CONJUGACY CLASSIFICATION FOR ONE-WORD SHIFTS OF FINITE TYPE

Abstract. We classify the shifts of finite type obtained by forbidding one word in a full shift over a fixed finite alphabet. Two irreducible such shifts are conjugate exactly when their forbidden words have the same border lengths. We construct conjugacies extending to full-shift automorphisms that preserve occurrence positions. The proof is an induction on the longest proper border, with explicit matrix factorizations for the binary base case. We also classify the reducible binary exceptions and describe the conjugacies by finite graph splittings and amalgamations.

1. Introduction

Consider the binary sequences that avoid 001001, and those that avoid 011011. In each forbidden word, the first three letters agree with the last three, and there is no shorter nonempty self-overlap. Does this common overlap structure allow one system of sequences to be transformed into the other by an invertible local rule? We prove that the same question has a positive answer whenever the words have the same border lengths and the shifts obtained by forbidding them are irreducible. The proof constructs the local rules and explains how to express them as maps between finite graphs.

Let Σ be a finite alphabet of size $q \geq 2$, and let $\sigma : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ be the left shift, $(\sigma x)_i = x_{i+1}$. For a nonempty word w, write

$$X_q(w) = \{x \in \Sigma^{\mathbb{Z}} : w \text{ occurs nowhere in } x\}.$$

This is a *shift of finite type*: a closed, shift-invariant space specified by finitely many forbidden words. We give these spaces the product topology. A conjugacy between two shift spaces is a homeomorphism commuting with σ; an automorphism is a conjugacy from a shift space to itself. Equivalently, it is an invertible sliding block code: each output coordinate is determined by a fixed finite window of input coordinates, and the inverse has the same property. A shift is *irreducible* if any two words appearing in it can be joined by a finite word so that their concatenation also appears in it.

Write $w_{[i,j)} = w_i \cdots w_{j-1}$ for a block, with the analogous notation for bi-infinite sequences. A *border* of $w = w_0 \cdots w_{n-1}$ is a nonempty word that is both a prefix and a suffix of w. We record its length, including the full length:

$$B(w) = \{b \in \{1, \ldots, n\} : w_{[0,b)} = w_{[n-b,n)}\}, \qquad \Phi_w(t) = \sum_{b \in B(w)} t^b.$$

Thus $B(001001) = B(011011) = \{3, 6\}$. A word is *unbordered* if its only border is itself. The polynomial Φ_w is the autocorrelation polynomial, with the normalization displayed above.

Theorem 1.1. *Fix $q \geq 2$. For nonempty words u, v such that $X_q(u)$ and $X_q(v)$ are irreducible, the following are equivalent:*

1. *$X_q(u)$ and $X_q(v)$ are conjugate;*
2. *they have the same number of points fixed by σ^r for every integer $r \geq 1$;*
3. *$B(u) = B(v)$, or equivalently $\Phi_u = \Phi_v$.*

When these conditions hold, $|u| = |v| =: n$, and there is an automorphism H of the full shift $\Sigma^{\mathbb{Z}}$ such that, for every x and i,

$$x_{[i,i+n)} = u \quad \Longleftrightarrow \quad H(x)_{[i,i+n)} = v.$$

Date: Draft, 8 November 2026.

In particular, the restriction of H is a conjugacy from $X_q(u)$ to $X_q(v)$. The word lengths need not be fixed in advance.

For $q \geq 3$, every one-word shift is irreducible. Over the binary alphabet, the reducible words of length $n \geq 2$ are precisely $0^{n-1}1$, $1^{n-1}0$, 10^{n-1}, and 01^{n-1}, with repetitions at $n = 2$. They are unbordered. For $n \geq 3$ they form two conjugacy classes. In each shift, a unique fixed point is isolated among the periodic points. The classes differ according to whether a nonconstant point can agree with this fixed point at all sufficiently large coordinates. The irreducibility hypothesis is automatic outside this unbordered binary case. For example, 0001 and 0011 have the same border set, but their avoiding shifts cannot be conjugate: the first is reducible and the second is irreducible.

Chandgotia, Marcus, Richey and Wu studied this conjugacy problem and constructed conjugacies between unbordered words [2, Section 4]. We use their chains of word swaps, in which occurrences of the two words being exchanged are disjoint, as a base case. The additional step here is to pass from words without proper borders to arbitrary border sets. This requires a stronger local statement than conjugacy of the two avoiding shifts: the maps must also act predictably on longer words that contain the words being transported at both ends. For binary words, explicit matrix factorizations supply that statement across the exceptional unbordered classes.

1.1. **Overview of the proof.** The construction has two steps: change the longest proper border, then change the letters between its two endpoint copies. The first step is recursive; the second uses simultaneous word replacements. Both must control the images of finite words independently of their exteriors.

For a word w of length n, let $[w]$ be the cylinder of sequences whose coordinates $0, \ldots, n-1$ spell w. Our maps are automorphisms of the full shift satisfying $H([u]) = [v]$. Commutation with the shift then gives the same equivalence at every position, so H restricts to a conjugacy from $X_q(u)$ to $X_q(v)$. Working on the full shift also lets us apply a map for a short word inside sequences containing that word, as the recursive construction will require.

The induction retains the following more precise assertion. If H is the map for words p, p' of length m, then for every word w of length $N \geq m$ beginning and ending in p, there is a word w' with

$$H([w]) = [w'], \qquad |w'| = N, \qquad w' \text{ begins and ends in } p'.$$

The same assertion holds for H^{-1}, with p, p' interchanged. Thus the two specified endpoint words ensure that the output inside the interval is independent of the input outside it. The endpoint copies may overlap or coincide. We call this the *two-sided cylinder property*. It gives a bijection between finite words with the specified endpoints, at every length, as well as a conjugacy of the avoiding shifts.

Here is how this property makes the induction work. Suppose $B(u) = B(v)$, with $|u| = |v| = n$, and let p, p' be their longest proper borders, of length m. Every shorter border of u is a border of p, and conversely, since the relevant prefixes and suffixes lie in the endpoint copies of p. Hence

$$B(p) = B(u) \setminus \{n\} = B(v) \setminus \{n\} = B(p').$$

Assume a map H for p, p' has already been constructed with the cylinder property. It gives $H([u]) = [z]$ for a length-n word z beginning and ending in p'. The cylinder identity also preserves every border length: a length-b overlap is possible exactly when $[u] \cap \sigma^{-(n-b)}[u]$ is nonempty, and H preserves these intersections. Consequently $B(z) = B(u) = B(v)$, and z, v have the *same actual longest border* p'. A comparison with that border fixed will finish the step.

Figure 1 shows this reduction over $\{0, 1, 2\}$. The first map swaps 010 and 020 wherever they occur: it interchanges 1 and 2 exactly when their two neighbors are zero. Every zero stays fixed, so applying the rule twice restores the sequence. For a word beginning and ending in 010, the boundary zeros stay fixed and every other output letter is determined by neighbors inside the word. This proves the cylinder property directly, including overlapping endpoints, as in $01010 \mapsto 02020$.

In the figure the first map changes the border 010 to 020; the second changes the intervening 00 to 11 inside the indicated length-eight occurrences.

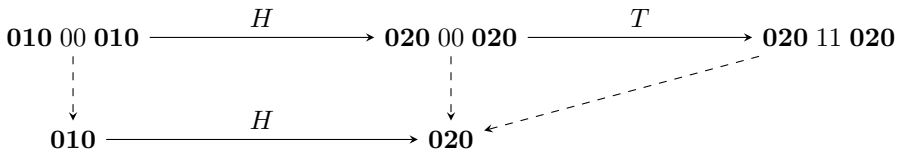


FIGURE 1. One recursive step. The upper words all have border set $\{1, 3, 8\}$; the lower words have border set $\{1, 3\}$. Solid arrows specify cylinder images under full-shift automorphisms. Dashed arrows take the longest proper border, shown in bold. After H transports that border, the two words on the right can be exchanged by a single swap T.

To understand the second map, consider two distinct length-n words r, t. An *ordered proper overlap* of r with t is a length $1 \le b < n$ at which a suffix of r equals a prefix of t. Call the pair *compatible* when the four ordered overlap sets, for $(r, r), (r, t), (t, r), (t, t)$, all agree. At each permitted length, the prefixes and suffixes of both words are the same. Thus simultaneous replacement of every r by t and every t by r leaves all intersections of occurrences unchanged. A newly created occurrence would contain a changed letter. Choose an original replacement interval containing that letter. These two intervals are distinct, and in the output both spell one of the swapped words. Their intersection is therefore a permitted proper overlap, which the original replacement leaves unchanged—a contradiction. The occurrence positions are preserved, so swapping again recovers the original sequence.

The same observation supplies the cylinder property. An occurrence crossing the left boundary of a word beginning in r meets its initial copy of r in a permitted overlap. Replacement fixes every letter of this intersection. The right boundary is identical, so the entire output word is determined internally; applying the inverse gives equality of cylinders. In Figure 1, the two right-hand words are compatible, with all four proper overlap sets equal to $\{1, 3\}$. The two outer words have the same self-overlaps but only a length-one cross-overlap. Transporting their actual border is what makes the final swap possible.

We now solve the fixed-border comparison in general. Suppose r, t have length n and the same actual longest proper border p, of length m. If $n \le 2m$, the two copies of p cover the word and force $r = t$. If $n > 2m$, put $s = n - 2m$ and choose a letter a such that

$$c = pa^s p \quad \text{contains } p \text{ only at its two endpoint positions.}$$

Call this a *safe* choice of spacer. Every proper overlap of length at most m among r, c is then exactly a border of p. An overlap of length $m < k < n$ would place an extra copy of p inside c: at position $k - m$ for an overlap from r to c, and at position $n - k$ in the other direction. The same reasoning excludes longer self-overlaps of c. Thus r, c are compatible when distinct, as are t, c. At most two swaps, through c, give the desired comparison.

The spacer exists over every alphabet of size at least two because:

> For every nonempty p and every $s \ge 1$, at most one letter a makes $pa^s p$ contain an extra occurrence of p.

For example, with $p = 010$ and $s = 1$, the spacer 1 gives 0101010, with copies starting at $0, 2, 4$; the spacer 0 gives 0100010, with copies only at $0, 4$. The reason for the general assertion is a conservation of letter counts among borders. If p is nonconstant, an extra occurrence must cross both endpoint copies. Indeed, a copy meeting only the left endpoint, at a displacement j, would give $p = p_{[j,|p|)}a^j$: the last j symbols are a, and this equality propagates them backwards through all of p. Meeting only the right endpoint gives the same conclusion by forward propagation; a copy wholly in the

spacer also forces p to be constant. The extra occurrence therefore expresses p as a border of length b, the block a^s, and a border of length d, where $b + d = |p| - s$.

Let $F(j)$ be the vector counting each letter in the first j symbols of p. For border lengths $x < y$, the first y letters end in the first x letters, so $F(y) - F(x) = F(y - x)$. If two border pairs satisfy $b + d = b' + d'$ and $b < b'$, this gives

$$F(b') - F(b) = F(b' - b) = F(d) - F(d').$$

Hence $F(b) + F(d)$ is fixed when $b + d$ is fixed. The intervening block has the fixed count vector $F(|p|) - F(b) - F(d)$, which can be the count vector of a^s for at most one letter a. For a constant word p, any other letter is safe. This proves the spacer assertion and completes the fixed-border construction (Lemmas 3.5–3.6).

Returning to the induction, compose the map H for the longest borders with these swaps to send $[u]$ onto $[v]$. The cylinder property is preserved: a longer word beginning and ending in u also begins and ends in p, so H determines its image internally. Its endpoint occurrences of u become occurrences of z at the same positions, by $H([u]) = [z]$ and shift commutation. The swaps then determine an image with endpoints v. The inverse has the identical property. If the common border set is $\{b_1 < \cdots < b_k = n\}$, we therefore construct maps between the corresponding prefixes at lengths $b_1, b_2, \ldots, b_k$, in that order. The length-b_1 prefixes are unbordered. For $q \geq 3$, the unbordered replacement chains of [2, Section 4] supply the first map; each subsequent step adds the next border length. This proves sufficiency for every border set over those alphabets.

The binary case requires a different starting point. The opening pair $001001, 011011$ shows why: their shortest borders $001, 011$ give nonconjugate reducible avoiding shifts, whereas the two longer words give irreducible shifts. We start instead with words having exactly one proper border. If their border set is $\{m, n\}$, then $n = 2m + s$ with $s \geq 0$, and the fixed-border comparison reduces each to pa^sp, where p is unbordered of length m. For $m \geq 2$ every constant spacer is safe, since p has no proper borders with which an extra copy could overlap. In the unbordered chains, the words being swapped have all four proper overlap sets empty. The only occurrences of either base inside pa^sp are then its two endpoints, so the swap sends it to $p'a^sp'$. The additional comparison needed for $m \geq 3$ is

$$W = p0^sp, \qquad p = 0^{m-1}1,$$
$$W' = p'0^sp', \qquad p' = 0^{m-2}11.$$

For $m = 3, s = 0$ this is exactly $001001 \leftrightarrow 011011$. Together with the unbordered chains, complementation, and the reversed construction, these comparisons join all words with one proper border. The cases $m = 1, 2$ use letter permutations and fixed-border swaps.

We obtain the maps for W, W' from finite graphs and matrix factorizations. A nonnegative integer matrix A represents the graph with A_{ij} edges from vertex i to vertex j. The elementary operation has a direct local meaning: when $A = RS$ and $B = SR$ with nonnegative integer factors, regard an edge of the graph of A as two consecutive edges (r_t, s_t), through an intermediate vertex. Regrouping gives

$$\cdots (r_t, s_t)(r_{t+1}, s_{t+1}) \cdots \quad \longmapsto \quad \cdots (s_t, r_{t+1})(s_{t+1}, r_{t+2}) \cdots .$$

This is an invertible local map between bi-infinite paths; the inverse regroups the same half-edges one position earlier. A finite sequence of such factor reversals constructs the required conjugacy. The issue for our induction is the missing half-edge at the end of a *finite* path, where regrouping could otherwise use the exterior.

To control it, encode a binary sequence in a graph whose state records the longest prefix of W matching the end of the past, including the completed word as a state. Every binary sequence has a unique path in this graph. Mark the edge that completes W; marks record exactly its completion times. The factorizations are chosen to transport this mark and to force the missing half-edges at marked endpoints. More precisely, one factor in each reversal is $U = I + E_{ij}$, where E_{ij} has a single

1 at (i, j) and zeros elsewhere. The edges of U are the diagonal edges together with one extra edge from i to j, with $i \neq j$. The other factor contains the marked edge, from ℓ to β; we choose $j \neq \ell$ and $i \neq \beta$. The only edge of U entering ℓ or leaving β is then the corresponding diagonal edge. Those are precisely the otherwise missing halves at the two ends. Thus a finite path between marked edges determines its whole image between the corresponding marks, in both directions (Lemma 4.1).

The constructed sequence takes the graph for W to that for W'. For $001001, 011011$, it consists of seven reversals, four of which move every mark one position earlier. Shifting the output four positions to the right aligns the occurrences. The general construction likewise has a fixed displacement, which we correct by a shift. A word beginning and ending in W determines the path between the completions of these two designated endpoint copies. The finite-path property determines the output between those marks, and the first designated output mark forces the entire initial copy of W'. Together these determine every output letter in the original interval. Applying the inverse gives equality of cylinders. For example, the aligned map for the opening pair satisfies $H([001001001]) = [011011011]$, where the endpoint copies overlap. This is the two-sided cylinder property needed by the induction. The binary construction therefore supplies exactly the stronger starting map that the border recursion uses.

Finally, periodic points recover the proposed invariant. Writing $p_r = |\{x \in X_q(w) : \sigma^r x = x\}|$, we prove

$$D_w(t) := \exp\left(-\sum_{r\geq 1} \frac{p_r}{r} t^r\right) = (1-qt) \sum_{b\in B(w)} t^{n-b} + t^n, \qquad n = |w|.$$

The degree is n, and $(D_w(t) - t^n)/(1-qt)$ records every border length. A conjugacy preserves all p_r, hence the length and border set. This gives necessity and completes the classification.

2. Periodic points and irreducibility

The constructions for Theorem 1.1 use equality of border sets. We first show that a conjugacy requires this equality, even if the forbidden words initially have different lengths. Periodic points give this information without any irreducibility assumption.

For a shift X, put

$$p_r(X) = |\{x \in X : \sigma^r x = x\}|, \qquad \zeta_X(t) = \exp\left(\sum_{r\geq 1} \frac{p_r(X)}{r}\, t^r\right).$$

The series counts points whose period divides r, rather than points of least period r. We regard ζ_X as a formal power series.

Proposition 2.1. *For every $w \in \Sigma^n$, the reciprocal zeta function is the polynomial*

$$D_w(t) := \zeta_{X_q(w)}(t)^{-1} = (1-qt) \sum_{b\in B(w)} t^{n-b} + t^n. \tag{1}$$

It has degree exactly n. Consequently, at fixed q, equality of zeta functions is equivalent to equality of full border sets, and conjugacy implies both equal word lengths and equal border sets.

Proof. We give the determinant argument for the single-word identity [2, Eq. (2)]. If $n = 1$, the shift is the full shift on $q-1$ symbols, so $p_r = (q-1)^r$ and $D_w(t) = 1-(q-1)t$, as required.

Suppose $n \geq 2$. Form the directed graph with vertices Σ^{n-1} and one edge for each word of length n, from its length-$(n-1)$ prefix to its length-$(n-1)$ suffix. Let M be its adjacency matrix, whose entries count directed edges. Deleting the edge corresponding to w gives

$$A = M - e_\alpha e_\beta^{\mathsf{T}}, \qquad \alpha = w_{[0,n-1)}, \quad \beta = w_{[1,n)},$$

where e_v denotes the coordinate vector at vertex v. Bi-infinite paths are in bijection with their symbol sequences, so $p_r(X_q(w)) = \operatorname{tr}(A^r)$. The formal identity $\log\det(I - tA) = -\sum_{r\geq 1} \operatorname{tr}(A^r)t^r/r$

gives $D_w(t) = \det(I - tA)$. Similarly, $\operatorname{tr}(M^r) = q^r$ and $\det(I - tM) = 1 - qt$. The determinant formula for a rank-one change therefore yields

$$D_w(t) = (1 - qt)\left(1 + t\, e_\beta^{\mathsf{T}}(I - tM)^{-1}e_\alpha\right).$$

For $0 \le r < n - 1$, a path of length r from β to α exists precisely when $n - 1 - r \in B(w)$, and is then unique. For $r \ge n - 1$, there are q^{r-n+1} such paths: the final $n - 1$ appended symbols are prescribed, and the preceding ones are free. It follows that

$$t\, e_\beta^{\mathsf{T}}(I - tM)^{-1}e_\alpha = \sum_{\substack{b \in B(w) \\ b < n}} t^{n-b} + \frac{t^n}{1 - qt},$$

which proves (1).

The coefficient of t^n in D_w is $1 - q\mathbf{1}_{\{1 \in B(w)\}}$, which is nonzero for $q \ge 2$. Thus $n = \deg D_w$, including when $n = 1$. Once n is known,

$$\frac{D_w(t) - t^n}{1 - qt} = \sum_{b \in B(w)} t^{n-b}$$

recovers every border indicator. Finally, a conjugacy bijects the points fixed by σ^r for every r. □

The remaining hypothesis in Theorem 1.1 is irreducibility. The following characterization locates the cases to which the constructive argument will apply.

Proposition 2.2. *Every $X_q(w)$ with $|w| = 1$ is irreducible. For $n \ge 2$, $X_q(w)$ is reducible exactly when $q = 2$ and*

$$w \in E_n := \{0^{n-1}1,\ 1^{n-1}0,\ 10^{n-1},\ 01^{n-1}\}.$$

In particular, every one-word shift over $q \ge 3$ symbols is irreducible, as is every binary one-word shift whose forbidden word has a proper border.

Proof. At length one we have the full shift on $q - 1 \ge 1$ symbols. For $n \ge 2$, Fact 3.3 and Proposition 3.5 of [2] supply a finite labeled graph whose bi-infinite path labels form $X_q(w)$, with a directed path between every ordered pair of vertices, in every case outside the list. Paths carrying any two allowed words can therefore be joined, proving irreducibility.

The listed binary exceptions are reducible directly from the word criterion. In $X_2(0^{n-1}1)$, no word can join 0^{n-1} to 1: the first subsequent 1 would complete the forbidden word. In $X_2(10^{n-1})$, no word can join 1 to 0^{n-1}: the last 1 before that zero block would complete the forbidden word. Both endpoint words occur in constant points of the respective shifts. Complementing symbols handles the other two cases. Every member of E_n is unbordered. □

3. Compatible replacements and recursion on borders

The recursive construction first changes the longest proper border of a word and then changes the letters between its two endpoint copies. To carry out the first step, we need more than a conjugacy between the shifts obtained by forbidding the borders. We construct automorphisms of the full shift that also determine the image of every finite word with the specified border at both ends. This finite-word property is the induction hypothesis used below.

3.1. Compatible swaps and their finite-word property.

A proper border has length strictly less than the word. For words r, s of the same length m, define the ordered proper overlap set by

$$O(r, s) = \{1 \le b < m : r_{[m-b,m)} = s_{[0,b)}\}.$$

Distinct words $p, p' \in \Sigma^m$ are *compatible* if the four sets $O(r, s)$, for $r, s \in \{p, p'\}$, are equal. If their common value is C, then for every $b \in C$ the length-b prefixes and suffixes of both words are all the

same word. When C is empty, we call the pair nonoverlapping; this excludes self-overlaps as well as overlaps between the two different words.

The following replacement rule is the compatible swap of [2, Proposition 4.11]. We include its proof to establish the finite-word property that follows.

Lemma 3.1 (Compatible swaps)**.** *For a compatible pair p, p' of length m, simultaneously replacing every occurrence of p by p' and every occurrence of p' by p defines an involutive automorphism S of the full shift. It preserves the set of occurrence positions of the pair and interchanges their types.*

Proof. Two distinct overlapping occurrences meet in a suffix of the left word and a prefix of the right word. The overlap length belongs to C, so both replacements leave every letter in their intersection unchanged. The simultaneous replacements are therefore consistent, and each original occurrence becomes the prescribed output word.

Let $y = Sx$. Suppose that an occurrence of p or p' appears in y on an interval J that was not an occurrence interval in x. Some coordinate of J must have changed. Choose such a coordinate and an original replacement interval I containing it. The intervals I, J are distinct and have the same length. In y both carry words of the pair, so their intersection has a length in C. The replacement on I leaves precisely that prefix or suffix unchanged, contradicting the choice of the coordinate. No new occurrence position appears. Interchanging the two types again, on the same intervals, recovers x. Hence S^2 is the identity.

To determine $(Sx)_i$, one need only inspect the length-m windows containing i, all of which lie in $[i - m + 1, i + m)$. This proves continuity. The definition commutes with σ, and the inverse is S itself. □

We now specify the additional property needed by the recursion. For equal-length words p, p', an automorphism H has the *two-sided cylinder property for* (p, p') if $H([p]) = [p']$ and the following two assertions hold. For every $N \geq |p|$ and every length-N word w beginning and ending in p, there is a length-N word w' beginning and ending in p' such that

$$H([w]) = [w'].$$

Conversely, every length-N word beginning and ending in p' has such a single-cylinder preimage, beginning and ending in p. The endpoint copies may overlap, or coincide when $N = |p|$.

Lemma 3.2 (Cylinder properties of swaps)**.** *The automorphism in Lemma 3.1 has the two-sided cylinder property for (p, p'). More generally, if an automorphism H sends $[w]$ onto $[w']$ and $|w| = |w'|$, then $B(w) = B(w')$.*

Proof. Fix a length-N word w beginning and ending in p. An occurrence of either swapped word that crosses the left boundary of $[0, N)$ ends inside the initial copy of p. Its intersection with this copy is a permitted proper overlap, on which replacement changes no letter. The same argument applies at the right boundary. An occurrence of length m cannot cross both boundaries, since $N \geq m$. Consequently, the output on $[0, N)$ is determined entirely by the occurrences contained in that interval, and hence by w alone. Denote this output word by w'. Its endpoint copies are p', and $S([w]) \subseteq [w']$.

Apply the same boundary argument to w' and the inverse swap. It gives $S^{-1}([w']) \subseteq [\widetilde{w}]$ for a word $\widetilde{w}$ of length N. An extension x of w satisfies $x \in S^{-1}([w'])$, so $[w] \cap [\widetilde{w}] \neq \varnothing$. These cylinders specify the same coordinates and length; therefore $\widetilde{w} = w$. The two inclusions give $S([w]) = [w']$. The argument starting with a word whose endpoints are p' proves the converse assertion as well.

For the last statement, put $n = |w|$. For each $1 \leq b \leq n$,

$$b \in B(w) \quad \Longleftrightarrow \quad [w] \cap \sigma^{-(n-b)}[w] \neq \varnothing.$$

Indeed, the two copies can be extended to a point of the full shift exactly when their overlap of length b agrees. A bijection commuting with the shift preserves these intersections and their nonemptiness. Applying H proves the border equality. □

Lemma 3.3 (Composition and inheritance)**.** *Alphabet permutations have the two-sided cylinder property. This property is preserved under composition. Moreover, suppose H has the property for (p, p'), w begins and ends in p, and $H([w]) = [w']$. Then H has the property for (w, w').*

Proof. An alphabet permutation sends every word cylinder to the cylinder obtained by permuting its letters; its inverse does the same. If H has the property for (p, p') and G for (p', p''), the successive images of a word with endpoints p are single cylinders with endpoints p' and p''. Apply the same argument to $H^{-1}G^{-1}$ for the reverse assertion.

For inheritance, put $n = |w|$. A word W of length $N \geq n$ beginning and ending in w also begins and ends in p. Thus $H([W]) = [W']$ for a length-N word W'. The containment

$$[W] \subseteq [w] \cap \sigma^{-(N-n)}[w]$$

implies $[W'] \subseteq [w'] \cap \sigma^{-(N-n)}[w']$, so W' begins and ends in w'. For the converse, w' begins and ends in p', and the identical argument with H^{-1} applies to every word with endpoints w'. This argument allows overlapping endpoint copies. □

Lemma 3.4 (Reversal)**.** *Suppose H has the two-sided cylinder property for (u, v). Let $\mathcal{R}(x)_i = x_{-i}$, and let w^{R} denote reversal of a finite word. Then $\mathcal{R}H\mathcal{R}$ is an automorphism with the property for $(u^{\mathrm{R}}, v^{\mathrm{R}})$.*

Proof. Since $\mathcal{R}\sigma = \sigma^{-1}\mathcal{R}$, conjugating H by $\mathcal{R}$ gives a homeomorphism commuting with σ. If $|w| = N$, then $\mathcal{R}([w^{\mathrm{R}}]) = \sigma^{N-1}[w]$. Consequently $H([w]) = [z]$ implies

$$(\mathcal{R}H\mathcal{R})([w^{\mathrm{R}}]) = \mathcal{R}\sigma^{N-1}[z] = [z^{\mathrm{R}}].$$

Apply this identity to each word with the prescribed endpoints and then to H^{-1}. It proves the assertion in both directions, with no translation of occurrence positions. Reversal alone is not used as a shift conjugacy. □

3.2. **Comparison with a fixed longest border.** If a word starts and ends in p with a gap between the two copies, we compare it with pa^sp. The letter a must be chosen so that this new word has no additional copy of p. This restriction is necessary: for $p = 010$ and $s = 1$, the word $p1p = 0101010$ has a third copy of p, whereas $p0p = 0100010$ has only the two endpoint copies. The following lemma gives a suitable choice over every alphabet of size at least two.

Lemma 3.5 (Safe constant spacers)**.** *Let p be a nonempty word and let $s \geq 1$. At most one symbol $a \in \Sigma$ makes pa^sp contain an occurrence of p other than its two endpoint occurrences.*

Proof. Put $m = |p|$. For $0 \leq k \leq m$, let $F(k) \in \mathbb{N}^\Sigma$ be the vector of letter counts in $p_{[0,k)}$. Include zero as an empty border length for this proof. If $x < y$ are border lengths of p, the length-y prefix has a suffix equal to the length-x prefix of p. Splitting that length-y prefix into its first $y - x$ letters and last x letters gives

$$F(y) - F(x) = F(y - x). \tag{2}$$

Consider positive proper border lengths b, c with $b + c = m - s$. The intervening block $p_{[b,b+s)}$ has count vector $F(m) - F(b) - F(c)$, because the final c letters have count vector $F(c)$. This vector is independent of the choice of (b, c). Indeed, if $b < b'$ and $b + c = b' + c'$, put $d = b' - b = c - c'$. Equation (2) gives

$$F(b') - F(b) = F(d) = F(c) - F(c').$$

Thus these intervening blocks cannot be constant words in two different symbols: the vectors se_a, $a \in \Sigma$, are distinct, where e_a is the unit vector at coordinate a.

Suppose first that p is nonconstant, and number starts in pa^sp from zero. An extra occurrence has a start j with $0 < j < m + s$. If it meets the left endpoint copy but not the right, then $0 < j < m$, $j \leq s$, and $p = p_{[j,m)}a^j$. Its last j letters and this equality, iterated backwards through the word, force $p = a^m$. If it meets only the right endpoint, then $p = a^k p_{[0,m-k)}$ for some $0 < k < m$, with the

same consequence by forward iteration. An occurrence entirely in the spacer also forces $p = a^m$. None of these cases is possible.

An extra occurrence must therefore meet both endpoint copies, so $s < j < m$. Put $b = m - j$ and $c = j - s$. Matching its parts in the left copy, the spacer, and the right copy shows that b, c are positive proper border lengths, $b + c = m - s$, and $p_{[b,b+s)} = a^s$. The count-vector conclusion above allows at most one such symbol a.

Finally, if $p = c^m$ is constant, every $a \neq c$ separates its two length-m runs. The only occurrences of p are then the endpoint copies. This also proves the claimed bound in the constant case. □

Lemma 3.6 (Fixed longest border)**.** *Fix n and a word p of length $m < n$. Any two length-n words whose actual longest proper border is p are connected by at most two compatible swaps. If $n \leq 2m$, there is at most one such word. If $n > 2m$, a common intermediate word is*

$$z = pa^{n-2m}p,$$

where a is any choice from Lemma 3.5 having only the two endpoint occurrences of p. The resulting map has the two-sided cylinder property for the two original words.

Proof. When $n \leq 2m$, the prescribed prefix and suffix cover the word, so there is at most one possibility. Suppose $n > 2m$, choose a as stated, and let w have actual longest proper border p.

All overlaps of lengths at most m among w, z are exactly $B(p)$: the relevant prefixes and suffixes lie in their common endpoint copies of p. There is no proper self-overlap of w longer than m, by hypothesis. If an overlap of length k, $m < k < n$, goes from w to z, the last m letters of that overlap place p in z starting at $k - m$. This is an internal occurrence, since $0 < k - m < n - m$. If the overlap goes from z to w, its first m letters place p in z starting at $n - k$, again strictly between the endpoint starts. The same argument excludes a self-overlap of z longer than m.

Hence the four proper overlap sets for w, z all equal $B(p)$. In particular, z itself has longest proper border p, and w, z are compatible whenever they are distinct. Apply this argument to each of the two given words and use at most one swap from each to z. Each swap has the cylinder property by Lemma 3.2, and composition preserves it. An equality of words requires no swap. □

3.3. **The stronger induction.** The longest-border step uses the following elementary nesting fact. If p is a border word of w, with $m = |p|$, then

$$B(p) = B(w) \cap \{1, \ldots, m\}. \tag{3}$$

For $b \leq m$, the length-b prefix and suffix of w are precisely the length-b prefix and suffix of p, which proves the identity. In particular, the shortest nonempty border of a word is unbordered.

Proposition 3.7 (One recursive step)**.** *Let $u, v \in \Sigma^n$ have the same border set and a nonempty proper border. Let p, p' be their actual longest proper borders, of common length m. If an automorphism H has the two-sided cylinder property for (p, p'), it can be followed by at most two compatible length-n swaps to obtain an automorphism with that property for (u, v). No additional swap is needed when $n \leq 2m$.*

Proof. The cylinder property gives $H([u]) = [z]$ for one word z of length n beginning and ending in p'. By Lemma 3.2, applied to the cylinder identity for H, we have $B(z) = B(u) = B(v)$. Thus p' is the actual longest proper border of z as well as of v. Inheritance in Lemma 3.3 gives the cylinder property for (u, z). Lemma 3.6 supplies the asserted swaps from z to v, or gives $z = v$ when $n \leq 2m$. Composition proves the result, including its inverse assertion. □

Theorem 3.8 (Recursion along the border lengths)**.** *Suppose*

$$B(u) = B(v) = \{b_1 < b_2 < \cdots < b_k = n\},$$

and let $u_i = u_{[0,b_i)}$ and $v_i = v_{[0,b_i)}$. Fix an index $1 \le r \le k$. If there is an automorphism H_r with the two-sided cylinder property for (u_r, v_r), then there is an automorphism H_k with that property for (u, v). It is obtained by following H_r with at most

$$2\,\#\{i : r < i \le k,\ b_i > 2b_{i-1}\}$$

compatible swaps, each of length b_i at its corresponding stage. In particular, H_k preserves the specified occurrence positions and conjugates $X_q(u)$ to $X_q(v)$.

Proof. We prove, successively for $i = r, \dots, k$, the stronger assertion that H_i has the two-sided cylinder property for (u_i, v_i). The assertion at $i = r$ is the hypothesis. By (3),

$$B(u_i) = B(v_i) = \{b_1, \dots, b_i\}.$$

For $i > r$, their actual longest proper borders are therefore u_{i-1}, v_{i-1}. Apply Proposition 3.7 to H_{i-1}. It gives H_i with the required property, using no new swap when $b_i \le 2b_{i-1}$ and at most two otherwise. This proves both the induction and the stated count. At $i = k$, the identity $H_k([u]) = [v]$ and shift commutation transport occurrences at every position and give the restriction to the two avoiding shifts. □

There are two useful starting points for this induction. Chains of nonoverlapping swaps and alphabet permutations provide the first; maps for words with exactly one proper border provide the second. The latter allows a matrix construction to supply a base case without requiring a swap chain between the shortest border words.

Corollary 3.9 (The two base-case reductions)**.** *Let u, v have the common border set in Theorem 3.8. If their shortest nonempty border words u_1, v_1 are connected by a finite chain of nonoverlapping swaps and alphabet permutations, then an ambient automorphism has the two-sided cylinder property for (u, v).*

If $k \ge 2$, it is instead enough to construct an automorphism with that property for (u_2, v_2). Consequently, for any fixed alphabet, the assertion that every pair of words with border set $\{m, n\}$, $m < n$, admits such a map implies the same assertion for every pair of bordered words with equal border sets.

Proof. Each swap in the first chain has the property by Lemma 3.2; alphabet permutations and composition are covered by Lemma 3.3. Apply Theorem 3.8 with $r = 1$. For the second assertion, (3) gives $B(u_2) = B(v_2) = \{b_1, b_2\}$; apply the same theorem with $r = 2$. □

The words needed in the second base case reduce to two copies of an unbordered word separated by a constant word. The next lemma identifies the finite-word assertion required of a construction for these words.

Lemma 3.10 (Reduction of the one-proper-border case)**.** *Suppose $B(w) = \{m, n\}$, where $m < n$, and let p be its border word of length m. Then p is unbordered and $n = 2m + s$ for some $s \ge 0$. If $s = 0$, then $w = pp$. If $s > 0$, choose any $a \in \Sigma$ when $m \ge 2$, and choose $a \ne p_0$ when $m = 1$. At most one compatible swap gives a map with the two-sided cylinder property from w to $pa^s p$.*

Consequently, a map with that property between the corresponding constant-spacer words for two words with border set $\{m, n\}$ gives such a map between the original words, using at most two additional compatible swaps.

Proof. The word p is unbordered by (3). Its two endpoint copies in w cannot overlap: if $m < n < 2m$, their overlap would give p a proper border of length $2m - n$. Therefore $n = 2m + s$ for $s \ge 0$, and $s = 0$ gives $w = pp$.

Suppose $s > 0$. For $m \ge 2$, the unbordered word p is nonconstant and has no positive proper borders. None of the extra occurrences in the proof of Lemma 3.5 can therefore arise, for any choice of a. For $m = 1$, choosing $a \ne p_0$ leaves exactly the two endpoint occurrences. Lemma 3.6

now compares w with pa^sp by one compatible swap or the identity. The swap has the required cylinder property. For two original words, compose the first comparison, the given map of constant-spacer words, and the inverse of the second comparison. Lemma 3.3 proves the property for this composition. □

3.4. **The unbordered starting point.** For the first border level we use the following combinatorial result of Chandgotia, Marcus, Richey and Wu. Its conclusion supplies the actual word operations needed by the cylinder induction.

Proposition 3.11 (Unbordered chains [2, Theorems 4.3 and 4.9])**.** *For $q \geq 3$, any two unbordered words of the same length over Σ are connected by a finite sequence of alphabet permutations and nonoverlapping swaps. For $q = 2$ and length $n \geq 4$, the same statement holds among the unbordered words outside E_n, defined in Proposition 2.2.*

For length one an alphabet permutation suffices directly. Every operation in the proposition has the two-sided cylinder property, by Lemmas 3.2 and 3.3. Thus for $q \geq 3$, Theorem 3.8 constructs the desired map for every equal-border pair: begin with the shortest border words and proceed through the border lengths. The binary exceptions prevent this starting point from treating every binary pair. The next section constructs maps for words with exactly one proper border, which permits the induction to start at the second border length instead.

4. Binary words with one proper border

The binary construction needs a map between words whose shortest border words belong to different nonoverlapping-swap classes. We construct such a map for

$$p = 0^{m-1}1, \qquad p' = 0^{m-2}11, \qquad W = p0^sp, \qquad W' = p'0^sp', \qquad m \geq 3,\ s \geq 0.$$

The matrix calculation alone gives a conjugacy between the avoiding shifts. To use it in the border recursion, we retain the completed-word state of each prefix graph and mark its incoming edge. This also determines the image of every finite word beginning and ending in W, including words whose endpoint copies overlap.

4.1. **Factor reversals and finite marked paths.** For a nonnegative integer matrix A, distinguish all parallel edges in the directed graph with adjacency matrix A. Its edge shift consists of bi-infinite paths in this graph. Suppose $A = RS$ and $B = SR$, where R is a nonnegative integer $k \times l$ matrix and S is a nonnegative integer $l \times k$ matrix. View R as edges from the k original states to l intermediate states, and S as edges back. Each entry of RS counts pairs of consecutive factor edges. Choose bijections between the edges of A and these pairs, and likewise between edges of B and the pairs in the opposite order. Writing an A-edge as $e_t = (r_t, s_t)$ gives the edge conjugacy

$$F(e)_t = (s_t, r_{t+1}), \qquad F^{-1}(f)_t = (\widehat{r}_{t-1}, \widehat{s}_t) \quad \text{when } f_t = (\widehat{s}_t, \widehat{r}_t). \tag{4}$$

The two compositions recover each original pair of half-edges at its original coordinate. These formulas prove continuity, shift commutation, and invertibility without any assumption about edge multiplicities. We call $RS \longmapsto SR$ an elementary factor reversal.

Write E_{ij} for the matrix with a single 1 in position (i, j) and zeros elsewhere, and set $U_{ij} = I + E_{ij}$ for $i \neq j$. We use the following two reversals:

move	new matrix	nonnegative complementary factor
$\mathrm{col}(i,j)$	$U_{ij}AU_{ij}^{-1}$	AU_{ij}^{-1}
$\mathrm{row}(i,j)$	$U_{ij}^{-1}AU_{ij}$	$U_{ij}^{-1}A$.

The column move is allowed when column j dominates column i; the row move is allowed when row i dominates row j. Indeed, $U_{ij}^{-1} = I - E_{ij}$, and these inequalities are exactly the entrywise nonnegativity of the displayed factors.

Lemma 4.1 (Finite paths with marked endpoints)**.** *Suppose a sequence of these row and column moves takes A to A'. Fix states ℓ, β, and assume every move has*

$$i \neq j, \qquad j \neq \ell, \qquad i \neq \beta.$$

The same moves take $A + zE_{\ell\beta}$ to $A' + zE_{\ell\beta}$ with nonnegative polynomial factors. Regard the unique z-edge as marked. The resulting edge conjugacy moves every mark κ coordinates earlier, where κ is the number of column moves.

Every finite path whose first and last edges are marked determines the entire image path between the corresponding shifted marks, independently of either exterior. This correspondence is a bijection, including singleton paths and paths containing additional marks.

Proof. Put $D = E_{\ell\beta}$ and $U = U_{ij}$. The index assumptions give

$$U^{\pm 1} D = D = D U^{\pm 1}.$$

For a column move put $T = AU^{-1} + zD$; the reversal is $TU \mapsto UT$. For a row move put $T = U^{-1}A + zD$; the reversal is $UT \mapsto TU$. Choose the edge–factor-pair bijections to preserve the degree in z. The only U-edge entering ℓ is its diagonal edge ϵ_ℓ, and the only U-edge leaving β is ϵ_β. Denote the marked T-edge by δ.

In the column reversal the old marked edge is (δ, ϵ_β) and the new one is (ϵ_ℓ, δ). Formula (4) therefore takes a mark at time t to a mark at time $t-1$. For a path with marked endpoints at times a, b, the output edge at $a-1$ is forced to be the new marked edge; its otherwise missing U-half is ϵ_ℓ. All the remaining output edges through time $b-1$ use halves of the given path. In the inverse direction, the missing half at the right endpoint is forced to be ϵ_β.

In the row reversal the old mark is (ϵ_ℓ, δ) and the new mark is (δ, ϵ_β). A mark remains at time t. For marked endpoints at a, b, the output through time $b-1$ uses only the given edges, and the output at b is the new mark because the missing right half is ϵ_β. For the inverse, the missing left half is ϵ_ℓ.

More explicitly, let $e_0, \dots, e_h$ be a finite path with marked endpoints, let f_* be the new marked edge, and let $\phi(e, e')$ be the forward regrouping of two consecutive edges. Its finite image is

$$\begin{cases} (f_*, \phi(e_0, e_1), \dots, \phi(e_{h-1}, e_h)), & \text{column move,} \\ (\phi(e_0, e_1), \dots, \phi(e_{h-1}, e_h), f_*), & \text{row move.} \end{cases}$$

The inverse regroups consecutive output edges and appends the old mark in the column case, or prepends it in the row case. The two half-edge formulas cancel on every interior edge, and the forced endpoint halves give cancellation there as well. When $h = 0$, both maps simply send the single mark to the single mark. No step excludes internal marks. Composition gives the stated finite-path bijection and adds the displacements of the column moves. □

4.2. **Explicit factors for the spacer words.** For a binary word w of length n, its proper-prefix matrix A_w has states $0, \dots, n-1$. From state i, reading a symbol a goes to the largest $j < n$ for which $w_{[0,j)}$ is a suffix of $w_{[0,i)}a$; if this reading completes w, the edge is omitted. Thus A_w counts the remaining symbol-labelled transitions. Let e_i denote a unit row.

Lemma 4.2 (The spacer matrices)**.** *For $W = p0^s p$ and $W' = p'0^s p'$ as above, with $n = 2m + s$, the matrices $A_W, A_{W'}$ are joined by $4s + 7$ allowed row and column moves. There are $2s + 4$ column moves. In every move (i, j),*

$$i \notin \{0, m+1\}, \qquad j \neq n-1. \tag{5}$$

Proof. Put $b = m-2$, $c = m-1$, and $d = m$. First take $s = 0$, and set $f = 2m-1$, $g = f-1$. The seven moves are

$$\begin{aligned} &\mathrm{col}(f, 0), \quad \mathrm{col}(d, c), \quad \mathrm{col}(c, 0), \quad \mathrm{col}(d, 0), \\ &\mathrm{row}(d, c), \quad \mathrm{row}(c, b), \quad \mathrm{row}(f, b). \end{aligned} \tag{6}$$

Initially the rows are $e_0 + e_{i+1}$ except $A_c = e_c + e_d$ and $A_f = e_c$. Column f has its only nonzero entry at g, whose row contains e_0. The first column subtraction is therefore allowed. Column d then has its only entry at c, where column c is also one, allowing the second subtraction. These two moves give

$$\begin{aligned} &A_b = e_0 + e_c, \quad A_c = e_d, && A_d = e_0 + e_d + e_{d+1}, \\ &A_g = e_f, && A_f = e_0 + e_1 + e_c. \end{aligned}$$

Unlisted rows still equal $e_0 + e_{i+1}$. Column c is supported at b, f, both with an e_0 term. After $\mathrm{col}(c, 0)$, column d is supported at c, d, again with e_0 terms. After $\mathrm{col}(d, 0)$ the rows at c, d are $e_1 + e_d$ and $e_0 + e_1 + e_d + e_{d+1}$, so the next row subtraction is allowed. These three moves leave

$$A_b = e_c, \quad A_c = e_1 + e_c + e_d, \quad A_d = e_0 + e_{d+1}, \quad A_g = e_f, \quad A_f = e_1 + e_c.$$

The move $\mathrm{row}(c, b)$ gives $A_b = e_b + e_c$, $A_c = e_1 + e_d$, and $A_f = e_1 + e_b + e_c$. Subtracting row b from row f is now allowed. Since column f has only its entry at g, the final exceptional rows are

$$A_b = e_b + e_c, \quad A_c = e_1 + e_d, \quad A_g = e_b + e_f, \quad A_f = e_1.$$

They are the prefix rows for $(p')^2$. The default rows have not changed. All sums here count multiplicity; at $m = 3$, $b = 1$ and the penultimate row at f contains two copies of e_1.

Now suppose $s \geq 1$, and put $h = 2m - 1$, $f = h + s = n - 1$. Use

$$\begin{aligned} &\mathrm{col}(t, d) \quad (t = f, f-1, \ldots, h+1), \qquad \mathrm{col}(h, 0); \\ &\mathrm{col}(d, c), \quad \mathrm{col}(c, 0), \quad \mathrm{col}(d, 0), \quad \mathrm{row}(d, c), \quad \mathrm{row}(c, b); \\ &\mathrm{row}(t, d), \quad \mathrm{col}(t, 0), \quad \mathrm{row}(t-1, c) \quad (t = h+1, \ldots, f-1); \\ &\mathrm{row}(f, b), \quad \mathrm{col}(f, 0), \quad \mathrm{row}(f, d), \quad \mathrm{row}(f-1, c). \end{aligned} \tag{7}$$

Initially the rows are $e_0 + e_{i+1}$ except

$$A_c = e_c + e_d, \qquad A_i = e_d + e_{i+1} \ (h \leq i < f), \qquad A_f = e_c.$$

These formulas follow by reading the two possible next symbols: after at least $m - 1$ zeros following the first copy of p, reading 1 returns to state d, unless it completes W. In the descending group, column t is supported only at $t - 1$, whose row still contains e_d, so each subtraction is allowed. The resulting rows are

$$A_h = e_{h+1}, \qquad A_i = e_0 + e_{d+1} + e_{i+1} \ (h < i < f), \qquad A_f = e_0 + e_{d+1} + e_c.$$

Column h has only its entry at $h - 1$, whose row contains e_0. Thus $\mathrm{col}(h, 0)$ gives $A_{h-1} = e_h$ and $A_h = e_0 + e_1 + e_{h+1}$.

For the next five moves, column d is first supported only at c, where column c is one. After $\mathrm{col}(d, c)$, column c is supported at b, f, both with e_0 terms. After $\mathrm{col}(c, 0)$, column d is supported at c, d, again with e_0 terms. The next two row subtractions are by $e_1 + e_d$ and e_c, as in the square calculation. At the end of this preparation,

$$\begin{aligned} A_b &= e_b + e_c, & A_c &= e_1 + e_d, & A_d &= e_0 + e_{d+1}, \\ A_{h-1} &= e_h, & A_h &= e_0 + e_1 + e_{h+1}, & & \\ A_i &= e_0 + e_{d+1} + e_{i+1} \quad (h < i < f), & A_f &= e_b + e_c + e_{d+1}. & & \end{aligned}$$

Every unlisted row still has its default value.

Before the propagation triple with index t, the three relevant rows are

$$A_{t-2} = e_{t-1}, \qquad A_{t-1} = e_0 + e_1 + e_t, \qquad A_t = e_0 + e_{d+1} + e_{t+1}.$$

Subtracting row $d = e_0 + e_{d+1}$ from row t is allowed. Column t has only its entry at $t - 1$, so that row acquires e_d. The next column move changes these two rows to $e_1 + e_d + e_t$ and $e_0 + e_1 + e_{t+1}$.

Subtracting row $c = e_1 + e_d$ from row $t-1$ is allowed; its column is supported only at $t-2$. The three rows become

$$A_{t-2} = e_c + e_{t-1}, \qquad A_{t-1} = e_t, \qquad A_t = e_0 + e_1 + e_{t+1}.$$

This proves the induction through all propagation triples, as well as the required dominations.

The last three rows before the terminal moves are e_{f-1}, $e_0 + e_1 + e_f$, and $e_b + e_c + e_{d+1}$. The last two change as follows:

	A_{f-1}	A_f
$\mathrm{row}(f, b)$	$e_0 + e_1 + e_b + e_f$	e_{d+1}
$\mathrm{col}(f, 0)$	$e_1 + e_b + e_f$	$e_0 + e_1 + e_{d+1}$
$\mathrm{row}(f, d)$	$e_1 + e_b + e_d + e_f$	e_1
$\mathrm{row}(f-1, c)$	$e_b + e_f$	e_1.

All the required differences are nonnegative. Columns f and $f-1$ are supported at their respective predecessors, so the last move also gives $A_{f-2} = e_c + e_{f-1}$. This argument includes $s = 1$, when the propagation range is empty. The final exceptional rows are

$$A_b = e_b + e_c, \qquad A_c = e_1 + e_d,$$
$$A_i = e_c + e_{i+1} \quad (h-1 \le i \le f-2), \qquad A_{f-1} = e_b + e_f, \qquad A_f = e_1.$$

Together with the default rows, these are the prefix transitions for $0^{m-2}11\,0^{m+s-2}11 = W'$. Again the formulas retain multiplicity when $b = 1$.

The list has $s + 6 + 3(s-1) + 4 = 4s + 7$ moves, of which $2s + 4$ are column moves. The square list has the same counts at $s = 0$. Every row index i used in either list belongs to

$$\{m-1, m\} \cup \{2m-1, \dots, n-1\},$$

which excludes $0, m+1$ for $m \ge 3$. No second index j is $n-1$. This proves (5). □

4.3. **The ambient map and its cylinders.** The completed prefix graph of a word w has states $0, \dots, n$. Reading a symbol a from state i goes to the longest prefix of w that is a suffix of $w_{[0,i)}a$, now allowing length n. Induction on the number of input symbols shows that the current state is the longest prefix of w that is a suffix of the initial prefix followed by those symbols. After n input symbols it therefore depends only on those symbols, since every candidate suffix has length at most n. Thus every two-sided binary sequence labels exactly one two-sided path, and the path is recovered by a finite window. The label map is a conjugacy from this edge shift onto the full binary shift.

For a nonconstant w, the only edge completing w is the edge $n-1 \longrightarrow n$. Indeed, a completion from state n would give w period one, and a shorter state has insufficient length. For our words W, W', the only proper border length is m. Both bases are unbordered: the nonempty proper prefixes of p end in 0; those of p' also end in 0, except for $0^{m-2}1$, which is not a suffix because p' ends in 11. Both bases are nonconstant, so each displayed word has only its two endpoint occurrences of its base: an extra occurrence would overlap an endpoint or lie in a constant spacer. A shorter border would border the base, and a longer proper border would give a third occurrence. In particular, after reading the completed word the possible next states are 0 and $m+1$: the next symbol 0 continues from its longest proper border, while 1 returns to the empty prefix. Therefore the marked completed matrix is

$$\widehat{A}_w(z) = \begin{pmatrix} A_w & 0 \\ a & 0 \end{pmatrix} + zE_{n-1,n}, \qquad a = e_0 + e_{m+1}, \qquad w \in \{W, W'\}. \tag{8}$$

Here a is a row of length n. Setting $z = 1$ gives the full graph; the z-edge records precisely the completion times of w. Paths avoiding that edge correspond exactly to points avoiding w.

Lemma 4.3 (The ambient spacer bridge)**.** *For every $m \geq 3$ and $s \geq 0$, an automorphism of the full binary shift has the two-sided cylinder property for*

$$(0^{m-1}1\,0^s\,0^{m-1}1,\ 0^{m-2}11\,0^s\,0^{m-2}11).$$

It is obtained from the $4s+7$ factor reversals of Lemma 4.2, followed by a shift correction of $2s+4$ coordinates.

Proof. Extend each shear U from Lemma 4.2 by the identity on state n. Its row index i avoids the support of a, so $aU^{\pm 1} = a$. For a column move the complementary constant factor is

$$\begin{pmatrix} A & 0 \\ a & 0 \end{pmatrix} U^{-1} = \begin{pmatrix} AU^{-1} & 0 \\ a & 0 \end{pmatrix} \geq 0.$$

For a row move it is

$$U^{-1} \begin{pmatrix} A & 0 \\ a & 0 \end{pmatrix} = \begin{pmatrix} U^{-1}A & 0 \\ a & 0 \end{pmatrix} \geq 0.$$

The conjugated matrices have the same bottom row and zero last column. The second index of each shear avoids $n-1$, and its first index avoids n. Lemma 4.1 therefore applies to $D = E_{n-1,n}$. At the last step we obtain exactly $\widehat{A}_{W'}(z)$ in (8).

At $z = 1$, encode binary sequences in their completed prefix graphs, apply the edge conjugacies, and decode target labels. This defines a full-shift automorphism K. There are $\kappa = 2s+4$ column moves, so K moves every completion time κ coordinates earlier. With the convention $(\sigma x)_i = x_{i+1}$, put

$$H = \sigma^{-\kappa} K.$$

Completions, and hence occurrence starts, now have their original coordinates in both directions. In particular $H([W]) = [W']$.

Let w be a word of length $N \geq n$ beginning and ending in W. Its designated completion edges occur at times $n-1, N-1$. The first edge is the unique marked edge. All the following edges through the second mark are determined by the remaining symbols of w, independently of the past. Lemma 4.1 determines the entire output path between the corresponding marks. After the shift correction, this determines the output labels on $[n-1, N)$. The first output completion forces W' on $[0, n)$, determining the remaining initial symbols. The last completion forces the final copy of W'. Thus one word v of length N, beginning and ending in W', satisfies $H([w]) \subseteq [v]$.

Apply the same finite-path argument to the inverse and v. It gives $H^{-1}([v]) \subseteq [\widetilde{w}]$ for one length-N word $\widetilde{w}$ beginning and ending in W. Any extension $x \in [w]$ also lies in $H^{-1}([v])$, so $[w] \cap [\widetilde{w}] \neq \varnothing$. Because the two cylinders specify the same coordinates and length, $\widetilde{w} = w$. The inclusions are therefore equalities. Starting with an arbitrary word beginning and ending in W' gives the inverse assertion. For $N = n$ the finite path is a singleton marked edge; for $n < N < 2n$ the endpoint words overlap. Both cases are included in the argument. □

4.4. **The binary one-border case.** Proposition 3.11 joins the nonexceptional unbordered binary bases by nonoverlapping swaps and complementation. We first check what these operations do to the constant-spacer words.

Lemma 4.4 (Lifting a nonoverlapping base swap)**.** *Let p, p' be distinct unbordered binary words of length $m \geq 2$ having no proper cross-overlap in either direction. For any symbol a and any $s \geq 0$, their ambient swap has the two-sided cylinder property for $(pa^sp, p'a^sp')$.*

Proof. Both bases are nonconstant. An occurrence of either base inside pa^sp, apart from its endpoint copies of p, would overlap an endpoint copy or lie wholly in the constant spacer. The first is excluded by the overlap hypotheses and the second by nonconstancy. A swapped occurrence crossing the boundary cannot overlap the endpoint copy. Thus the restriction of the output to this interval is $p'a^sp'$. The inverse gives the reverse containment, so the swap carries $[pa^sp]$ onto $[p'a^sp']$. The cylinder property for the bases and inheritance give the asserted property for the longer words. □

Theorem 4.5 (One proper border over the binary alphabet)**.** *Let u, v be binary words with $B(u) = B(v) = \{m, n\}$, where $m < n$. There is an automorphism of the full binary shift with the two-sided cylinder property for (u, v). In particular it takes occurrences of u to occurrences of v at the same positions and conjugates $X_2(u)$ to $X_2(v)$.*

Proof. By Lemma 3.10, $n = 2m + s$ with $s \geq 0$, and each word can be compared with $pa^s p$, where p is its unbordered length-m border. For $s = 0$ there is no spacer; for $s > 0$, any symbol a is allowed when $m \geq 2$, and the symbol different from p_0 is required when $m = 1$. All comparisons have the two-sided cylinder property. When $m \geq 2$, Lemma 3.6 also compares the two choices of constant spacer for a fixed base p.

If $m = 1$, complementation identifies the two possible constant-spacer words, for $s = 0$ as well as for $s > 0$. If $m = 2$, the unbordered bases are $01, 10$; complementation identifies them, and a fixed-base swap adjusts the spacer when necessary. These cases are therefore settled.

Suppose $m \geq 3$. Lemma 4.4 lifts every nonoverlapping base step to the corresponding constant-spacer words. Complementation also lifts, possibly changing the spacer, which a fixed-base swap can adjust. For $m \geq 4$, the cited connectivity theorem consequently joins all constant-spacer words with nonexceptional bases. The base $0^{m-2}11$ is unbordered and nonexceptional. After adjusting the spacer to 0, Lemma 4.3 joins the exceptional base $0^{m-1}1$ to this class. Complementation attaches $1^{m-1}0$. Conjugating the bridge by reversal attaches 10^{m-1} to the nonexceptional base 110^{m-2}; complementation attaches 01^{m-1}. Reversal is used only in this conjugated form, which is an automorphism by Lemma 3.4.

For $m = 3$, the four unbordered bases are $001, 011, 100, 110$. The marked bridge joins the words with bases $001, 011$, and complementation joins 001 to 110 and 011 to 100, with spacer adjustment when needed. Thus these four bases also give a single class of constant-spacer words. Composition with the initial and final compatible swaps proves the theorem, including the inverse cylinder assertion. □

The theorem supplies the second-border starting point in Corollary 3.9. The recursion then transports the same two-sided cylinder property through every longer border. Thus no additional matrix family is needed for the bordered binary case.

5. Completion of the classification

We now have both starting points for the border induction. Periodic points give the converse, while the exceptional unbordered binary words require a separate invariant.

Proof of Theorem 1.1. A conjugacy preserves each set of points fixed by σ^r. Proposition 2.1 proves that equality of these counts is equivalent to equality of border sets, including equality of the forbidden-word lengths.

It remains to prove sufficiency and the ambient assertion. Suppose $B(u) = B(v) = \{b_1 < \cdots < b_k = n\}$. For $q \geq 3$, the shortest border words $u_{[0,b_1)}$ and $v_{[0,b_1)}$ are unbordered. Proposition 3.11 and Theorem 3.8, starting at $r = 1$, give a full-shift automorphism with the two-sided cylinder property for (u, v).

Let $q = 2$. If $k = 1$ and $n = 1$, use complementation if necessary. If $k = 1$ and $n \geq 2$, irreducibility excludes E_n. No such unbordered words exist at lengths two and three; for $n \geq 4$, Proposition 3.11 supplies the required map. If $k \geq 2$, the words $u_{[0,b_2)}$ and $v_{[0,b_2)}$ have border set $\{b_1, b_2\}$. Theorem 4.5 supplies their map, and Theorem 3.8, starting at $r = 2$, extends it to (u, v).

In every case the resulting automorphism H satisfies $H([u]) = [v]$. Commuting with the shift transports this identity to every coordinate, so H preserves the specified occurrence positions and restricts to a conjugacy $X_q(u) \to X_q(v)$. □

The reducible binary words all have the same border set at a fixed length, and hence the same periodic counts. We distinguish their two orientations by the behavior of points approaching a fixed point under forward iteration.

Proposition 5.1. *For each $n \geq 3$, the reducible binary shifts with forbidden-word length n form exactly two conjugacy classes, represented by the following pairs of forbidden words:*

$$\{0^{n-1}1,\, 1^{n-1}0\}, \qquad \{10^{n-1},\, 01^{n-1}\}.$$

For $n = 2$, the pairs coincide and give a single class.

Proof. Complementing symbols conjugates the two shifts within each pair. For $n \geq 3$, put

$$X^+ = X_2(0^{n-1}1), \qquad X^- = X_2(10^{n-1}), \qquad Y = X_2(0^{n-1}), \qquad z = 0^{\mathbb{Z}}.$$

Let $\mathrm{Per}(X) = \bigcup_{r\geq 1}\{x : \sigma^r x = x\}$. A periodic point using both symbols and containing $n-1$ consecutive zeros contains both forbidden orientations, at the two ends of a maximal zero run. Therefore

$$\mathrm{Per}(X^+) = \mathrm{Per}(X^-) = \mathrm{Per}(Y) \cup \{z\}.$$

Requiring $n-1$ consecutive zeros at specified coordinates isolates z in this periodic-point set. The other fixed point, $1^{\mathbb{Z}}$, is not isolated there: periodic points with isolated zeros can agree with it on arbitrarily long central intervals, and belong to Y because $n - 1 \geq 2$. Thus z is the unique fixed point isolated in the relative topology of $\mathrm{Per}(X^{\pm})$. This is an intrinsic property, so any conjugacy $X^+ \to X^-$ must send z to z.

For a fixed point z define

$$W^s_X(z) = \{x \in X : \sigma^r x \longrightarrow z \text{ as } r \longrightarrow +\infty\}.$$

A conjugacy carries this set onto the corresponding set at its image fixed point, by continuity of the conjugacy and its inverse. For $z = 0^{\mathbb{Z}}$, convergence means exactly that $x_i = 0$ for all sufficiently large i. The point with $x_i = 1$ for $i < 0$ and $x_i = 0$ for $i \geq 0$ belongs to $W^s_{X^+}(z) \setminus \{z\}$. In X^-, any eventually-zero point containing a 1 has a last 1, followed by at least $n-1$ zeros, which is forbidden. Hence

$$W^s_{X^+}(z) \neq \{z\}, \qquad W^s_{X^-}(z) = \{z\}.$$

The two classes cannot be conjugate. When $n = 2$, the only exceptional words are 01 and 10, which complementation interchanges. □

Corollary 5.2. *Among binary shifts obtained by forbidding an unbordered word of length n, the numbers of conjugacy classes are*

n	1	2	3	$n \geq 4$
number of classes	1	1	2	3.

Classes belonging to different lengths cannot coincide.

Proof. At length one, the two shifts each consist of one point. At length two the unbordered words are $01, 10$; at length three they are $001, 110, 100, 011$, all exceptional. For $n \geq 4$ the two reducible classes remain, and there is a nonempty irreducible class: the word $0^{n-2}11$ is unbordered and nonexceptional. Indeed, its only occurrence of 11 is terminal, so no proper prefix of length at least two can be a suffix, and its endpoint letters differ. All nonexceptional unbordered words of a fixed binary length define conjugate shifts by [2, Theorem 4.3]. Irreducibility is preserved by conjugacy: its word criterion is equivalent to requiring, for any two nonempty open sets U, V, an $r \geq 0$ with $\sigma^r(U) \cap V \neq \varnothing$. This separates that class from the other two. Proposition 2.1 separates different lengths. □

6. Finite graph constructions

The local maps constructed above can also be expressed by finite sequences of graph splittings and amalgamations. We give the presentations and the partitions that specify these moves. For a directed graph G, write $E(G)$ for its edge set and X_G for the shift of bi-infinite edge paths in G.

For $w = w_0 \cdots w_{n-1}$ over Σ, define $F(w)$ to have vertices $0, \ldots, n-1$. From vertex i, the edge labelled $a \in \Sigma$ ends at

$$\max\{0 \le j < n : w_{[0,j)} \text{ is a suffix of } w_{[0,i)}a\}.$$

The prefix of length zero is empty. Edges with different labels remain distinct even when their endpoints agree. Mark the edge δ_w leaving $n-1$ with label w_{n-1}.

The labels give a conjugacy from the edge shift of $F(w)$ to $\Sigma^{\mathbb{Z}}$. Indeed, immediately before reading x_t, the vertex is the longest proper prefix of w that is a suffix of the past $x_{[t-n+1,t)}$. Iterating the transition rule shows that every path has these vertices, so every label sequence has exactly one lift. The edge at time t is δ_w exactly when $x_{[t-n+1,t+1)} = w$. Consequently $G(w) = F(w) \setminus \{\delta_w\}$ presents $X_q(w)$ by a conjugacy. Every edge lies on a bi-infinite path: vertex zero has a loop labelled by any symbol different from w_0, every vertex is reached from zero by its prefix, and every vertex has at least $q-1$ outgoing edges.

An *in-splitting* partitions the edges entering a vertex, replaces that vertex by one copy for each class, and assigns each incoming edge to its class. Each copy receives a copy of every outgoing edge. For loops, the outgoing copy determines the source and the incoming class determines the target. Forgetting the copy names is a conjugacy: the preceding edge determines the copy. An *out-splitting* instead partitions outgoing edges and duplicates incoming edges. The inverse of either operation is an *amalgamation*.

Lemma 6.1. *For each $a \in \Sigma$, the graph $F(wa)$ is obtained from $F(w)$ by one in-splitting. Split the target of δ_w according to the incoming classes $\{\delta_w\}$ and its complement. Name the first copy n and retain the old name for the other copy. The edge from n labelled a is δ_{wa}.*

Thus if p is a length-m prefix of a length-n word u, precisely $n-m$ specified in-splittings build $F(u)$ from $F(p)$. If p is also a suffix of u, occurrences of u are detected in $F(p)$ by a path of length $n-m+1$ whose first and last edges are δ_p.

Proof. The proper prefixes of wa are those of w, together with w itself. The latter matches the past exactly when the incoming edge is δ_w. This is precisely the information recorded by the new copy n, and proves all transitions, including those involving loops. The edge labelled a from that copy completes wa.

Both incoming classes are nonempty. Write b for the target of δ_w, the longest proper-border length of w. If $b > 0$, the edge from $b-1$ labelled w_{b-1} also enters b and differs from δ_w. If $b = 0$, a loop at zero with label different from w_0 belongs to the other class. This also covers $n = 1$.

Iterating proves the splitting count. For the last assertion, start with δ_p and then follow the labels $u_{[m,n)}$. The first edge certifies the initial copy of p, and the remaining labels certify the rest of u. The last edge is δ_p because u ends in p. The initial $m-1$ letters of the occurrence precede the displayed path. The same argument applies when the endpoint copies overlap. □

For example, in $F(01)$ the edge completing 01 enters vertex zero. Separating it from the other incoming edges creates vertex two of $F(010)$; the edge labelled zero from this vertex is the new marked edge. No edge is deleted; all lifts of the old completion edge are retained.

There is a useful distinction between extending a detector and restricting a graph map. If a splitting has collapse map $\pi : G' \to G$ and $D \subseteq E(G)$, then it restricts to a conjugacy

$$X_{G' \setminus \pi^{-1}(D)} \longrightarrow X_{G \setminus D}.$$

This follows because a path uses an edge of D exactly when its unique lift uses an edge of $\pi^{-1}(D)$. Every copy of a deleted edge must therefore be removed. By contrast, to forbid wa one deletes only

δ_{wa} after constructing $F(wa)$. Deleting δ_w first forbids the shorter word. For instance, $X_2(0)$ has one fixed point, whereas $X_2(01)$ has two, so these two operations cannot in general be interchanged.

Proposition 6.2. *Let $u, v \in \Sigma^n$. Suppose an automorphism H of $\Sigma^{\mathbb{Z}}$ satisfies $H([u]) = [v]$, where the cylinders are based at coordinate zero. Given local rules for H and H^{-1}, there is a finite, specified sequence of graph splittings, graph isomorphisms, and amalgamations realizing the restriction $H : X_q(u) \to X_q(v)$ on their prefix presentations.*

Proof. Since H commutes with the shift, it transports occurrences at every position. Its restriction is therefore a conjugacy of the avoiding shifts. Say a local rule has *memory* r and *anticipation* s if its output at zero reads only $x_{[-r,s+1)}$, where $r, s \geq 0$. If H has these parameters, the induced rule on prefix-edge symbols reads the block on $[-r-n+1, s+1)$: compute the current output letter and the preceding $n-1$ output letters, which determine its initial vertex. The inverse rule is obtained in the same way.

The classical splitting decomposition now applies to this restricted conjugacy; see [1, Theorem 2.12]. We specify its partitions here. A full splitting puts each incoming, or each outgoing, edge in a separate class. First fully out-split both prefix graphs. The resulting vertices are the old edges e, and the edges are pairs (e, f) that can be concatenated. Collapse sends (e, f) to e, so the vertex at time t records the original edge at time t, without a shift. Parallel old edges remain distinct vertices. Every state and transition occurs on a bi-infinite path, and splittings into nonempty classes preserve this property.

Let $R = r + n - 1$ and $S = s$ be the memory and anticipation of the induced edge rule. Apply S full outgoing splittings and R full incoming splittings to its source presentation. The new state records the old block on $[-R, S+1)$, so the forward rule becomes a one-state map f: each output symbol depends only on the current input symbol.

Suppose its inverse has positive anticipation N and memory M. Partition the outgoing edges $a \to a'$ of the source by $f(a')$, and those $b \to c$ of the target by c. The new source and target states are (a, c) and (b, c), and the new rule is

$$(a, c) \longmapsto (f(a), c).$$

On corresponding paths these states record (x_t, y_{t+1}) and (y_t, y_{t+1}). A target block on $[-M, N)$ supplies the old block $y_{[-M,N+1)}$, so the old inverse recovers x_0; the current pair supplies y_1. The new inverse therefore has anticipation at most $N-1$ and memory at most M.

Dually, partition incoming source edges $a' \to a$ by $f(a')$, and incoming target edges by their predecessor. The new rule sends (c, a) to $(c, f(a))$ and reduces positive inverse memory by one. Iterating leaves one-state rules in both directions. Since all vertices and transitions occur on bi-infinite paths, these rules give inverse graph isomorphisms. Follow the source splittings by this isomorphism and the inverses of the target splittings, including the initial edge-symbol conversion. The partition constructions commute with collapse, so the resulting map is the given restriction of H. □

References

[1] M.-P. Béal, J. Berstel, S. Eilers, and D. Perrin, *Symbolic dynamics*, 2011, arXiv:1006.1265v3.

[2] N. Chandgotia, B. Marcus, J. Richey, and C. Wu, *Shifts of finite type obtained by forbidding a single pattern*, 2024, arXiv:2409.09024v1.