq A^*$ be a recurrent language. If X is a right F-complete prefix code, then F_X is nonempty.*

Proof: Let $(u_n)_n$ be a sequence of elements of F converging in $\widehat{A^*}$ to $u \in J_A(F)$. As X is right F-complete, for each n we have a factorization $u_n = x_n s_n$ with $x_n \in X^*$ and $s_n \in A^*$ such that s_n is a proper prefix of an element of X. Take an accumulation point (x, s) of (x_n, s_n). Then $u = xs$ and, since X is finite, one has $s \in A^*$. It follows from Corollary 2.3 that $x \in J_A(F) \cap \overline{X^*}$. Hence x is regular and $x = xe$ for some idempotent $e \in J_A(F)$. Since X is a prefix code, applying Proposition 3.1 we get $e \in E_X(F)$, and so $E_X(F)$ and F_X are nonempty. ■

Proposition 5.3 guarantees that $F_X = F \cap X^*$ when X is a finite F-complete bifix code. The inclusion $F_X \subseteq F \cap X^*$ may be strict when X is just a right F-complete prefix code, as $F \cap X^*$ may not be prolongable over X (cf. Example 4.4).

6. Relationship between $J_A(F)$ and $J_X(F \cap X^*)$

In this section we seek to clarify the relationship between the $\mathcal{J}_A$-class $J_A(F)$ and the union of $\mathcal{J}_X$-classes $J_X(F \cap X^*)$ when F is recurrent and X is a finite F-complet bifix code.

The following lemma helps dealing with factorizations of pseudowords.

Lemma 6.1. *Let $u, v \in \widehat{A^*}$. If the sequence $(w_n)_n$ converges to uv in $\widehat{A^*}$, then there are factorizations $w_n = u_n v_n$ such that $u_n \to u$ and $v_n \to v$.*

A proof of this lemma appears in [ACKP20b, Exercise 4.20]; the result was discovered independently in [AC09] and [HRS10] (cf. also [ACCZ19, Section 3]). We apply the lemma in the proof of the next proposition, the bulk of this section.

Proposition 6.2. *Let F be a recurrent language over the alphabet A, and suppose that X is a finite F-complete bifix code. Then we have the following properties:*

(i) the equality $J_X(F \cap X^*) = J_A(F) \cap \widehat{X^*}$ holds;
(ii) $J_X(F \cap X^*)$ is a union of finitely many regular $\mathcal{J}_X$-classes of $\widehat{X^*}$;
(iii) for each $\mathcal{K} \in \{\mathcal{R}, \mathcal{L}, \mathcal{H}\}$, the $\mathcal{K}_X$-classes contained in $J_X(F \cap X^*)$ are the sets of the form $V \cap \widehat{X^*}$ that are nonempty and such that V is a $\mathcal{K}_A$-class contained in $J_A(F)$;
(iv) the maximal subgroups of $\widehat{X^*}$ contained in $J_X(F \cap X^*)$ are the intersections $H \cap \widehat{X^*}$ in which H is a maximal subgroup of $\widehat{A^*}$ contained in $J_A(F)$;
(v) if $yX^*x \cap F \subseteq X^*$ whenever $xy \in X^*$, then the maximal subgroups of $\widehat{X^*}$ contained in $J_X(F \cap X^*)$ are isomorphic profinite groups.

Remark 6.3. Proposition 6.2 holds if $X = F \cap A^n$; more generally, if Z is a group code with $Z^* = \eta_{Z^*}^{-1}(N)$ for a normal subgroup N of $M(Z^*)$, then Proposition 6.2 applies to $X = F \cap Z$. Indeed, $X = Z \cap F$ is then finite F-complete (cf. Theorem 3.6), and $xy \in Z^* \Leftrightarrow \eta_{Z^*}(y)N = \eta_{Z^*}(x)^{-1}N$, yielding

$$xy \in Z^* \Rightarrow \eta_{Z^*}(yZ^*x)N = \eta_{Z^*}(y)N\eta_{Z^*}(x)N = N \Rightarrow yZ^*x \subseteq \eta_{Z^*}^{-1}(N) = Z^*,$$

and thus $yX^*x \cap F \subseteq X^*$, ensuring that Proposition 6.2(v) holds for such X.

We proceed with the proof of Proposition 6.2. For showing its item (v) we adapt the proof that maximal subgroups in a regular $\mathcal{J}$-class of a compact monoid are all isomorphic compact groups, see for instance [ACKP20b, Proposition 3.6.11].

Proof of Proposition 6.2: Let $u \in J_X(F \cap X^*)$. There is a finite set I of idempotents in $E_X(F) = J_A(F) \cap \widehat{X^*}$ such that $F \cap X^* = \bigcup_{e \in I} \mathrm{Fin}_X(e)$, by Propositions 5.3 and 5.4. Hence, as $J_X(F \cap X^*) \subseteq \overline{F \cap X^*}$, there is $e \in I$ with $u \in \overline{\mathrm{Fin}_X(e)}$. Since $\leq_{\mathcal{J}_X}$ is a closed relation of $\widehat{X^*}$, we have $e \leq_{\mathcal{J}_X} u$. Note that $e \in \overline{F} \cap \widehat{X^*} = \overline{F \cap X^*}$, as $J_A(F) \subseteq \overline{F}$. It follows from the $\mathcal{J}_X$-minimality of u as an element of $\overline{F \cap X^*}$ that $e \mathrel{\mathcal{J}_X} u$. This shows (ii), as well the inclusion $J_X(F \cap X^*) \subseteq J_A(F) \cap \widehat{X^*}$ in (i).

Let $u \in J_A(F) \cap \widehat{X^*}$. Since F is recurrent, the set $J_A(F)$ is a regular $\mathcal{J}_A$-class, and so we may take some idempotent $e \in J_A(F)$ that is $\mathcal{R}_A$-equivalent to u. Let $s \in \widehat{A^*}$ be such that $e = us$. As X is an F-complete bifix code, we know that $e \in \widehat{X^*}$ by Proposition 3.8, which together with the equality $e = us$ implies $s \in \widehat{X^*}$ by Proposition 3.1. Since we also have $u = eu$, we deduce that $u \mathrel{\mathcal{R}_X} e$. We claim that $e \in J_X(F \cap X^*)$. Since $e \in J_A(F) \cap \widehat{X^*} \subseteq \overline{F \cap X^*}$, and we already showed that $J_X(F \cap X^*)$ is a union of regular $\mathcal{J}_X$-classes, we know that there is some idempotent $f \in J_X(F \cap X^*)$ and some $x, y \in \widehat{X^*}$ such that $f = xey$ (cf. Remark 2.7). Then the idempotent $e' = (eyx)^2$ is $\mathcal{J}_X$-equivalent to f and satisfies $e' \leq_{\mathcal{R}_X} e$, whence $e' \leq_{\mathcal{R}_A} e$. Since $e, e' \in J_A(F)$ as the inclusion $J_X(F \cap X^*) \subseteq J_A(F)$ was already established, by stability of $\widehat{A^*}$ we actually have $e \mathrel{\mathcal{R}_A} e'$, which means that $e = e'e$ and $e' = ee'$, thus $e \mathrel{\mathcal{R}_X} e'$. This shows the claim that $e \in J_X(F \cap X^*)$. Hence, the above conclusion that $u \mathrel{\mathcal{R}_X} e$ yields $J_A(F) \cap \widehat{X^*} \subseteq J_X(F \cap X^*)$, finishing the proof of (i).

For showing (iii), it suffices to consider $\mathcal{K} = \mathcal{R}$, as case $\mathcal{K} = \mathcal{L}$ is symmetric and case $\mathcal{K} = \mathcal{H}$ follows straightforwardly from the other two cases. Since $\mathcal{R}_X \subseteq \mathcal{R}_A$, it is clear that every $\mathcal{R}_X$-class contained in $J_X(F \cap X^*)$ is contained in an intersection $V \cap \widehat{X^*}$ for some $\mathcal{R}_A$-class V, with $V \subseteq J_A(F)$, as we saw that $J_X(F \cap X^*) \subseteq J_A(F)$. Conversely, let V be any $\mathcal{R}_A$-class contained in $J_A(F)$ such that $V \cap \widehat{X^*} \neq \emptyset$. Note that $V \cap \widehat{X^*} \subseteq J_X(F \cap X^*)$, since we showed $J_A(F) \cap \widehat{X^*} \subseteq J_X(F \cap X^*)$. Let $v, w \in V \cap \widehat{X^*}$, and take $s, t \in \widehat{A^*}$ such that $v = ws$ and $w = vt$. Because X is bifix, we actually have $s, t \in \widehat{X^*}$, by

Proposition 3.1, whence $v \mathcal{R}_X w$. Therefore, $V \cap \widehat{X^*}$ is an $\mathcal{R}_X$-class contained in $J_X(F \cap X^*)$. This concludes the proof of (iii).

Since every idempotent in $J_A(F)$ belongs to $\widehat{X^*}$ by Proposition 3.8, Property (iv) follows immediately from Property (iii).

Finally, suppose that $xy \Rightarrow yX^*x \cap F \subseteq X^*$. We claim that

$$xy \in \overline{X^*} \Rightarrow y\widehat{X^*}x \cap \overline{F} \subseteq \widehat{X^*}. \tag{6.1}$$

Assuming that $xy \in \overline{X^*}$, let $u \in y\widehat{X^*}x \cap \overline{F}$. Take $v \in \widehat{X^*}$ such that $u = yvx$ and a sequence $(u_n)_n$ of elements of F converging to u. By Lemma 6.1, for every n there are factorizations $u_n = y_n v_n x_n$ such that $\lim y_n = y$, $\lim v_n = v$ and $\lim x_n = x$. Since $\widehat{X^*}$ is a clopen subset of $\widehat{A^*}$, we may as well suppose that the words v_n and $x_n y_n$ belong to X^* for every n. By hypothesis, $x_n y_n \in X^*$ implies $y_n X^* x_n \cap F \subseteq X^*$, whence $y_n v_n x_n \in X^*$ for every n. This shows that $u = yvx \in \widehat{X^*}$, establishing the claim.

Let e, f be idempotents in $J_X(F \cap X^*)$, and let H_e, H_f be the maximal subgroups of $\widehat{A^*}$ respectively containing e, f. We already saw that $J_X(F \cap X^*) \subseteq J_A(F) \cap \widehat{X^*}$. Since F is recurrent, $J_A(F)$ is a regular $\mathcal{J}_A$-class. Therefore, there are $x, y \in J_A(F)$ such that $e = xy$ and $f = yx$. Moreover, the mappings $\varphi\colon H_e \to H_f$ and $\psi\colon H_f \to H_e$ such that $\varphi(g) = ygx$ and $\psi(h) = xhy$, for every $g \in H_e$ and $h \in H_f$, are mutually inverse continuous isomorphisms. In view of (iv) we only need to check that $\varphi(H_e \cap \widehat{X^*}) = H_f \cap \widehat{X^*}$. Since $xy = e \in \widehat{X^*}$ and $J_A(F) \subseteq \overline{F}$, it follows from (6.1) that $\varphi(H_e \cap \widehat{X^*}) \subseteq H_f \cap \widehat{X^*}$. Similarly, the inclusion $\psi(H_f \cap \widehat{X^*}) \subseteq H_e \cap \widehat{X^*}$ also holds. Hence, $\varphi(H_e \cap \widehat{X^*}) = H_f \cap \widehat{X^*}$ indeed holds, concluding the proof of (v). ∎

7. Decoding of uniformly recurrent languages by finite codes

For the sake of the proof of Theorem 4.5, we establish in this section some results concerning the decoding of uniformly recurrent languages by arbitrary finite codes, not necessarily bifix.

Proposition 7.1. *Let F be a uniformly recurrent language over the alphabet A. Let X be any finite code contained in F. For every $e \in E_X(F)$, the set* $\mathrm{Fin}_X(e)$ *is a uniformly recurrent language over the alphabet X.*

Proof: Let $e \in E_X(F)$. Suppose that $u \in \widehat{X^*}$ is an infinite pseudoword such that $e \leq_{\mathcal{J}_X} u$. Take elements x and y of $\widehat{X^*}$ such that $e = xuy$. Then the pseudoword $f = (u \cdot yx)^2 \in \widehat{X^*}$ is an idempotent such that $e \mathrel{\mathcal{J}_X} f$. Since u is an infinite factor of e in $\widehat{A^*}$, and F is uniformly recurrent over A, we obtain from Theorem 2.8 that $u \mathrel{\mathcal{J}_A} e$. As $\widehat{A^*}$ is stable, it follows that $u \mathrel{\mathcal{R}_A} f$, whence

$$u = f \cdot u = uyxuyx \cdot u = \underbrace{uy}_{\in \overline{X^*}} \cdot e \cdot \underbrace{xu}_{\in \overline{X^*}}.$$

Therefore $e \mathrel{\mathcal{J}_X} u$, showing that e is a $\mathcal{J}_X$-maximal infinite pseudoword of $\overline{X^*}$. This shows that there is a minimal subshift $\mathcal{S} \subseteq X^{\mathbb{Z}}$ such that $e \in J_X(\mathcal{B}(\mathcal{S}))$, by Theorem 2.8. As $\overline{\mathcal{B}(\mathcal{S})}$ is a factorial subset of $\widehat{X^*}$, all factors of $J_X(\mathcal{B}(\mathcal{S}))$ in X^* belong to $\mathcal{B}(\mathcal{S})$, whence $\mathrm{Fin}_X(e) \subseteq \mathcal{B}(\mathcal{S})$. Since $\mathrm{Fin}_X(e)$ is a subshift language (cf. Lemma 5.1), we deduce from the minimality of $\mathcal{S}$ that $\mathrm{Fin}_X(e) = \mathcal{B}(\mathcal{S})$, thus showing that $\mathrm{Fin}_X(e)$ is uniformly recurrent. ∎

Corollary 7.2. *Let F be a uniformly recurrent language over the alphabet A. Let X be any finite code contained in F. If F_X is recurrent over the alphabet X, then it is uniformly recurrent over X.*

Proof: As seen in Proposition 5.4, the set F_X is the union of the finite collection $\{\mathrm{Fin}_X(e) \mid e \in E_X(F)\}$. By Proposition 7.1, the elements of this collection are uniformly recurrent languages over X.

It remains to observe that if a union of a finite collection of uniformly recurrent sets is recurrent, then it is actually uniformly recurrent. We give a "profinite" proof for this known fact. Suppose that $F = \bigcup_{i=1}^{k} F_i$ is a finite union of k distinct sets $F_i \subseteq X^*$ that are uniformly recurrently over X, and that F is recurrent over X. For each i, let $u_i \in J(F_i)$. As $\overline{F}$ is recurrent, there are $x_1, \ldots, x_{k-1} \in \widehat{X^*}$ such that the pseudoword $w = u_1 x_1 u_2 x_2 \cdots u_{k-1} x_{k-1} u_k$ belongs to $\overline{F}$. Since $\overline{F} = \bigcup_{i=1}^{k} \overline{F_i}$, we then have $w \in \overline{F_j}$ for some j. Because F_j is uniformly recurrent, we know that w is an $\mathcal{J}_X$-maximal infinite pseudoword, and so $u_i \mathrel{\mathcal{J}_X} w$ for every $i \in \{1, \ldots, k\}$. But $J(F_i) = J(F_j)$ implies $F_i = F_j$, and so necessarily $k = 1$. ∎

Corollary 7.3. *If the complete bifix decoding of a uniformly recurrent language is recurrent, then it is uniformly recurrent.*

Proof: This is a special case of Corollary 7.2: for a uniformly recurrent language F and a finite F-complete bifix code X, we have $F \cap X^* = F_X$ by Proposition 5.3. ■

8. Proof of Theorem 4.5

In this section we prove that every charged complete bifix decoding of a uniformly recurrent language is uniformly recurrent (Theorem 4.5). The following fact is needed.

Lemma 8.1. *Let F be a factorial language of A^*. If Z is a rational language of A^*, then the inclusion $\overline{Z^*} \cap \overline{F} \subseteq \overline{(Z \cap F)^*}$ holds in $\widehat{A^*}$.*

Proof: Let $u \in \overline{Z^*} \cap \overline{F}$. Take a sequence $(u_n)_{n\in\mathbb{N}}$ of elements of F converging to u. Since Z^* is rational, $\overline{Z^*}$ is open, and so there is $N \in \mathbb{N}$ such that $n \geq N$ implies $u_n \in Z^* \cap F$. As F is factorial, we have $Z^* \cap F = (Z \cap F)^*$, thus $u \in \overline{(Z \cap F)^*}$. ■

In the next proof, we denote by H_u the $\mathcal{H}$-class of an element u of a monoid M.

Proposition 8.2. *Let F be a recurrent language over the alphabet A. If X is a finite F-complete F-charged bifix code, then $F \cap X^* = \mathrm{Fin}_X(e)$ for every $e \in E_A(F)$.*

Proof: Let $Z \subseteq A^+$ be an F-charged rational complete bifix code for which we have $X = Z \cap F$. Let $e, f \in E_A(F)$. Recall that $e, f \in \overline{X^*}$ by Proposition 3.8. By Proposition 5.3, it suffices to show that $e \mathrel{\mathcal{J}_X} f$.

As $e \mathrel{\mathcal{J}_A} f$, we may take $x, y \in \widehat{A^*}$ such that $e = xfy$, with $x = exf$ and $y = fye$. Since Z is F-charged, the inclusion $\hat{\eta}_{Z^*}(J_A(F))$ is contained in the minimum ideal of the syntactic monoid M_{Z^*}, and so

$$\hat{\eta}_{Z^*}(x)^\omega \mathrel{\mathcal{H}} \hat{\eta}_{Z^*}(x) \mathrel{\mathcal{R}} \hat{\eta}_{Z^*}(e).$$

Therefore, since $\hat{\eta}_{Z^*}(e) \cdot \hat{\eta}_{Z^*}(x) = \hat{\eta}_{Z^*}(x)$, we know from Green's Lemma (cf. [RS09, Lemma A.3.1]) that there is $s \in H_{\hat{\eta}_{Z^*}(e)}$ such that $s \cdot \hat{\eta}_{Z^*}(x) = \hat{\eta}_{Z^*}(x)^\omega$. Because Z is F-charged, we have $\hat{\eta}_{Z^*}(H_e) = H_{\hat{\eta}_{Z^*}(e)}$ and we may take some $u \in H_e$ such that $s = \hat{\eta}_{Z^*}(u)$. Since $u = ue = uxfy \in J_A(F)$, we have $ux \in \overline{F}$. As $\hat{\eta}_{Z^*}(ux) = \hat{\eta}_{Z^*}(x)^\omega$ is idempotent, applying Proposition 3.8 we conclude that $ux \in \overline{Z^*}$.

Symmetrically, there is $v \in H_e$ such that $yv \in \overline{Z^*}$.

Consider the pseudoword $w = ux \cdot f \cdot yv$. Because $xfy = e$ and $u, v \in H_e$, we have $w = uv \in H_e$. In particular, the equality

$$(ux \cdot f \cdot yv)^{\omega-1} \cdot w = e \tag{8.1}$$

holds. Again by Proposition 3.8, the idempotents e, f belong to $\overline{Z^*}$, and since ux and yv also belong to $\overline{Z^*}$, we conclude by the second implication in Proposition 3.1 that $w \in \overline{Z^*}$. And since $\overline{F}$ is factorial, the pseudowords ux, yv, f, e and w belong to $\overline{F}$, and so they belong to $\overline{X^*}$ by Lemma 8.1. Therefore, the equality (8.1) yields $e \leq_{\mathcal{J}_X} f$. By symmetry, we also have $f \leq_{\mathcal{J}_X} e$, concluding the proof. ∎

Proof of Theorem 4.5: Combining Propositions 8.2 and 7.1, we immediately obtain Theorem 4.5. ∎

9. Proof of Theorem 4.6

We proceed to show that every charged complete bifix decoding of a recurrent language is recurrent (Theorem 4.6).

Proof of Theorem 4.6: Let F be a recurrent language over the alphabet A, and let Z be an F-charged rational complete bifix code. Consider the code $X = Z \cap F$, and suppose that X is finite. We want to show that $F \cap X^*$ is a recurrent language over the alphabet X. Let $e \in E_A(F)$. By Proposition 8.2, what we want to show is that $\mathrm{Fin}_X(e)$ is recurrent. Let $u, v \in \mathrm{Fin}_X(e)$. Take $x, z, v, t \in \overline{X^*}$ such that $e = xuy = zvt$. Take a sequence $(w_n)_{n \in \mathbb{N}}$ of elements of F converging to e. Applying Lemma 6.1 to the equality

$$e = x \cdot u \cdot y \cdot z \cdot v \cdot t$$

in the free profinite monoid $\widehat{X^*}$, we conclude that, for every $n \in \mathbb{N}$, there are factorizations

$$w_n = x_n \cdot u_n \cdot y_n \cdot z_n \cdot v_n \cdot t_n$$

in X^* such that, in $\widehat{X^*}$, we have

$$x_n \to x, \quad u_n \to u, \quad y_n \to y, \quad z_n \to z, \quad v_n \to v, \quad t_n \to t.$$

Since $u, v \in X^*$, by taking subsequences we may as well suppose that $u_n = u$ and $v_n = v$ for every $n \in \mathbb{N}$. Because uy_nz_nv is a factor of $w_n \in F$, and F is factorial, we know that uy_nz_nv belongs to $F \cap X^*$. By Lemma 5.2, there is an idempotent $f \in J_A(F)$ of the form

$$f = uy_nz_nvr, \tag{9.1}$$

for some $r \in \widehat{A^*}$. Recall that $f \in \overline{X^*}$ by Proposition 3.8. As $uy_nz_nv \in X^*$ and X is a prefix code, applying Proposition 3.1 we conclude that $r \in \overline{X^*}$. Therefore, from (9.1) we deduce that $uy_nz_nv \in \mathrm{Fin}_X(f)$. Since by Proposition 8.2 the equality $\mathrm{Fin}_X(f) = \mathrm{Fin}_X(e)$ holds, we have thus established that $\mathrm{Fin}_X(e)$ is a recurrent language over the alphabet X. ∎

10. The Schützenberger group of an irreducible subshift is invariant under eventual conjugacy

In the category of symbolic dynamical systems, a morphism between a subshift $\mathcal{S}$ of $A^{\mathbb{Z}}$ and a subshift $\mathcal{T}$ of $B^{\mathbb{Z}}$ is a continuous mapping $\varphi\colon \mathcal{S} \to \mathcal{T}$ such that $\sigma_B(\varphi(x)) = \varphi(\sigma_A(x))$ for every $x \in \mathcal{S}$. If there is a map $\Phi\colon A \to B$ such that $\varphi(x) = (\Phi(x_i))_{i\in\mathbb{Z}}$ for each $x \in \mathcal{S}$, then φ is said to be a *one-block code* with associated *letter-to-letter block map* Φ.

An isomorphism of subshifts is called a *conjugacy*, and isomorphic subshifts are *conjugate*. A *one-block conjugacy* is a one-block code that is a conjugacy. It turns out that conjugacy is the equivalence relation generated by one-block conjugacies. More precisely, the following holds (cf. [LM95, Proposition 1.5.12]).

Proposition 10.1. *If $\varphi\colon \mathcal{S} \to \mathcal{T}$ is a conjugacy, then there is a subshift $\mathcal{R}$ and one-block conjugacies $\alpha\colon \mathcal{R} \to \mathcal{S}$ and $\beta\colon \mathcal{R} \to \mathcal{T}$ such that $\varphi = \beta \circ \alpha^{-1}$.*

Two subshifts $\mathcal{S}$ and $\mathcal{T}$ are *eventually conjugate* when there is a positive integer N such that the higher power shifts $\mathcal{S}^n$ and $\mathcal{T}^n$ are conjugate for every $n \geq N$. If the subshifts $\mathcal{S}$ and $\mathcal{T}$ are conjugate, then $\mathcal{S}$ and $\mathcal{T}$ are eventually conjugate.

10.1. The Schützenberger group of an irreducible subshift and its higher powers.

Let $\mathcal{S}$ be an irreducible subshift of $A^{\mathbb{Z}}$. Since the $\mathcal{J}$-class $J_A(\mathcal{B}(S))$ is regular, we may consider its maximal subgroups, and identify them as a single profinite group $G(\mathcal{S})$, which since [AC13] is called the *Schützenberger group* of $\mathcal{S}$. In this section, we show that $G(\mathcal{S})$ is an invariant of eventual conjugacy, for any irreducible subshift $\mathcal{S}$. The invariance of $G(\mathcal{S})$ under conjugacy was first proved in [Cos06].

We next generalize the definition of Schützenberger group of an irreducible subshift to (possibly not irreducible) higher powers of irreducible subshifts, as follows. Given an irreducible subshift $\mathcal{S}$ of $A^{\mathbb{Z}}$, let $F = \mathcal{B}(\mathcal{S})$. For every positive integer n and irreducible subshift $\mathcal{S} \subseteq A^{\mathbb{Z}}$, the equality $\mathcal{B}(S^n) = F \cap (A^n)^*$ holds. Moreover, by Proposition 6.2:

- the set $J(\mathcal{S}^n) = J_{A^n}(F \cap (A^n)^*)$ of $\leq_{\mathcal{J}_{A^n}}$-minimal elements of $\overline{F \cap (A^n)^*}$ is a union of regular $\mathcal{J}_{A^n}$-classes;
- any two maximal subgroups contained in $J(\mathcal{S}^n)$ are isomorphic profinite groups.

We identify the maximal subgroups contained in $J(\mathcal{S}^n)$ as a single profinite group, which we call the *Schützenberger group of* $\mathcal{S}^n$ and denote $G(\mathcal{S}^n)$. Note that if $\mathcal{S}^n$ is irreducible (which happens if $n = 1$), then we get the Schützenberger group of $\mathcal{S}^n$ as initially defined, and so we indeed have a consistent generalization.

To show the invariance under eventual conjugacy of $G(\mathcal{S})$ when $\mathcal{S}$ is irreducible, we need to use not only the invariance under conjugacy of $G(\mathcal{S})$, but the following stronger result.

Theorem 10.2. *If $\mathcal{S}$ and $\mathcal{T}$ are conjugate irreducible subshifts, then $G(\mathcal{S}^n)$ and $G(\mathcal{T}^n)$ are isomorphic profinite groups, for every $n \geq 1$.*

Theorem 10.2 is also from [Cos06], where one finds a more general result concerning the restriction of the quasi-order $\leq_{\mathcal{J}}$ to $\overline{\mathcal{B}(\mathcal{S})}$: if $\mathcal{S}$ is a (possibly not irreducible) subshift of $A^{\mathbb{Z}}$, then we have invariance under conjugacy of the set of isomorphism classes of the maximal subgroups containing idempotents that are $\leq_{\mathcal{J}_A}$-minimal among the idempotents in $\overline{\mathcal{B}(\mathcal{S})}$ (cf. [Cos06, Theorem 3.11]).

The free procyclic group on one generator $\widehat{\mathbb{Z}}$ is the inverse limit $\varprojlim \mathbb{Z}/\mathbb{Z}_n$, in which $\mathbb{Z}$ embeds densely (cf. [RZ10]). For each alphabet A and pseudoword u in $\widehat{A^*}$, the *procyclic image* of u (with respect to A) is the image of u in $\widehat{\mathbb{Z}}$ by the unique continuous homomorphism $\ell^A \colon \widehat{A^*} \to \widehat{\mathbb{Z}}$ such that $\ell^A(a) = 1$ for every $a \in A$. Note that if $u \in A^*$, then $\ell^A(u)$ is just the length of u. Therefore, the notion of procyclic image of a pseudoword is a generalization of the notion of word length.

The proof in [Cos06] that $G(\mathcal{S})$ is a conjugacy invariant of irreducible subshifts gives for free that the procyclic image of the elements of $G(\mathcal{S})$ is preserved by conjugacy, as seen next.

Theorem 10.3. *Suppose that $\mathcal{S} \subseteq A^{\mathbb{Z}}$ and $\mathcal{T} \subseteq B^{\mathbb{Z}}$ are conjugate irreducible subshifts. If H is a maximal subgroup of $J(\mathcal{S})$ and K is a maximal subgroup of $J(\mathcal{T})$, then there is an isomorphism of profinite groups $\psi \colon H \to K$ that* *preserves procyclic images, that is to say, such that*

$$\ell^B(\psi(h)) = \ell^A(h)$$

for every $h \in H$.

Proof: It suffices to assume that there is a one-block conjugacy $\varphi\colon \mathcal{S} \to \mathcal{T}$, by Proposition 10.1. Let $\Phi\colon A \to B$ be the associated letter-to-letter block map. For every maximal subgroup H contained in $J(\mathcal{S})$, the unique continuous homomorphism $\widehat{\Phi}\colon \widehat{A^*} \to \widehat{B^*}$ extending Φ restricts to a continuous isomorphism from $H \subseteq J(\mathcal{S})$ onto a maximal subgroup $N \subseteq J(\mathcal{T})$, a property of $\widehat{\Phi}$ shown in [Cos06]. (More precisely, applying Proposition 3.10 from [Cos06] we see that $N = \widehat{\Phi}(H)$ is a maximal subgroup, and Theorem 3.11 from [Cos06] guarantees that $N \subseteq J(\mathcal{T})$.) Note that $\ell^B(\widehat{\Phi}(u)) = \ell^A(u)$ for every $u \in \widehat{A^*}$, since that clearly holds when $u \in A^*$.

Finally, let K be any maximal subgroup of $\widehat{B^*}$ contained in $J(\mathcal{T})$. Then there are $x, y \in J(\mathcal{T})$ such that xy is the idempotent in N, yx is the idempotent in K, and the mapping $\lambda\colon N \to K$ given by $\lambda(g) = ygx$ is a continuous isomorphism. As xy is idempotent, we have $\ell^B(x) = -\ell^B(y)$. The fact that $\widehat{\mathbb{Z}}$ is Abelian furthermore yields $\ell^B(\lambda(g)) = \ell^B(g)$ for every $g \in N$. Therefore, the homomorphism $\psi\colon H \to K$ given by $\psi = \lambda \circ (\widehat{\Phi}|_H)$ to H is a continuous isomorphism such that $\ell^B(\psi(h)) = \ell^A(h)$ for every $h \in H$. ■

10.2. Proof of the invariance of $G(\mathcal{S})$ under eventual conjugacy. For each positive integer n, denote ρ_n the unique continuous homomorphism $\widehat{\mathbb{Z}} \to \mathbb{Z}/n\mathbb{Z}$ such that $\rho_n(1)$ is the class modulo n of 1. The composition $\rho_n \circ \ell^A$ is the homomorphism $\ell_n^A\colon \widehat{A^*} \to \mathbb{Z}/n\mathbb{Z}$ introduced before Corollary 4.7.

Remark 10.4. The length of a word u over the alphabet A is a multiple of n if and only if $\ell_n^A(u) = 0$. Therefore, by continuity of the homomorphism $\ell_n^A\colon \widehat{A^*} \to \mathbb{Z}/n\mathbb{Z}$, the equality

$$(\ell_n^A)^{-1}(0) = \overline{(A^n)^*}$$

holds.

Lemma 10.5. *If n and m are relatively prime positive integers, then for every $u \in \widehat{A^*}$ such that $u \in \overline{(A^m)^*}$, the equivalence*

$$\ell_n^A(u) = 0 \Leftrightarrow \ell_n^{A^m}(u) = 0$$

holds.

Proof: Because $(A^n)^*$ and $(A^m)^*$ are rational, we have

$$\overline{(A^n)^*} \cap \overline{(A^m)^*} = \overline{(A^n)^* \cap (A^m)^*}.$$

Since n and m are relatively prime, we also have

$$(A^n)^* \cap (A^m)^* = (A^{nm})^*.$$

Notice also that $(A^{nm})^* = ((A^m)^n)^*$. Therefore, for every $u \in \overline{(A^m)^*}$, the chain of equivalences

$$\ell_n^A(u) = 0 \Leftrightarrow u \in \overline{(A^n)^*} \cap \overline{(A^m)^*} \Leftrightarrow u \in \overline{((A^m)^n)^*} \Leftrightarrow \ell_n^{A^m}(u) = 0$$

holds. ∎

Lemma 10.6. *Let $\mathcal{S}$ be an irreducible subshift of $A^{\mathbb{Z}}$. Let n be a positive integer. Suppose that the maximal subgroup H of $J(\mathcal{S})$ satisfies $\ell_n^A(H) = 0$. Then H is a maximal subgroup of $J(\mathcal{S}^n)$.*

Proof: Let $X = \mathcal{B}(\mathcal{S}) \cap A^n$. By Proposition 6.2, the intersection $H \cap \widehat{X^*}$ is a maximal subgroup of $J(\mathcal{S}^n) \subseteq \overline{\mathcal{B}(\mathcal{S})}$. On the other hand, the equality $\ell_n^A(H) = 0$ gives $H \subseteq \overline{(A^n)^*} \cap \overline{\mathcal{B}(\mathcal{S})}$. Since $\overline{(A^n)^*} \cap \overline{\mathcal{B}(\mathcal{S})} \subseteq \overline{(A^n)^* \cap \mathcal{B}(\mathcal{S})} = \widehat{X^*}$, we are done. ∎

For each positive integer k and closed subgroup H of $\widehat{A^*}$, we denote by $H(k)$ the intersection $H \cap \overline{(A^k)^*}$. In particular, we have $H(1) = H$.

Remark 10.7. Note that we always have $H(nm) \subseteq H(n) \cap H(m)$. Moreover, the equality $H(nm) = H(n) \cap H(m)$ holds whenever m, n are relatively prime.

Note that $H(k) = H \cap (\ell_k^A)^{-1}(0)$ is a closed normal subgroup of H. Hence, we may consider for all $n, m \geq 1$ the profinite semidirect product $H(n) \ltimes H(m)$ induced by the left action in which ${}^x u = xux^{-1}$ for all $x \in H(m)$ and $u \in H(n)$.

Remark 10.8. For every closed normal subgroup H of $\widehat{A^*}$, the semidirect product $H(m) \ltimes H(n)$ is a closed normal subgroup of $H \ltimes H$: it is the kernel of the continuous homomorphism $H \ltimes H \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ mapping each $(u, x) \in H \ltimes H$ to $(\ell_m^A(u), \ell_n^A(x))$.

The following lemma is crucial for our proof that the Schützenberger group of an irreducible subshift is invariant under eventual conjugacy. The notation $\cong$ stands for isomorphism of compact groups.

Lemma 10.9. *Let $\mathcal{S}$ be an irreducible subshift of $A^{\mathbb{Z}}$, and let H be a maximal subgroup contained in $J(\mathcal{S})$. Let m, n be relatively prime positive integers such that $\ell_m(H) = \mathbb{Z}/m\mathbb{Z}$ and $\ell_n(H) = \mathbb{Z}/n\mathbb{Z}$. Then we have*

$$H \cong (H(m) \ltimes H(n))/(H(mn) \ltimes H(mn)).$$

More precisely, the mapping

$$\begin{array}{rccc} \varphi\colon & H(m) \ltimes H(n) & \to & H \\ & (u, x) & \mapsto & ux \end{array}$$

is a continuous onto homomorphism of profinite groups, and the kernel of φ is the profinite group $H(mn) \ltimes H(mn)$.

Proof: The map φ is clearly continuous, and it is also straightforward to see that it is a homomorphism: $\varphi((u, x)(v, y)) = \varphi(uxvx^{-1}, xy) = uxvy = \varphi(u, x)\varphi(v, y)$.

We proceed to show that φ is onto. Let $h \in H$. Since m, n are relatively prime, there is a natural number k_0 such that, for every natural $k > k_0$, there are natural numbers k_1, k_2 such that $k = k_1 m + k_2 n$. Therefore, every word of A^* of length greater than k_0 belongs to $(A^m)^* \cdot (A^n)^*$, and so every infinite pseudoword of $\widehat{A^*}$ belongs to $\overline{(A^m)^*} \cdot \overline{(A^n)^*}$. In particular, if $h \in H$, then there are pseudowords $u \in \overline{(A^m)^*}$ and $x \in \overline{(A^n)^*}$ such that $h = ux$. Let e be the idempotent in H. Then we have $h = euxe$, whence $eu \mathrel{\mathcal{R}_A} h \mathrel{\mathcal{L}_A} xe$. Since $J(\mathcal{S})$ is regular, we may take an idempotent $f \in J(\mathcal{S})$ such that $eu = euf$.

By the Chinese Remainder Theorem, we have $\mathbb{Z}/mn\mathbb{Z} = \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ and $\ell_{mn}(w) = (\ell_m(w), \ell_n(w))$ for every $w \in \widehat{A^*}$. Hence, the equalities $\ell_m(H) = \mathbb{Z}/m\mathbb{Z}$ and $\ell_n(H) = \mathbb{Z}/n\mathbb{Z}$ entail $\ell_{mn}(H) = \mathbb{Z}/mn\mathbb{Z}$. In particular, $\mathcal{S}^{mn}$ is irreducible by Corollary 4.7. Therefore, letting $X = \mathcal{B}(\mathcal{S}) \cap A^{mn}$, we may consider the $\mathcal{J}_X$-class $J(\mathcal{S}^{mn})$. Note that $\ell_{mn}(e) = \ell_{mn}(f) = 0$, and thus e, f belong to $\overline{\mathcal{B}(\mathcal{S})} \cap \widehat{X^*}$, that is, $e, f \in \overline{\mathcal{B}(\mathcal{S}^{mn})}$. It follows from Proposition 6.2 that e, f are both in the $\mathcal{J}_X$-class $J(\mathcal{S}^{mn})$. Hence, there are $z, t \in \widehat{X^*}$ such that $f = fzetf$.

Consider the pseudowords $u' = eufze$ and $x' = etfxe$. We have

$$u'x' = eu \cdot fzetf \cdot xe = euf \cdot xe = euxe = h.$$

Note that u', x' are $\mathcal{H}$-equivalent to e. As $\ell_{mn}(\widehat{X^*}) = 0$ and $\ell_m(u) = 0$, we have $\ell_m(u') = \ell_m(eufze) = 0$. Similarly, we also have $\ell_n(x') = 0$. Therefore, the pair (u', x') belongs to $H(m) \ltimes H(n)$ and $\varphi(u', x') = h$, thus showing that φ is onto.

It remains to check that $\ker\varphi = H(mn) \ltimes H(mn)$. Take $(u,x) \in H(m) \ltimes H(n)$ for which $\varphi(u,x) = e$. Observe that $0 = \ell_n(ux) = \ell_n(u) + \ell_n(x) = \ell_n(u)$, which together with $\ell_m(u) = 0$ gives $\ell_{mn}(u) = 0$, that is, $u \in H(mn)$. Similarly, we also have $x \in H(mn)$. We conclude that (u,x) belongs to $H(mn) \ltimes H(mn)$. ∎

We are now ready to show the main result of this section.

Theorem 10.10. *If $\mathcal{S}_1$ and $\mathcal{S}_2$ are eventually conjugate irreducible subshifts, then $G(\mathcal{S}_1)$ and $G(\mathcal{S}_2)$ are isomorphic profinite groups.*

Proof: We may consider an alphabet A such that both $\mathcal{S}_1$ and $\mathcal{S}_2$ are subshifts of $A^{\mathbb{Z}}$ (let A for example be the set of letters appearing in some element of $\mathcal{S}_1$ or $\mathcal{S}_2$).

Since $\mathcal{S}_1$ and $\mathcal{S}_2$ are eventually conjugate, we may fix a positive integer N such that $\mathcal{S}_1^n$ and $\mathcal{S}_2^n$ are conjugate for every $n \geq N$.

For each $i \in \{1,2\}$, take a maximal subgroup H_i of $\widehat{A^*}$ contained in $J(\mathcal{S}_i)$. Consider the set Q_i of prime numbers greater than N and such that $\ell_p^A(H_i) = 0$.

Lemma 10.11. *The sets $Q_2 \setminus Q_1$ and $Q_1 \setminus Q_2$ have at most one element.*

Proof: Suppose the lemma is false. Without loss of generality, suppose that $Q_2 \setminus Q_1$ has two distinct elements p, q. Since $\ell_p^A(H_1) \neq 0$ and p is prime, we must have $\ell_p^A(H_1) = \mathbb{Z}/p\mathbb{Z}$. By Corollary 4.7, the subshift S_1^p is irreducible. Furthermore, the intersection $H_1 \cap \overline{(A^p)^*}$ is a maximal subgroup of the $\mathcal{J}_{A^p}$-class $J(S_1^p)$ by Proposition 6.2. Because S_1^p and S_2^p are conjugate (as $p > N$), the subshift S_2^p is also irreducible. Moreover, since $\ell_p^A(H_2) = 0$, it follows from Lemma 10.6 that H_2 is a maximal subgroup of the $\mathcal{J}_{A^p}$-class $J(\mathcal{S}_2^p)$. Therefore, by Theorem 10.3, there is a continuous isomorphism $\psi\colon H_2 \to H_1 \cap \overline{(A^p)^*}$ such that $\ell^{A^p} \circ \psi = \ell^{A^p}|_{H_2}$. Hence, applying Lemma 10.5, we have the chain of equivalences

$$\begin{aligned}
\ell_q^A(H_2) = 0 &\Leftrightarrow \ell_q^{A^p}(H_2) = 0 \qquad (\text{because } H_2 \subseteq \overline{(A^p)^*}, \text{ as } p \in Q_2)\\
&\Leftrightarrow \ell_q^{A^p}(H_1 \cap \overline{(A^p)^*}) = 0 \qquad (\text{as } \ell^{A^p} \circ \psi = \ell^{A^p}|_{H_2})\\
&\Leftrightarrow H_1 \cap \overline{(A^p)^*} \subseteq \overline{((A^p)^q)^*}.
\end{aligned}$$

Since $q \in Q_2$, we indeed have $\ell_q^A(H_2) = 0$, whence

$$H_1 \cap \overline{(A^p)^*} \subseteq H_1 \cap \overline{(A^q)^*}. \tag{10.1}$$

As $q \notin Q_1$, we may take some $h \in H_1$ such that $\ell_q^A(h) \neq 0$. Because $p \neq q$, we must have $p\ell_q^A(h) \neq 0$, that is, $h^p \notin \overline{(A^q)^*}$. But this contradicts (10.1).

To avoid a contradiction, each of the sets $Q_2 \setminus Q_1$ and $Q_1 \setminus Q_2$ must have at most one element. ∎

We proceed with the proof of the theorem.

Suppose first that at least one of the sets Q_1 and Q_2 is infinite. Then $Q_1 \cap Q_2$ is infinite, by Lemma (10.11). Let $p \in Q_1 \cap Q_2$. As $p > N$, the subshifts $\mathcal{S}_1^p$ and $\mathcal{S}_2^p$ are conjugate. Since $\ell_p(H_i) = 0$, it follows from Lemma 10.6 that H_i is a maximal subgroup of the $\mathcal{J}_{A^p}$-class of $J(\mathcal{S}_i^p)$, for each $i \in \{1, 2\}$. Applying Theorem 10.2, we conclude that $H_1 \cong H_2$, settling the theorem for this case.

We now consider the remaining case where Q_1 and Q_2 are finite. Then there are distinct prime numbers p, q greater than N such that $p, q \notin Q_1 \cup Q_2$. Let $r \in \{p, q\}$. Since r is prime, we must have $\ell_r(H_i) = \mathbb{Z}/r\mathbb{Z}$, for each $i \in \{1, 2\}$. By Corollary 4.7, the subshift $\mathcal{S}_i^r$ is irreducible and, in view of Proposition 6.2, the profinite group $H_i(r)$ is a maximal subgroup of the $\mathcal{J}_{A^r}$-class $J(\mathcal{S}_i^r)$. Therefore, since S_1^r is conjugate to S_2^r (as $r > N$), it follows from Theorem 10.3 that there is a continuous isomorphism $\psi_r \colon H_1(r) \to H_2(r)$ such that

$$\ell^{A^r} \circ \psi_r = \ell^{A^r}|_{H_1(r)} \qquad \text{(for each } r \in \{p, q\}). \tag{10.2}$$

Since p, q are distinct primes, for each $i \in \{1, 2\}$ we have

$$H_i(pq) = H_i(p) \cap H_i(q) = H_i(p) \cap \overline{(A^q)^*}.$$

Therefore, for every $h \in H_1(p)$, we have

$$\begin{aligned}
\psi_p(h) \in H_2(pq) &\Leftrightarrow \psi_p(h) \in H_2(q) = H \cap \overline{(A^q)^*}. \\
&\Leftrightarrow \ell_q^A(\psi_p(h)) = 0 \\
&\Leftrightarrow \ell_q^{A^p}(\psi_p(h)) = 0 \qquad \text{(by Lemma 10.5, as } \psi_p(h) \in \overline{(A^p)^*}) \\
&\Leftrightarrow \ell_q^{A^p}(h) = 0 \qquad \text{(by (10.2))} \\
&\Leftrightarrow \ell_q^A(h) = 0 \qquad \text{(by Lemma 10.5, as } h \in \overline{(A^p)^*}) \\
&\Leftrightarrow h \in H_1(p) \cap \overline{(A^q)^*} \\
&\Leftrightarrow h \in H_1(pq).
\end{aligned}$$

Similarly, for every $h \in H_1(q)$, the equivalence

$$\psi_q(h) \in H_2(pq) \Leftrightarrow h \in H_1(pq)$$

holds. Therefore, the isomorphism $\psi\colon H_1(p) \ltimes H_1(q) \to H_2(p) \ltimes H_2(q)$ given by the formula $\psi(u,x) = (\psi_p(u), \psi_q(x))$ satisfies

$$\psi(u,x) \in H_2(pq) \ltimes H_2(pq) \Leftrightarrow (u,x) \in H_1(pq) \ltimes H_1(pq).$$

Since $H_2(pq) \ltimes H_2(pq)$ is a closed normal subgroup of $H_2(p) \ltimes H_2(q)$ (cf. Remark 10.8), it follows that

$$(H_1(p) \ltimes H_1(q))/(H_1(pq) \ltimes H_1(pq)) \cong (H_2(p) \ltimes H_2(q))/(H_2(pq) \ltimes H_2(pq)).$$

Applying Lemma 10.9, we conclude that $H_1 \cong H_2$. ■

10.3. A digression on relatively free profinite monoids. We finish by explaining how to generalize Theorem 10.10 to *relatively* free profinite monoids over certain pseudovarieties. The reader may wish to consult [RS09] or [Alm05b] for background on pseudovarieties of monoids (or semigroups) and their relatively free objects. A *pseudovariety of monoids* V is a class of finite monoids closed under taking submonoids, homomorphic images, and finite direct products. If H is a pseudovariety of finite groups, then the class $\overline{\mathsf{H}}$ of finite monoids whose subgroups belong to H is a pseudovariety of monoids. A *pro-*V *monoid* is an inverse limit of monoids from V, in the category of compact monoids. The class of A-generated compact monoids has free objects (as we have done in the rest of the paper, we assume that A is finite). We let $\widehat{F}_{\mathsf{V}}(A)$ be the *free pro-*V *monoid* generated by A. Note that for the pseudovariety M of all finite monoids, we have $\widehat{F}_{\mathsf{M}}(A) = \widehat{A^*}$.

Let H be a pseudovariety of finite groups that is extension-closed (i.e., closed under taking semidirect products). Roughly speaking, from the viewpoint adopted in this paper, $\widehat{F}_{\overline{\mathsf{H}}}(A)$ behaves pretty much like $\widehat{A^*}$ does.[§] For example, the free semigroup A^* embeds densely in $\widehat{F}_{\overline{\mathsf{H}}}(A)$ and the elements of A^* are isolated in $\widehat{F}_{\overline{\mathsf{H}}}(A)$. A version of Theorem 2.1 holds, where we replace "*rational languages* $L \subseteq A^*$" by "*languages* $L \subseteq A^*$ *recognized by monoids in* $\overline{\mathsf{H}}$". The hypothesis that H is extension-closed is needed to guarantee that the analog of Theorem 4.3, which is Corollary 2.2 of the paper [MSW98],

[§]When checking references in the literature, the reader should bear in mind that $\overline{\mathsf{H}}$ is often viewed as a *pseudovariety of semigroups*, namely the pseudovariety of finite semigroups whose subgroups belong to H. The two perspectives are essentially the same, for the purposes of this paper. As a pseudovariety of finite semigroups, $\overline{\mathsf{H}}$ is a *monoidal* pseudovariety of monoids, that is, it is generated by the monoids in $\overline{\mathsf{H}}$. In particular, the free profinite *semigroup* over the pseudovariety of *semigroups* $\overline{\mathsf{H}}$ is the free profinite *monoid* over the pseudovariety of *monoids* $\overline{\mathsf{H}}$ minus the empty word, just as we have the equality $\widehat{A^+} = \widehat{A^*} \setminus \{\varepsilon\}$.

holds for codes $X \subseteq A^+$ such that X^* is recognized by a monoid in $\overline{\mathsf{H}}$. Also, roughly speaking, we may say that all the "combinatorics on pseudowords" used so far in this paper also works in $\widehat{F}_{\overline{\mathsf{H}}}(A)$: we have a version of Proposition 2.2 (details may be found in [CS21, Section 2]), and also a version of Lemma 6.1 (cf. [ACCZ19, Section 3]).

If $\mathcal{S}$ is an irreducible subshift of $A^{\mathbb{Z}}$, then we may still consider an analog of $G(\mathcal{S})$ in $\widehat{F}_{\overline{\mathsf{H}}}(A)$, denoted $G_{\overline{\mathsf{H}}}(\mathcal{S})$, which is actually the image of $G(\mathcal{S})$ in $\widehat{F}_{\overline{\mathsf{H}}}(A)$ under the canonical projection $\widehat{A^*} \to \widehat{F}_{\overline{\mathsf{H}}}(A)$. It turns out that $G_{\overline{\mathsf{H}}}(\mathcal{S})$ is also invariant under conjugacy, a fact observed in [Cos06], and crucially used to show the main result of [CS11], which states that $G_{\overline{\mathsf{H}}}(\mathcal{S})$ is a free pro-H group of rank $\aleph_0$ if $\mathcal{B}(\mathcal{S})$ is recognized by a monoid from H and H is an extension-closed pseudovariety of groups containing $\mathbb{Z}/p\mathbb{Z}$ for infinitely many primes p. In fact, $G_{\overline{\mathsf{H}}}(\mathcal{S})$ is even invariant under flow equivalence, for every irreducible subshift $\mathcal{S}$ [CS21].

With minimal adaptations, the proof of Theorem 10.10 yields the following generalization.

Theorem 10.12. *Let H be an extension-closed pseudovariety of finite groups such that H contains $\mathbb{Z}/p\mathbb{Z}$ for infinitely many primes p. If $\mathcal{S}_1$ and $\mathcal{S}_2$ are eventually conjugate irreducible subshifts, then $G_{\overline{\mathsf{H}}}(\mathcal{S}_1)$ and $G_{\overline{\mathsf{H}}}(\mathcal{S}_2)$ are isomorphic profinite groups.*

Proof: Since the language $(A^n)^*$ is recognized by $\mathbb{Z}/n\mathbb{Z}$ for every positive integer n, our proof that $G(\mathcal{S})$ is invariant under eventual conjugacy also works in this more general setting. It suffices to pay attention to the following adaptations.

- We replace the natural projection $\ell^A\colon \widehat{A^*} \to \widehat{\mathbb{Z}}$, by the natural projection $\ell^{A,\overline{\mathsf{H}}}\colon \widehat{F}_{\overline{\mathsf{H}}}(A) \to \widehat{\mathbb{Z}}_{\mathsf{H}}$ where $\widehat{\mathbb{Z}}_{\mathsf{H}}$ is the inverse limit of cyclic groups from H;
- For each $n \in \mathbb{N}$ such that $\mathbb{Z}/n\mathbb{Z} \in \mathsf{H}$, we replace the natural projection $\ell^A_n\colon \widehat{A^*} \to \mathbb{Z}/n\mathbb{Z}$, by the natural projection $\ell^{A,\overline{\mathsf{H}}}_n\colon \widehat{F}_{\overline{\mathsf{H}}}(A) \to \mathbb{Z}/n\mathbb{Z}$;
- We use only positive integers n such that $\mathbb{Z}/n\mathbb{Z} \in \mathsf{H}$. In particular, the set Q_i used in the proof of Lemma 10.11 is now the set of all primes p such that $\mathbb{Z}/p\mathbb{Z} \in \mathsf{H}$ and $\ell^{A,\overline{\mathsf{H}}}_p(H_i) = 0$. This requirement explains why we need that H contains $\mathbb{Z}/p\mathbb{Z}$ for infinitely many primes p: it is to ensure that the proof of Theorem 10.10 works in the case where Q_1 and Q_2 are finite, more precisely to guarantee that if Q_1 and Q_2

are finite, then there are distinct prime numbers p, q not belonging to $Q_1 \cup Q_2$ and greater than the threshold N such that both $\mathbb{Z}/p\mathbb{Z}$ and $\mathbb{Z}/q\mathbb{Z}$ belong to H.

All arguments adapted from the proof of Theorem 10.10 rely on the aforementioned generalizations/adaptations of Theorem 2.1, Proposition 2.2, Lemma 6.1 and Theorem 4.3, which entail transparent generalizations of the results of the previous sections on which Theorem 10.10 depends. ■

References

[AC09] J. Almeida and A. Costa, *Infinite-vertex free profinite semigroupoids and symbolic dynamics*, J. Pure Appl. Algebra **213** (2009), 605–631.

[AC13] ———, *Presentations of Schützenberger groups of minimal subshifts*, Israel J. Math. **196** (2013), 1–31.

[AC16] ———, *A geometric interpretation of the Schützenberger group of a minimal subshift*, Ark. Mat. **54** (2016), 243–275.

[ACCZ19] J. Almeida, A. Costa, J. C. Costa, and M. Zeitoun, *The linear nature of pseudowords*, Publ. Mat. **63** (2019), 361–422.

[ACKP20a] J. Almeida, A. Costa, R. Kyriakoglou, and D. Perrin, *On the group of a rational maximal bifix code*, Forum Math. **32** (2020), 553–576.

[ACKP20b] ———, *Profinite semigroups and symbolic dynamics*, Lect. Notes in Math., vol. 2274, Springer, Cham, 2020.

[Alm95] J. Almeida, *Finite semigroups and universal algebra*, World Scientific, Singapore, 1995, English translation.

[Alm02] ———, *Dynamics of finite semigroups*, Semigroups, Algorithms, Automata and Languages (Singapore) (G. M. S. Gomes, J.-E. Pin, and P. V. Silva, eds.), World Scientific, 2002, pp. 269–292.

[Alm05a] ———, *Profinite groups associated with weakly primitive substitutions*, Fundamentalnaya i Prikladnaya Matematika (Fundamental and Applied Mathematics) **11** (2005), 13–48, In Russian. English version in J. Math. Sciences **144**, No. 2 (2007) 3881–3903.

[Alm05b] ———, *Profinite semigroups and applications*, Structural theory of automata, semigroups and universal algebra (New York) (V. B. Kudryavtsev and I. G. Rosenberg, eds.), Springer, 2005, pp. 1–45.

[AV06] J. Almeida and M. V. Volkov, *Subword complexity of profinite words and subgroups of free profinite semigroups*, Int. J. Algebra Comput. **16** (2006), 221–258.

[BDFD+15a] V. Berthé, C. De Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer, and G. Rindone, *Acyclic, connected and tree sets*, Monatsh. Math. **176** (2015), 521–550.

[BDFD+15b] ———, *The finite index basis property*, J. Pure Appl. Algebra **219** (2015), 2521–2537.

[BDFD+15c] ———, *Maximal bifix decoding*, Discrete Math. **338** (2015), no. 5, 725–742.

[BDFP+12] J. Berstel, C. De Felice, D. Perrin, Ch. Reutenauer, and G. Rindone, *Bifix codes and Sturmian words*, J. Algebra **369** (2012), 146–202.

[BF77] R. Bowen and J. Franks, *Homology for zero-dimensional nonwandering sets*, Ann. of Math. (2) **106** (1977), no. 1, 73–92.

[Boy00] M. Boyle, *Algebraic aspects of symbolic dynamics*, Topics in Symbolic Dynamics and Applications (Temuco, 1997) (Cambridge) (F. Blanchard, A. Maass, and A. Nogueira, eds.), London Math. Soc. Lecture Note Ser., no. 279, Cambridge Univ. Press, 2000.

[Boy02] ———, *Positive K-theory and symbolic dynamics*, Dynamics and randomness (Santiago, 2000), Nonlinear Phenom. Complex Systems, vol. 7, Kluwer Acad. Publ., Dordrecht, 2002, pp. 31–52.

[Boy22] ———, *Personal communication*, 2022.

[BPR10] J. Berstel, D. Perrin, and Ch. Reutenauer, *Codes and automata*, Encyclopedia of Mathematics and its Applications, vol. 129, Cambridge University Press, Cambridge, 2010.

[Cos06] A. Costa, *Conjugacy invariants of subshifts: an approach from profinite semigroup theory*, Int. J. Algebra Comput. **16** (2006), 629–655.

[CS11] A. Costa and B. Steinberg, *Profinite groups associated to sofic shifts are free*, Proc. London Math. Soc. **102** (2011), 341–369.

[CS21] ———, *The Karoubi envelope of the mirage of a subshift*, Comm. Algebra **49** (2021), 4820–4856.

[Fog02] N. Pytheas Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lect. Notes in Math., vol. 1794, Springer-Verlag, Berlin, 2002.

[Fra84] John Franks, *Flow equivalence of subshifts of finite type*, Ergodic Theory Dynam. Systems **4** (1984), no. 1, 53–66. MR **MR758893 (86j:**58078)

[GO21] H. Goulet-Ouellet, *Freeness of Schützenberger groups of primitive substitutions*, arXiv preprint arXiv:2109.11957 (2021), To appear.

[GO22] ———, *Pronilpotent quotients associated with primitive substitutions*, J. Algebra **606** (2022), 341–370.

[HRS10] K. Henckell, J. Rhodes, and B. Steinberg, *A profinite approach to stable pairs*, Int. J. Algebra Comput. **20** (2010), 269–285.

[Hun83] R. P. Hunter, *Some remarks on subgroups defined by the Bohr compactification*, Semigroup Forum **26** (1983), 125–137.

[KP17] R. Kyriakoglou and D. Perrin, *Profinite semigroups*, Tech. report, 2017, arXiv:1703.10088 [math.GR].

[KR92] K. H. Kim and F. W. Roush, *Williams conjecture is false for reducible subshifts*, J. Amer. Math. Soc. **5** (1992), 213–215.

[KR99] ———, *The Williams conjecture is false for irreducible subshifts*, Ann. of Math. (2) **149** (1999), 545–558.

[LM95] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995.

[MSW98] S. Margolis, M. Sapir, and P. Weil, *Irreducibility of certain pseudovarieties*, Comm. Algebra **26** (1998), 779–792.

[Per18] D. Perrin, *Groups, languages and dendric shifts*, Developments in Language Theory. DLT 2018. (Hoshi M. and Seki S., eds.), Lect. Notes in Comput. Sci., vol. 11088, Springer, Cham, 2018.

[PS75] B. Parry and D. Sullivan, *A topological invariant of flows on* 1*-dimensional spaces*, Topology **14** (1975), no. 4, 297–299.

[RS09] J. Rhodes and B. Steinberg, *The q-theory of finite semigroups*, Springer Monographs in Mathematics, Springer, 2009.

[RZ10] L. Ribes and P. A. Zalesskii, *Profinite groups*, second ed., Ergeb. Math. Grenzgebiete. 3, vol. 40, Springer-Verlag, Berlin, 2010.

[Wil73] R. F. Williams, *Classification of subshifts of finite type*, Ann. of Math. (2) **98** (1973), 120–153, Erratum, Ann. of Math. 99 (1974), 380-381.

Alfredo Costa
University of Coimbra, CMUC, Department of Mathematics, Apartado 3008, EC Santa Cruz, 3001-501 Coimbra, Portugal.

E-mail address: amgc@mat.uc.pt