Journal of Statistical Physics (2025) 192:28
https://doi.org/10.1007/s10955-025-03403-2

Hierarchical Structure of Periodic Orbits of a Hyperbolic Automorphism on the 2-Torus

Huynh M. Hien [1] · **Nguyen B. Tran** [1] · **Tran N. Nguyen** [1]

Received: 13 January 2024 / Accepted: 13 January 2025

Abstract
This paper studies the hierarchical structure of periodic orbits of the automorphism induced by the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ on the torus $\mathbb{T}^2$. The induced symbolic dynamics is not trivial with forbidden sequences. We show that the periodic orbits of the system is hierarchically structured by clusters. We establish the number of clusters via symbolic dynamics and digraphs. Algorithms that group all periodic orbits in clusters are given.

Keywords Hierarchical structure · Periodic orbits · Symbolic dynamics · Hyperbolic toral automorphisms

Mathematics Subject Classification primary 37B10 · 37D20; secondary 05C20 · 37C55

1 Introduction

Bohigas, Giannoni and Schmit [7] argued that periodic orbit clustering leads to universal spectral fluctuations for chaotic quantum systems. Altland et al. [1] stated that periodic orbits including encounters form orbit clusters. The clusters of periodic orbits have hierarchical structures due to the near indistinguishability of different orbits of a cluster within links, and one can get a bigger cluster when one orbit in a cluster closely encounters an orbit from another cluster. The simplest orbit cluster is a Sieber–Richter pair [17] where each orbit has 2 stretches mutually close, which is called a 2-antiparallel encounter.

Roughly speaking, a periodic orbit cluster is a family of periodic orbits which visit the same parts in the phase space with the same number of times but with different orders. There is a similar notion in graph theory, namely *degeneracy class*. The problem of counting degeneracy classes was studied in [6, 9, 16, 18] for some classes of metric graphs. The leading asymptotic (for large n) contribution to the number of degeneracy classes in a general connected graph

This paper is dedicated to Professor Dinh Thanh Duc on the occasion of his 65th birthday.

Communicated by Marco Lenci.

✉ Huynh M. Hien
huynhminhhien@qnu.edu.vn

[1] Department of Mathematics and Statistics, Quy Nhon University, 170 An Duong Vuong, Quy Nhon, Vietnam

was obtained by Berkolaiko [6] and it was applied to calculating the two-point spectral statistics for this class of graphs in the limit as the number of vertices tends to infinity. The number of degeneracy classes and the number of closed paths in each class (so-called *degeneracy*) were obtained by Tanner [18] for balanced, binary directed graphs (i.e., directed graphs with two incoming and two outgoing edges for each vertex). The spectral form factor can be written in terms of degenerate periodic orbit pairs only via the periodic orbit length degeneracy function. Gavish and Smilansky [9] considered an indirected, fully connected graph (i.e., a graph in which each vertex is connected by a single edge to any other vertices beside itself) with no loop and the lengths of edges being rationally independent and they obtained asymptotics for the average size of a degeneracy class. Sharp [16] established an asymptotic formula for the number of pairs of closed cycles with the same metric length for directed, connected graph with non-backtracking and the lengths of edges are rationally independent. Gutkin and Osipov [10] discovered the relation between clusters of periodic orbits and degeneracy classes of closed paths. The authors used the symbolic dynamics of the baker's map and introduced the notion of *p-closeness* to define clustering of periodic orbits and showed that periodic orbits create hierarchy of clustering. The counting of clusters is equivalent to the one of closed paths in the corresponding de Brujin graph. The symbolic dynamics of the baker's map is the simplest chaotic dynamical system with two-letter alphabet $x_i \in \{0, 1\}$ without forbidden sequences, i.e., each symbol in the sequence can be followed by any other symbols, and all sequences are periodic. The authors studied the distribution of cluster sizes for the baker's map in the asymptotic limit of long trajectories via counting degeneracies in the spectrum of the associated de Bruijn graphs and derived the probability $\mathcal{P}_k$ that k randomly chosen periodic orbits belong to the same cluster. The graphs considered are balanced and binary. In addition, they provided an asymptotic formula for the number of clusters as the length of sequences tends to infinity via a system of linear equations. Again Gutkin and Osipov [11] showed that the counting of cluster size can be turned to spectral problem for matrices from truncated unitary ensemble and they gave an asymptotic formula for the average size of clusters through the average number of encounters and a conjecture. However, up to now, there have been no researches considering dynamical systems whose symbolic dynamics are non-trivial. In the present paper, we consider a classical hyperbolic dynamical system, namely the diffeomorphism T on the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ induced by matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. More concretely, $T(x+\mathbb{Z}^2) = Ax + \mathbb{Z}^2$ for all $x + \mathbb{Z}^2 \in \mathbb{T}^2$. The 2-torus can be viewed as the unit square $[0, 1] \times [0, 1]$ with opposite sides identified: $(x_1, 0) \sim (x_1, 1)$ and $(0, x_2) \sim (1, x_2)$, $x_1, x_2 \in [0, 1]$. The map T is given in coordinates by

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} (2x_1 + x_2) \bmod 1 \\ (x_1 + x_2) \bmod 1 \end{pmatrix}.$$

Note that $\mathbb{T}^2$ is a commutative group and A is an automorphism on $\mathbb{T}^2$, since A^{-1} is also an integer matrix. A Markov partition for T was constructed by Katok and Hasselblatt [12], then Barreira [4] proposed the symbolic dynamics induced from this Markov partition. The adjacency matrix is a 5×5 matrix with entries 0 and 1, so the respective symbolic dynamics is with forbidden sequences, i.e., each symbol in a sequence is not allowed to follow by any other symbols. This is completely different from the one induced from the baker's map considered by Gutkin and Osipov in [10, 11].

The aim of this paper is to study the hierarchical property of clustering of periodic orbits of T via its respective symbolic dynamics and with the help of digraphs. We exploit the notion *p-closeness* between periodic sequences introduced by Gutkin and Osipov [10]. This is an equivalence relation and each equivalence class is called a p-cluster. There is a one-to-one

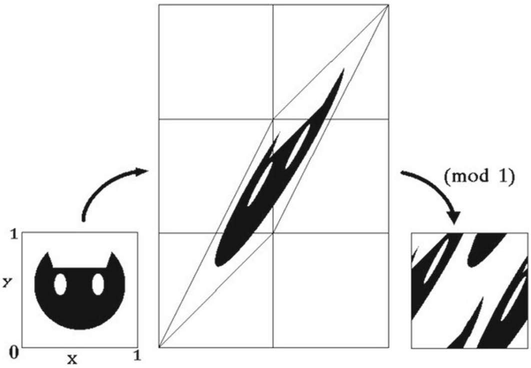

Fig. 1 Graphical representation of the Arnold's cat map [19]

correspondence between the set of n-periodic sequences in the shift space and the set of n-*periodic* points of the automorphism T. Each cluster of periodic sequences corresponds to a cluster of periodic orbit of T. Furthermore, if two periodic orbits are $p+1$-close, then they are p-close. So, the periodic orbits of the system can be represented as a hierarchy structure of clusters. We obtain an asymptotic formula for the number of p-clusters for given p. In addition, we give algorithms to list all periodic orbits and clusters of periodic orbits for the respective subshift of finite type.

Periodic sequences with the same period are grouped by p-clusters, so we know the accurate size of each p-cluster as well. The largest $1 \leq p \leq n$ for which at least one p-cluster contains more than one sequence is given. The argument in this paper can be applied for any dynamical system whose adjacency matrix is specific.

The paper is organized as follows. In the next section we give a brief construction of symbolic dynamics for T. A periodic point of T associates to a periodic sequence with symbols $0, 1, 2, 3, 4$. Section 3 investigates the clustering of periodic orbits of T. Periodic orbits are hierarchically structured by p-clusters. We show that the number of p-clusters is the one of degeneracy classes in the corresponding de Bruijn graph. An asymptotic formula for the number of p-clusters for sufficiently large period is established. In the last part of this paper, we present algorithms to list all periodic orbits and arrange them in p-clusters.

2 Symbolic Dynamics of *T*

The eigenvalues of matrix A are

$$\lambda = \frac{3-\sqrt{5}}{2} \quad \text{and} \quad \lambda^{-1} = \frac{3+\sqrt{5}}{2}, \tag{2.1}$$

which do not lie on the unit circle. So, the corresponding isomorphism T is hyperbolic [4, Example 6.1]. In addition, the fact that the set of periodic points of T is $\mathbb{Q}^2/\mathbb{Z}^2$ implies that set of its non-wandering points is the full space $\mathbb{T}^2$. It follows that T is an Axiom A diffeomorphism (see [8]). An explicit Markov partition of T is constructed by Katok and Hasselblatt [12, p. 84] including five rectangles R_0, R_1, R_2, R_3, R_4. Let

$$\mathscr{A} = \{0, 1, 2, 3, 4\}$$

and

$$\mathscr{A}^{\mathbb{Z}} = \{x = (x_i)_{i\in\mathbb{Z}} : x_i \in \mathscr{A} \quad \text{for all} \quad i \in \mathbb{Z}\}.$$

Definition 2.1 The map $\sigma : \mathscr{A}^{\mathbb{Z}} \longrightarrow \mathscr{A}^{\mathbb{Z}}$ defined by

$$(\sigma x)_i = x_{i+1} \quad \text{for all} \quad i \in \mathbb{Z}$$

is called the *shift map*.

A distance on $\mathscr{A}^{\mathbb{Z}}$ is defined by

$$\varrho(x, y) = \begin{cases} 2^{-k} & \text{if } x \neq y \text{ and } k \text{ is maximum so that } x_{-k} \dots x_0 \dots x_k = y_{-k} \dots y_0 \dots y_k \\ 0 & \text{if } x = y \end{cases}$$

for $x = (x_n)$, $y = (y_n) \in \mathscr{A}^{\mathbb{Z}}$. Then $(\mathscr{A}^{\mathbb{Z}}, \varrho)$ is a compact metric space and the shift map σ is a homeomorphism on $\mathscr{A}^{\mathbb{Z}}$. The adjacency matrix $\mathcal{A} = (a_{i,j})_{i,j=0}^4$ is defined by

$$a_{i,j} = \begin{cases} 1 & \text{if} \quad \text{int } T(R_i) \cap \text{int } R_j \neq \varnothing, \\ 0 & \text{if} \quad \text{int } T(R_i) \cap \text{int } R_j = \varnothing \end{cases}$$

and according to [4],

$$\mathcal{A} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

From the matrix $\mathcal{A}$, we define a set of allowed sequences as follows. The set

$$\Lambda_{\mathcal{A}} = \{(x_n) \in \mathscr{A}^{\mathbb{Z}} : a_{x_i, x_{i+1}} = 1 \ \text{ for all } i \in \mathbb{Z}\}$$

is a closed set of $\mathscr{A}^{\mathbb{Z}}$ and invariant under the shift map σ. The map $\sigma_{|\Lambda_{\mathcal{A}}} : \Lambda_{\mathcal{A}} \to \Lambda_{\mathcal{A}}$ is called the *subshift of finite type* induced by $\mathcal{A}$.

Definition 2.2 For a given $n \in \mathbb{N}$, a sequence $x = (x_i)_{i=-\infty}^{\infty} \in \mathscr{A}^{\mathbb{Z}}$ is called *periodic of period* n or n*-periodic* if $\sigma^n(x) = x$, i.e., $x_{i+n} = x_i$, for all $i \in \mathbb{Z}$. Then we write $x = [x_0 x_1 \dots x_{n-1}]$. The set consisting of all periodic sequences in $\Lambda_{\mathcal{A}}$ of period n and its cardinality are denoted by P_n and $p(n)$, respectively.

Let

$$X_n := \left\{x_0 \dots x_{n-1} : x_0, \dots, x_{n-1} \in \mathscr{A},\ a_{x_i, x_{i+1}} = 1,\ i \in \{0, \dots, n-2\}\right\}$$

be the set of subsequences of length n in $\Lambda_{\mathcal{A}}$. To calculate the cardinality of X_n and $p(n)$, we first observe that

$$\mathcal{A}^n = \begin{pmatrix} a_n & a_n & b_n & a_n & b_n \\ a_n & a_n & b_n & a_n & b_n \\ a_n & a_n & b_n & a_n & b_n \\ b_n & b_n & c_n & b_n & c_n \\ b_n & b_n & c_n & b_n & c_n \end{pmatrix},$$

where

$$a_n = \frac{5+\sqrt{5}}{10}\lambda^n + \frac{5-\sqrt{5}}{10}\lambda^{-n},\ b_n = -\frac{5+3\sqrt{5}}{10}\lambda^n + \frac{-5+3\sqrt{5}}{10}\lambda^{-n},$$
$$c_n = \frac{5+2\sqrt{5}}{5}\lambda^n + \frac{5-2\sqrt{5}}{5}\lambda^{-n}. \tag{2.2}$$

The number of subsequences of length n is the sum of all entries in $\mathcal{A}^{n-1}$, which is

$$\mathrm{card}(X_n) = 9a_{n-1} + 12b_{n-1} + 4c_{n-1} = \frac{5-2\sqrt{5}}{5}\lambda^n + \frac{5+2\sqrt{5}}{5}\lambda^{-n}. \tag{2.3}$$

The number of n-periodic sequences in $\Lambda_{\mathcal{A}}$ is

$$p(n) = \mathrm{tr}(\mathcal{A}^n) = \lambda^n + \lambda^{-n}.$$

Denote by

$$Q_n = \{x \in \Lambda_{\mathcal{A}} : \sigma^n(x) = x, \sigma^k(x) \neq x \text{ for } 1 \leq k < n\}$$

the set of primitive periodic points of period $n \in \mathbb{N}$ and by $q(n)$ its cardinality. It is obvious that $P_n = \bigcup_{k|n} Q_k$ and therefore $p(n) = \sum_{k|n} q(k)$. To calculate $q(k)$, we need the Möbius inversion formula:

$$q(k) = \sum_{m|k} \mu(m) p\left(\frac{k}{m}\right),$$

where μ is the Möbius function defined as follows. If $m = p_1^{s_1} \dots p_r^{s_r}$ is the prime factorisation, then

$$\mu(m) = \begin{cases} (-1)^r & \text{if } m \text{ is square-free,} \\ 1 & \text{if } m = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Recall that a positive integer is said to be square-free if no prime number occurs more than once in its prime factorisation; see [2, Theorem 2.9] for a proof of (2.4). We have shown the following result.

Proposition 2.3 *The number of periodic sequences with primitive period n is*

$$q(n) = \sum_{k|n} \mu\left(\frac{n}{k}\right)(\lambda^k + \lambda^{-k}).$$

We define a metric on $[0, 1)$ by

$$\rho(x, y) = \min\{|x-y|, |1-x+y|, |1+x-y|\}, \quad \text{for} \quad x, y \in [0, 1)$$

and consider metric d on $\mathbb{T}^2$ defined by

$$d(\mathsf{x}, \mathsf{y}) = \max\{\rho(x_1, y_1), \rho(x_2, y_2)\},$$

where $\mathsf{x} = (x_1, x_2) + \mathbb{Z}^2$, $\mathsf{y} = (y_1, y_2) + \mathbb{Z}^2$, $(x_1, x_2), (y_1, y_2) \in [0, 1) \times [0, 1)$. The next result will be useful later.

Lemma 2.4 *For $p \geq 2$ and $x = (x_i) \in \Lambda_{\mathcal{A}}$, the set*

$$Q_p(x) := \bigcap_{k=-p}^{p} T^{-k}(R_{x_k})$$

has diameter at most λ^p.

Proof Every rectangle $R \in \{R_0, \dots, R_4\}$ is formed by two directions, in which the shorter parallel sides go along the stable direction (the eigenvector w.r.t. λ^{-1}), whereas the longer parallel sides go along the unstable direction (the eigenvector w.r.t. λ). The width (resp. length) of R is contracting (resp. expanding) by the factor λ^{-1} (resp. λ) after the action

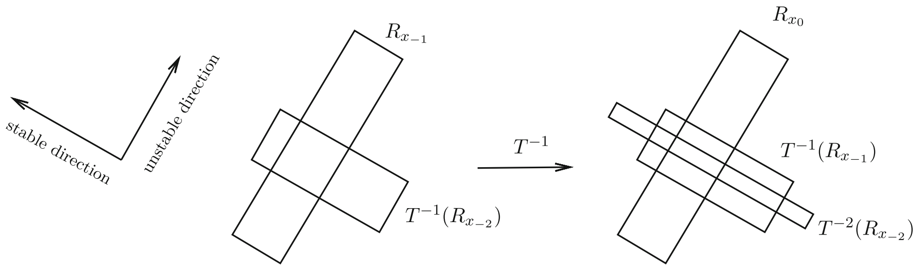

Fig. 2 Intersection of rectangles along the unstable direction [4]

of T. Conversely, the action of T^{-1} makes the width (resp. length) of R expanding (resp. contracting) by the factor λ^{-1} (resp. λ). Recall from the Markov property that if $R \cap T(R) \neq \varnothing$, then $T(R)$ intersects R the whole unstable direction, and if $R \cap T^{-1}(R) \neq \varnothing$, then $T^{-1}(R)$ intersects R the whole stable direction. Therefore, the non-empty set $Q_p^u(x) := \bigcap_{k=0}^{p} T^{-k}(R_{x_k})$ is a single rectangle stretching all the way across R_{x_0} in the expanding direction, whereas $Q_p^s(x) := \bigcap_{k=-p}^{0} T^{-k}(R_{x_k})$ is a rectangle stretching all the way across R_{x_0} in the contracting direction; see Fig. 2. This implies that the diameter of $Q_p(x) = Q_p^s(x) \cap Q_p^u(x)$ is at most λ^p.

□

Next we present the relation between the periodic sequences in $\Lambda_{\mathcal{A}}$ and the periodic orbits of T. Define

$$h : \Lambda_{\mathcal{A}} \to \mathbb{T}^2,\ h(x) = \bigcap_{n\in\mathbb{Z}} T^{-n}(R_{x_n}) \quad \text{for all} \quad x = (x_n)_{n\in\mathbb{Z}} \in \Lambda_{\mathcal{A}}. \tag{2.5}$$

Note that (2.5) is well-defined since the intersection $\bigcap_{n\in\mathbb{Z}} T^{-n}(R_{x_n})$ is a single point in $\mathbb{T}^2$ due to the fact that R_0, R_1, R_2, R_3, R_4 forms a Markov partition (see [4, p.189]). Then h is a continuous surjection and satisfies

$$h \circ \sigma = T \circ h.$$

Then

$$h \circ \sigma^n = T^n \circ h \quad \text{for all} \quad n \geq 1. \tag{2.6}$$

This implies that if $x \in \Lambda_{\mathcal{A}}$ is an n-periodic point, then $h(x)$ is an n-periodic point of T. It is clear that h is not one-to-one, but it is finite-to-one and this does not affect the investigation of periodic orbits. Indeed, the number of n-points of T is

$$\text{card}\{\mathsf{x} \in \mathbb{T}^2 : T^n(\mathsf{x}) = \mathsf{x}\} = \lambda^n + \lambda^{-n} - 2;$$

see [15].

Remark 2.5 (a) If $x = (x_i) \in P_n$, then $h(x) = \mathsf{x} \in \mathbb{T}^2$ satisfies $T^n(\mathsf{x}) = \mathsf{x}$ and

$$T^k(\mathsf{x}) \in R_{x_k}, k = 0, \ldots, n-1.$$

(b) The map $\sigma|_{\Lambda_{\mathcal{A}}}$ has three fixed points, which are 1-periodic points, including [0], [1], [4]. However, only $0 + \mathbb{Z}^2$ is the unique fixed point of T. Therefore,

$$h([0]) = h([1]) = h([4]) = 0 + \mathbb{Z}^2. \tag{2.7}$$

These fixed points are counted as n-periodic points with any $n > 0$. This explains the relation

$$\mathrm{card}\{x \in \Lambda_{\mathcal{A}} : \sigma^n(x) = x\} = \mathrm{card}\{\mathsf{x} \in \mathbb{T}^2 : T^n(\mathsf{x}) = \mathsf{x}\} + 2. \tag{2.8}$$

(c) It follows from (2.7) and (2.8) that the map

$$h : P_n \setminus \{[0], [1], [4]\} \to \{\mathsf{x} \in \mathbb{T}^2 \setminus \{0 + \mathbb{Z}^2\} : T^n(\mathsf{x}) = \mathsf{x}\}$$

is a bijection. Therefore, instead of considering periodic points of T, we shall work on periodic sequences in $\Lambda_{\mathcal{A}}$ as there are more advantages. ◇

For any n, we define two one-to-one maps

$$l : X_n \to X_{n-1}, l(x_0 \dots x_{n-1}) = x_1 \dots x_{n-1}, \tag{2.9}$$

$$r : X_n \to X_{n-1}, r(x_0 \dots x_{n-1}) = x_0 x_1 \dots x_{n-2}, \tag{2.10}$$

which remove the first symbol and the last symbols of elements in X_n, respectively. For $i, j, k \in \mathscr{A}$, define

$$X_{i;n} = \{i x_1 \dots x_{n-1} \in X_n\} \tag{2.11}$$

consisting of elements in X_n beginning with i, and

$$X_{i,j;n} = X_{i;n} \cup X_{j;n}, \quad X_{i,j,k;n} = X_{i;n} \cup X_{j;n} \cup X_{k;n} \tag{2.12}$$

the set of elements in X_n whole first symbols are in $\{i, j\}$ and in $\{i, j, k\}$, respectively. The following lemma is useful in the next section.

Lemma 2.6 (a) $l(X_{0;n}) = l(X_{1;n}) = l(X_{2;n})$ *and* $l(X_{3;n}) = l(X_{4;n})$*;*
(b) $l(X_{i;n}) \cup l(X_{j;n}) = X_{n-1}$ *for* $i \in \{0, 1, 2\}$ *and* $j \in \{3, 4\}$.

Proof Recall that if $i \in \{0, 1, 2\}$, then i is followed by 0, 1, 3, and if $j \in \{3, 4\}$, then j is followed by 2 and 4. This leads to

$$l(X_{i;n}) = X_{0;n-1} \cup X_{1;n-1} \cup X_{3;n-1} \quad \text{and} \quad l(X_{j;n}) = X_{2;n-1} \cup X_{4;n-1},$$

which proves the lemma. □

Denote by $\omega_n, \alpha_n, \beta_n, \gamma_n, \zeta_n, \eta_n$ the cardinalities of $X_n, X_{0;n}, X_{1;n}, X_{2;n}, X_{3;n}, X_{4;n}$, respectively. The following result will be useful in the next section.

Proposition 2.7 *One has*

(a) $\zeta_n = \eta_n = \omega_{n-2}$;
(b) $\alpha_n + \zeta_n = \omega_{n-1}$;
(c) $\alpha_n = \beta_n = \gamma_n = \frac{1}{3}(\omega_n - 2\omega_{n-2}) = \omega_{n-1} - \omega_{n-2}$.

Proof Since 3, 4 are followed by 2, 4 and 0, 1, 2 are followed by 0, 1, 3, one has

$$\begin{aligned} \zeta_n = \eta_n &= \gamma_{n-1} + \eta_{n-1} \\ &= (\alpha_{n-2} + \beta_{n-2} + \zeta_{n-2}) + (\gamma_{n-2} + \eta_{n-2}) = \omega_{n-2}, \end{aligned}$$

which is (a). Similarly,

$$\alpha_n + \zeta_n = (\alpha_{n-1} + \beta_{n-1} + \xi_{n-1}) + (\gamma_{n-1} + \eta_{n-1}) = w_{n-1}$$

and (b) is proved. For (c), $\alpha_n = \beta_n = \gamma_n$ is obvious. It follows from (a) and $\alpha_n + \beta_n + \gamma_n + \zeta_n + \eta_n = \omega_n$ that

$$\alpha_n = \frac{1}{3}(\omega_n - 2\omega_{n-2}). \tag{2.13}$$

For the last equality of (c), observe that

$$\alpha_n = \alpha_{n-1} + \alpha_{n-1} + \zeta_{n-1} = 3\alpha_{n-1} + 2\zeta_{n-1} - (\alpha_{n-1} + \zeta_{n-1}) = \omega_{n-1} - \omega_{n-2}.$$

The proof is complete. □

Corollary 2.8 *One has*

(a) $\omega_n = 3\omega_{n-1} - \omega_{n-2}$*;*
(b) $\alpha_n = \omega_n - 2\omega_{n-1}$ *and* $\zeta_n = 3\omega_{n-1} - \omega_n$*. In particular,* $2\omega_{n-1} < \omega_n < 3\omega_{n-1}$*;*
(c) $\omega_n = 2\omega_{n-1} + \omega_{n-2} + \cdots + \omega_2 + 8$.

Proof (a) This follows immediately from Proposition 2.7 (c).
(b) This is obtained due to $3\alpha_n + 2\zeta_n = \omega_n$ and $\alpha_n + \zeta_n = \omega_{n-1}$.
(c) Indeed,

$$\begin{aligned}\omega_n &= 3\omega_{n-1} - \omega_{n-2} = 2\omega_{n-1} + 2\omega_{n-2} - \omega_{n-3}\\ &= 2\omega_{n-1} + \omega_{n-2} + 2\omega_{n-3} - \omega_{n-4} + \cdots\\ &= 2\omega_{n-1} + \omega_{n-2} + \omega_{n-3} + \cdots + 2\omega_3 - \omega_2\\ &= 2\omega_{n-1} + \omega_{n-2} + \cdots + \omega_2 + 8,\end{aligned}$$

using $\omega_3 - \omega_2 = \omega_2 + 8$.

□

Remark 2.9 (a) Relation in Corollary 2.8 (a) is a homogeneous linear difference equation of second order. This allows us to calculate ω_n without using matrix $\mathcal{A}^n$. The characteristic equation has two eigenvalues λ, λ^{-1} as its roots; and with $\omega_1 = 5$ and $\omega_2 = 13$, we obtain again the explicit formula (2.3).
(b) From Corollary 2.8 (b) and in conjunction with (2.2)–(2.3), we have

$$\alpha_n = 3a_{n-1} + 2b_{n-1},$$

which is the sum of entries in the first row of $\mathcal{A}^n$ and this agrees with [14, Prop. 2.2.12].

3 Hierarchy of Clusters of Periodic Orbits

In this section we first show that the number of p-clusters of periodic orbits of T is the number of p-degeneracy classes. We derive an asymptotic formula for the number of p-clusters. Algorithms that list all periodic orbits and grouping them to clusters are also given.

3.1 Cluster Counting as Degeneracy Class Counting

Let $\mathsf{x} \in \mathbb{T}^2$ be an n-periodic point of T. The orbit of T through x is defined by

$$\mathcal{O}(\mathsf{x}) = \{T^i(\mathsf{x}), i = 0, 1, \ldots, n-1\}.$$

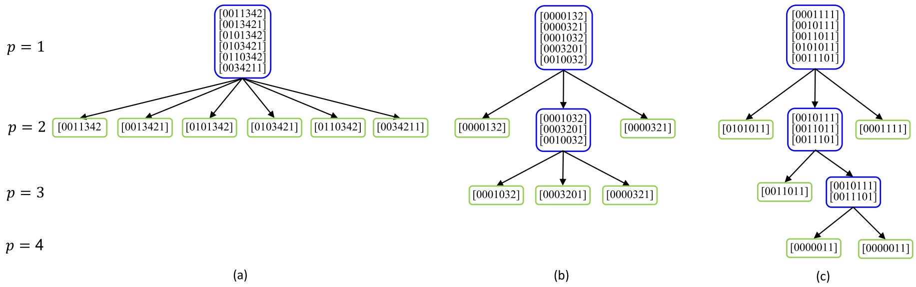

Fig. 3 Some clusters of periodic orbits with $n = 7$

Definition 3.1 Let x, y be n-periodic points of T and $p \in \mathbb{N}^*$. We say $\mathcal{O}(\mathsf{x})$ and $\mathcal{O}(\mathsf{y})$ are *p-close* if there exists a permutation $\alpha : \{0, 1 \ldots, n-1\} \to \{0, 1 \ldots, n-1\}$ such that

$$d(T^k(\mathsf{x}), T^{\alpha(k)}(\mathsf{y})) < \lambda^p, \text{ for all } k \in \{0, \ldots, n-1\}.$$

Roughly speaking, from every point on the orbit of x, there exists a point on the one of y such that the distance between them is less than λ^p. This means that these two orbits enter the same parts of $\mathbb{T}^2$ but with different order. When $\mathcal{O}(\mathsf{x})$ and $\mathcal{O}(\mathsf{y})$ are p-close, we say that they belong to the same p-cluster. So, one p-cluster contains orbits which are p-close to each other.

Next we recall an equivalence relation $\sim$ in P_n as follows. For $x, y \in P_n$, we write $x \sim y$ if there is $k \in \{0, \ldots, n-1\}$ such that $\sigma^k(x) = y$. Let $\mathcal{P}_n = P_n / \sim$. For convenience, we also write $x = [x_0 x_1 \ldots x_{n-1}] \in \mathcal{P}_n$.

Definition 3.2 [10] Let $1 \leq p \leq n$ be an integer. Two periodic sequences $x = [x_0 \ldots x_{n-1}]$, $y = [y_0 \ldots y_{n-1}] \in \mathcal{P}_n$ are *p-close* if any element in X_p appears the same number of times both in x and y. Then we write $x \overset{p}{\sim} y$.

Proposition 3.3 [10] *(i) The relation $\overset{p}{\sim}$ is an equivalence relation.*
(ii) If $x \overset{p+1}{\sim} y$, then $x \overset{p}{\sim} y$.

The first property (i) divides the set $\mathcal{P}_n$ into equivalence classes. Each equivalence class consists of sequences which are p-close to each other and it is called a *p-cluster*. The second property (ii) allows us to arrange periodic sequences as a hierarchical structure; see Fig. 3 below.

Example 3.4 According to Proposition 3.3, all periodic sequences in $\mathcal{P}_n$ can be illustrated by a line chart like Figure 2 in [10]. We present here some clusters in Fig. 3 in the case of $n = 7$.

(a) Six sequences [0011342], [0013421], [0101342], [0103421], [0110342], [0034211] belong to the same 1-cluster. This is separated into six different 2-clusters; see Fig. 3a.
(b) Five periodic sequences [0000132], [0000321], [0001032], [0003201], [0010032] are in the same 1-cluster because the numbers of times 0, 1, 2, 3, 4 appear in these sequences are 4, 1, 1, 1, 0, respectively. This 1-cluster is divided into three 2-clusters: [0000132] and [0000321] are in independent clusters, whereas [0001032], [0003201], [0010032] belong to the same 2-cluster since 00 appears twice in all three sequences, 01, 10, 03, 32, 20 appear once, and 11, 13, 21, 23, 34, 42, 44 are absent. This 2-cluster divides into three disjoint 3-clusters; see Fig. 3b.

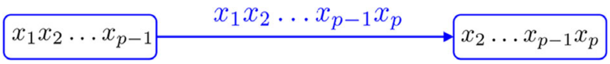

Fig. 4 Connection rule in G_p

(c) For $n = 7, p = 3$, only two clusters contain more than one element, namely: {[0001011], [0100011]} and {[0010111], [0011101]}. For $n = 7$ and $p \geq 4$, each cluster has only one element (see Fig. 3a and [10, Fig. 2]).

Remark 3.5 Since $\mathcal{A}$ is a 5×5 matrix with $a_{00} = a_{01} = a_{10} = a_{11} = 1$ and the adjacency matrix of the baker's map is $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, the hierarchy of periodic orbits with arbitrary length n consists of the hierarchy of respective periodic orbits of the baker's map. $\diamond$

Each n-periodic sequence $x \in \Lambda_{\mathcal{A}}$ corresponds to an n-periodic point of T. Therefore, each equivalence class in $\mathcal{P}_n$ corresponds to an n-periodic orbit of T. The following result allows us to study the clustering of periodic sequences in $\mathcal{P}_n$ instead of working on the clustering of periodic orbits of T.

Proposition 3.6 *If two equivalence classes $x, y \in \mathcal{P}_n$ are $(2p+1)$-close, then the respective orbits of T in $\mathbb{T}^2$, which are $\mathcal{O}(h(x))$ and $\mathcal{O}(h(y))$, are p-close.*

Proof Suppose that $x = [x_0 \ldots x_{n-1}]$, $y = [y_0 \ldots y_{n-1}] \in \mathcal{P}_n$ are $(2p+1)$-close and let $h(x) = \mathsf{x}$, $h(y) = \mathsf{y} \in \mathbb{T}^2$. Then, for every $i \in \{0, \ldots, n-1\}$, there exists $i' \in \{0, \ldots, n-1\}$ such that $x_{i+k} = y_{i'+k}$ for all $k \in \{-p, \ldots, p\}$. We can choose i' such that if $i \neq j$, then $i' \neq j'$. Define permutation $\alpha : \{0, \ldots, n-1\} \to \{0, \ldots, n-1\}$ by $\alpha(i) = i'$. We are going to show that $d(T^i(\mathsf{x}), T^{\alpha(i)}(\mathsf{y})) < \lambda^{-p}$ for all $i \in \{0, \ldots, n-1\}$. For $i \in \{0, \ldots, n-1\}$, note that $T^{i+k}(\mathsf{x}) \in R_{x_{i+k}}$ and $T^{\alpha(i)+k}(\mathsf{y}) \in R_{y_{\alpha(i)+k}}$ for all $k \in \{-p, \ldots, p\}$. Equivalently, $T^i(\mathsf{x}) \in T^{-k}(R_{x_{i+k}})$ and $T^{\alpha(i)}(\mathsf{y}) \in T^{-k}(R_{y_{\alpha(i)+k}})$ for all $k \in \{-p, \ldots, p\}$. Since $x_{i+k} = y_{\alpha(i)+k}$ for all $k \in \{-p, \ldots, p\}$, it follows that $T^i(\mathsf{x}), T^{\alpha(i)}(\mathsf{y}) \in \bigcap_{k=-p}^{p} T^{-k}(R_{i+k})$. By Lemma 2.4, $d(T^i(\mathsf{x}), T^{\alpha(i)}(\mathsf{y})) < \lambda^p$, completing the proof. $\square$

Next we are going to show that the counting problem of p-close periodic orbits is as the one of counting closed paths on the de Bruijn graph G_p passing the same number of times through its edges.

Definition 3.7 For $p \geq 2$. The *de Bruijn graph* G_p is a digraph defined as follows:

(i) the set of vertices $V(G_p)$ are given by elements in X_{p-1} and
(ii) the set of edges $E(G_p)$ are given by elements in X_p:

$$E(G_p) = \big\{(x_1 \ldots x_{p-1}, y_1 \ldots y_{p-1}) \in X_{p-1} \times X_{p-1} : x_{i+1} = y_i, \text{ for } i \in \{1, \ldots, p-2\}\big\}.$$

More instantly, the edge $x_1x_2 \ldots x_{p-1}x_p$ goes from the vertex $x_1x_2 \ldots x_{p-1}$ to the vertex $x_2 \ldots x_{p-1}x_p$ (see Fig. 4).

The graph G_p has ω_{p-1} vertices and ω_p edges. The number of incoming and outgoing edges at each vertex depends on its beginning and ending symbols, so the graph is not balanced. The vertex $x_1 \ldots x_{p-1}$ with $x_1 \in \{0, 1, 3\}$ (resp. $x_1 \in \{2, 4\}$) has 3 (resp. 2) incoming edges, namely $0x_1 \ldots x_{p-1}, 1x_1 \ldots x_{p-1}, 2x_1 \ldots x_{p-1}$ (resp. $3x_1 \ldots x_{p-1}, 4x_1 \ldots x_{p-1}$).

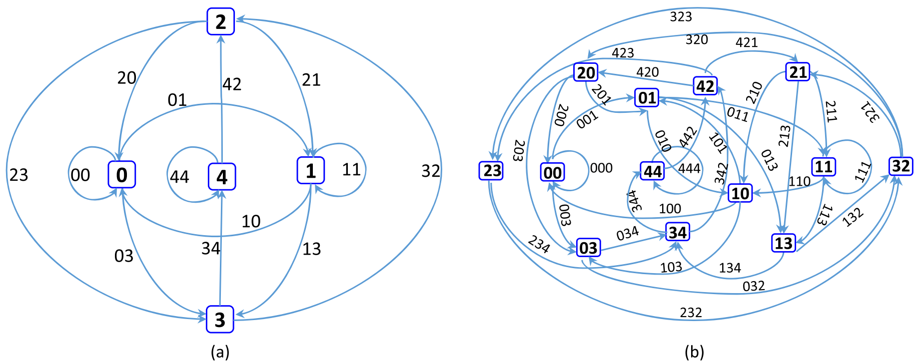

Fig. 5 Edges of graph G_2 (**a**) are vertices of graph G_3 (**b**)

The vertex $x_1 \dots x_{p-1}$ with $x_{p-1} \in \{0, 1, 2\}$ (resp. $x_{p-1} \in \{3, 4\}$) has 3 (resp. 2) outgoing edges, namely $x_1 \dots x_{p-1}0, x_1 \dots x_{p-1}1, x_1 \dots x_{p-1}3$ (resp. $x_1 \dots x_{p-1}2, x_1 \dots x_{p-1}4$). For instance, let us consider $p = 2$. The de Bruijn graph G_2 has vertices 0, 1, 2, 3, 4 and edges 00, 01, 03, 10, 11, 13, 20, 21, 23, 32, 34, 42, 44 (see Fig. 5a), whereas the graph G_3 has vertices which are edges of graph G_2, and has 34 edges which are the elements of X_3 (see Fig. 5b).

We say that a path on the graph G_p has the length of n if it passes through n edges with multiplicity. A path on G_p is called closed if the first and last vertices coincide, i.e., it starts and ends at the same vertex.

Proposition 3.8 *There is a one-to-one correspondence between the set of closed paths of length n in G_p and the set of n-periodic sequences $\mathcal{P}_n$.*

Proof Let g be a path of length n in G_p and suppose that its 1st edge, 2nd edge, ..., nth edge have the symbols $x_0 \dots x_{p-1}$, $x_1 \dots x_p$, $x_2 \dots, x_{p+1}$, $x_{n-1} \dots x_{n+p-2}$, respectively. If the path g is closed, then $x_n \dots x_{n+p-2} = x_0 \dots x_{p-2}$. We associate this closed path of length n with the n-periodic sequence $x = [x_0 x_1 \dots x_{n-1}] \in \mathcal{P}_n$. In this way, the ith edge of G_p associates to the symbol $x_i x_{i+1} \dots x_{i+p-1}$ of sequence $x = [x_0 \dots x_{n-1}]$.

Conversely, let $x = [x_0 x_1 \dots x_{n-1}] \in \mathcal{P}_n$. Then $x_i x_{i+1} \dots x_{i+p-1} \in X_p$ for $i = 0 \dots n - 1$, here $x_{n+j} = x_j$ for $j = 0, \dots, p-1$. There exists a unique path passing through the n edges $x_i x_{i+1} \dots x_{p-1+i}$, $i = 0 \dots n-1$, and since $x_n \dots x_{n+p-2} = x_0 \dots x_{p-2}$, the path is closed and has the length of n. The proof is complete. □

For $x \in \mathcal{P}_n$, denote by g_x the closed path associated the sequence x.

Definition 3.9 A *p-degeneracy class* in G_p is a family of closed paths which visit each edge of G_p with the same number of times.

For $x \in \mathcal{P}_n$ and $a \in X_p$, we denote by $n_a(x)$ the number of times $a \in X_p$ appears in x. We associate with each $x \in \mathcal{P}_n$ a ω_p-dimensional integer vector with non-negative components

$$x \mapsto \boldsymbol{n}_p(x) = \{n_a(x)\}_{a \in X_p} \in \mathbb{Z}_+^{\omega_p}.$$

Proposition 3.10 *Let $x, y \in \mathcal{P}_n$. The following assertions are equivalent:*

(a) $x \overset{p}{\sim} y$;

(b) $\boldsymbol{n}_p(x) = \boldsymbol{n}_p(y)$;
(c) g_x *and* g_y *belong to the same* p*-degeneracy class.*

Proof (a) $\Leftrightarrow$ (b): This is obvious. (b) $\Leftrightarrow$ (c): By Definition 3.9, g_x and g_y are in the same p-degeneracy class if and only if g_x and g_y visit each edge the same number of times. By the proof of Proposition 3.8, this is equivalent to $n_a(x) = n_a(y)$ for all $a \in X_p$, or $\boldsymbol{n}_p(x) = \boldsymbol{n}_p(y)$, which is (b). □

The above result shows that the problem of counting number of clusters of periodic sequences is equivalent to the one of counting number of degeneracy classes. The latter will be solved in the next section.

3.2 The Number of *p*-Degeneracy Classes

We treat the elements in X_p as p-digit-numbers in the decimal numerical system and arrange them in the ascending order with the convention of $2 > 3$. This convention makes things easier to get the explicit forms of the matrices S_p and R_p below. For instance, let us consider some small values of n:

$$\begin{aligned} X_1 =&\{0, 1, 3, 2, 4\}, \quad X_2 = \{00, 01, 03, 10, 11, 13, 32, 34, 20, 21, 23, 42, 44\}, \\ X_3 =& \{000, 001, 003, 010, 011, 013, 032, 034, 100, 101, 103, 110, 111, 113, 132, 134, 320, \\ & 321, 323, 342, 344, 200, 201, 203, 210, 211, 213, 232, 234, 420, 421, 423, 442, 444\}. \end{aligned}$$

For $p \geq 2$, let $S_p = \{S_p(a, b)\}_{a \in X_{p-1}, b \in X_p}$ be a matrix with rows labelled by the elements in X_{p-1} and columns labelled by the elements in X_p, defined by

$$S_p(a, b) = \begin{cases} 1 & \text{if } a \text{ is the last } p-1 \text{ symbols of } b, \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

and let $R_p = \{R_p(a, b)\}_{a \in X_p, b \in X_{p-1}}$ be a matrix with rows labelled by the elements in X_p and columns labelled by the elements in X_{p-1}, defined by

$$R_p(a, b) = \begin{cases} 1 & \text{if } b \text{ is the first } p-1 \text{ symbols of } a, \\ 0 & \text{otherwise.} \end{cases} \tag{3.15}$$

In other words,

$$S_p(a, b) = \begin{cases} 1 & \text{if } l(b) = a, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad R_p(a, b) = \begin{cases} 1 & \text{if } r(a) = b, \\ 0 & \text{otherwise,} \end{cases} \tag{3.16}$$

where the maps l and r are defined in (2.9)–(2.10). Recall that the elements in X_p and X_{p-1} indicate the edges and vertices of G_p, respectively. So $S_p(a, b) = 1$ means that the edge having the code a is possible to enter the vertex with the code b, whereas $R_p(a, b) = 1$ means that the edge with the code b can exit from the vertex having the code a. Intuitively, for $p = 2$ the connection rule of S_2 is illustrated in Table 1.

Table 1 The connection rule of S_2

$X_1 \backslash X_2$	00	01	03	10	11	13	32	34	20	21	23	42	44
0	1	0	0	1	0	0	0	0	1	0	0	0	0
1	0	1	0	0	1	0	0	0	0	1	0	0	0
3	0	0	1	0	0	1	0	0	0	0	1	0	0
2	0	0	0	0	0	0	1	0	0	0	0	1	0
4	0	0	0	0	0	0	0	1	0	0	0	0	1

The matrices S_2 and R_2 are given by

$$S_2 = \begin{pmatrix} 1&0&0&1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&1&0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&1&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1&0&0&0&0&1 \end{pmatrix}, R_2 = \begin{pmatrix} 1&0&0&0&0 \\ 1&0&0&0&0 \\ 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 0&0&1&0&0 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&1&0 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \\ 0&0&0&0&1 \end{pmatrix}.$$

Define

$$F = \begin{pmatrix} 1&0&0 \\ 1&0&0 \\ 1&0&0 \\ 0&1&0 \\ 0&1&0 \\ 0&1&0 \\ 0&0&1 \\ 0&0&1 \end{pmatrix}.$$

For $p \geq 3$, we obtain the following result.

Theorem 3.11 *One has*

$$S_p = \begin{pmatrix} I_{\alpha_p} & I_{\alpha_p+\zeta_p} & I_{\alpha_p+\zeta_p} \\ \star & & \end{pmatrix}, \quad p \geq 3, \tag{3.17}$$

$$R_3 = \begin{pmatrix} F & \star & \star \\ \star & R_2 & \star \\ \star & \star & R_2 \end{pmatrix}, \tag{3.18}$$

Table 2 The connection rule of S_p

X_{p-1} \ X_p	$0X_{0,1,3;p-1}$	$1X_{0,1,3;p-1}$	$3X_{2,4;p-1}$	$2X_{0,1,3;p-1}$	$4X_{2,4;p-1}$
$X_{0,1,3;p-1}$	I_{α_p}	I_{α_p}	$\star$	I_{α_p}	$\star$
$X_{2,4;p-1}$	$\star$	$\star$	I_{ζ_p}	$\star$	I_{ζ_p}

$$R_p = \begin{pmatrix} F & \star & \star & \dots & \star & \star & \star \\ \star & R_2 & \star & \dots & \star & \star & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \star & \star & \star & \dots & R_{p-2} & \star & \star \\ \star & \star & \star & \dots & \star & R_{p-1} & \star \\ \star & \star & \star & \dots & \star & \star & R_{p-1} \end{pmatrix}, \quad p \geq 4, \tag{3.19}$$

where the symbol $\star$ *denotes the matrices of entries* 0 *of suitable dimensions; recall* α_p *and* ζ_p *from Proposition 2.7.*

Proof We first verify the form of S_p. Recall that S_p is formed by rows and columns labelled by the elements in X_{p-1} and X_p, respectively. Decompose

$$X_{p-1} = X_{0,1,3;\,p-1} \cup X_{2,4;\,p-1}$$

and

$$X_p = 0X_{0,1,3;\,p-1} \cup 1X_{0,1,3;\,p-1} \cup 3X_{2,4;\,p-1} \cup 2X_{0,1,3;\,p-1} \cup 4X_{2,4;\,p-1};$$

recall (2.12). Note that the set X_p is obtained by adding labels 0, 1, 2 to the left of $X_{0,1,3;\,p-1}$ and 3, 4 to the left of $X_{2,4;\,p-1}$. The matrix S_p is illustrated by Table 2.

Recall from the proof of Proposition 2.7 that $\mathrm{card}(X_{0,1,3;\,p-1}) = \alpha_p$ and $\mathrm{card}(X_{2,4;\,p-1}) = \zeta_p$. Then

$$\mathrm{card}(0X_{0,1,3;\,p-1}) = \mathrm{card}(1X_{0,1,3;\,p-1}) = \mathrm{card}(2X_{0,1,3;\,p-1}) = \alpha_p$$

and

$$\mathrm{card}(3X_{2,4;\,p-1}) = \mathrm{card}(4X_{2,4;\,p-1}) = \zeta_p.$$

It is clear that

$$S_p(a,b) = \begin{cases} I_{\alpha_p}(a,b) & \text{if } a \in X_{0,1,3;\,p-1}, b \in 0X_{0,1,3;\,p-1} \cup 1X_{0,1,3,\,p-1} \cup 2X_{0,1,3;\,p-1}, \\ I_{\zeta_p}(a,b) & \text{if } a \in X_{2,4,\,p-1}, b \in 2X_{0,1,3;\,p-1} \cup 4X_{2,4;\,p-1}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.20}$$

This leads to the form of S_p in (3.17).

Next we verify the form of R_p. Let us start with $p = 3$. The symbols of columns of F are the first three elements of X_2, namely

$$C_F = \{00, 01, 03\}.$$

The symbols of rows of F are the first eight elements in X_3, obtained by adding allowed symbols to the right of C_F, including

$$R_F = \{000, 001, 003, 010, 011, 013, 032, 034\}.$$

Table 3 The connection rule in R_p

X_p \ X_{p-1}	$X_{0;p-1}$	$X_{1,3;p-1}$	$X_{2,4;p-1}$
$X_{0;p}$	F_p	$\star$	$\star$
$X_{1,3;p}$	$\star$	R_{p-1}	$\star$
$X_{2,4;p}$	$\star$	$\star$	R_{p-1}

This explains the presence of F in R_3. We also obtain that $R_3(a, b) = 0$ with $a \notin R_F$, $b \in C_F$.

The next elements of X_2 after R_F are

$$C_{R_2} = \{10, 11, 13, 32, 34\}.$$

The next elements of X_3 after R_F are obtained by adding allowed symbols to the right of elements in C_{R_2}, including

$$R_{R_2} = \{100, 101, 103, 110, 111, 113, 132, 134, 320, 321, 323, 342, 344\}.$$

This leads to the appearance of the first matrix R_2 and $R_3(a, b) = 0$ for $b \in C_{R_2}, b \notin R_{R_2}$. An analogous argument is applied to have the second appearance of R_2 and we obtain the form of R_3 in (3.18).

For $p \geq 4$, we decompose the rows

$$X_p = X_{0;p} \cup X_{1,3;p} \cup X_{2,4;p}$$

and the columns

$$X_{p-1} = X_{0;p-1} \cup X_{1,3;p-1} \cup X_{2,4;p-1}.$$

We first calculate the entries of R_p in rows $X_{1,3;p} \cup X_{2,4;p}$ and columns $X_{1,3;p-1} \cup X_{2,4;p-1}$. By the definition of R_p in (3.15) we observe that if $a \in X_p$ and $b \in X_{p-1}$ have the same first symbol, then $R_p(a, b) = R_{p-1}(l(a), l(b))$. In other words, for any $i \in \{0, 1, 2, 3, 4\}$

$$R_p(a, b) = R_{p-1}(l(a), l(b)) \quad \text{for} \quad (a, b) \in X_{i,p} \times X_{i,p-1}. \tag{3.21}$$

By Lemma 2.6(b) that

$$l(X_{1,3;p}) = X_{p-1} \quad \text{and} \quad l(X_{1,3;p-1}) = X_{p-2},$$

and $X_{1,3;p} = X_{1;p} \cup X_{3;p}$ and $X_{2,4;p} = X_{2,4;p-1}$, the relation (3.21) leads to the first and the second presences of R_{p-1} in R_p; see Table 3 for an illustration.

Next, for the entries with rows labelled by $X_{0;p}$ and columns labelled by $X_{0;p-1}$, we define

$$F_p(a, b) = R_p(a, b) \text{ for } (a, b) \in X_{0;p} \times X_{0;p-1}; \tag{3.22}$$

see Table 3. This means that F_p is the matrix formed from the first α_p rows and the first α_{p-1} columns of R_p. It remains to verify the following claim:

$$F_p = \text{diag}(F, R_2, \ldots, R_{p-2}) \quad \text{for} \quad p \geq 4. \tag{3.23}$$

Formula (3.23) is rewritten as

$$F_p = \text{diag}(F_{p-1}, R_{p-2}) \quad \text{for} \quad p \geq 4. \tag{3.24}$$

Table 4 The connection rule of F_p

$X_{0;p}$ \ $X_{0;p-1}$	$X_{00;p-1}$	$X_{01;p-1}$ $X_{03;p-1}$
$X_{00;p}$	F_{p-1}	$\star$
$X_{01;p}$ $X_{03;p}$	$\star$	R_{p-2}

Decompose the rows

$$X_{0;p} = X_{00;p} \cup X_{01;p} \cup X_{03;p}$$

and the columns

$$X_{0;p-1} = X_{00;p-1} \cup X_{01;p-1} \cup X_{03;p-1}.$$

Observe from property (3.21) and definition (3.22) that

$$F_p(a,b) = R_p(a,b) = R_{p-1}(l(a), l(b)) \quad \text{for} \quad (a,b) \in X_{0;p} \times X_{0;p-1}. \tag{3.25}$$

In particular,

$$F_p(a,b) = F_{p-1}(l(a), l(b)) \quad \text{for} \quad (a,b) \in X_{00;p} \times X_{00;p-1}; \tag{3.26}$$

note that (3.26) is well-defined since $(l(a), l(b)) \in X_{0;p-1} \times X_{0;p-2}$. In addition, owing to $l(X_{00;p}) = X_{0;p-1}$ and $l(X_{00;p-1}) = X_{0;p-2}$, relation (3.26) leads to the presence of F_{p-1} in F_p. From (3.25) one also has

$$F_p(a,b) = R_{p-2}(l^2(a), l^2(b)) \quad \text{for} \quad (a,b) \in (X_{01;p} \times X_{01;p}) \cup (X_{03;p-1} \times X_{03;p-1}). \tag{3.27}$$

If $(a,b) \in (X_{01;p} \times X_{01;p}) \cup (X_{03;p-1} \times X_{03;p-1})$, then $(l^2(a), l^2(b)) \in (X_{p-2}, X_{p-3})$. Furthermore, by Lemma 2.6 (b),

$$l^2(X_{01;p} \cup X_{03;p}) = X_{p-2} \quad \text{and} \quad l^2(X_{01;p-1} \cup X_{03;p-1}) = X_{p-3}$$

the presence of R_{p-2} in F_p. Finally, is clear that if $a \in X_{0;p}$ and $b \in X_{0,p-1}$ have different first symbols, then $F_p(a,b) = R_p(a,b) = 0$. We have proved (3.24). For an illustration, see Table 4.

□

Remark 3.12 (a) Each row of S_p as well as R_p has only one entry 1 and each column has two or three entries 1. The columns with two entries 1 correspond to the elements ending with 3 or 4 since 3 and 4 are followed by 2 and 4.
(b) Formula (3.19) explains the equality in Corollary 2.8(c).
(c) This argument can be generalized for arbitrary adjacency matrix with entries 0 and 1. ◇

Theorem 3.13 *The number of p-clusters in $\mathcal{P}_n$ is the number of vectors $\boldsymbol{n}_p = (n_a)_{a \in X_p} \in \mathbb{Z}_+^{\omega_p}$ satisfying the following constraints*

(i)

$$\sum_{a \in X_p} n_a = n; \tag{3.28}$$

(ii)

$$S_p \boldsymbol{n}_p = R_p^T \boldsymbol{n}_p; \tag{3.29}$$

(iii) *If there is* $a \in \{[0], [1], [4]\}$ *with* $0 < n_a < n$*, then* $n_b > 0$ *for some* $b \in X_{k;p} \setminus \{a\}$*, where* k *is the first symbol of* a.

Proof Due to the fact that the matrices S_p and R_p^T have dimension $\omega_{p-1} \times \omega_p$, system (3.28)–(3.29) consists of $\omega_{p-1} + 1$ linear equations with ω_p variables. Since each vector $\boldsymbol{n}_p$ corresponds to a periodic sequence of length n, its coordinates must satisfy the following:

(i) The length of the closed path is equal to n, so (3.28) holds;
(ii) The number of times the closed path visiting a vertex of G_p is the same to the one of this path exiting this vertex. This condition is expressed by equation (3.29).

Furthermore, if $(a, b) \in \{([0], [0]), ([1], [1]), ([4], [4])\} \subset X_{p-1} \times X_p$, then $S_p(a, b) = R_p^T(a, b)$. This means that the visiting and exiting at each vertex [0], [1], [4] are the same, and the variables $n_a, a \in \{[0], [1], [4]\}$ in (3.29) are eliminated after simplifying. Therefore, constraint (iii) is necessary to guarantee the path is connected since otherwise the subpath visiting the edge a would be isolated. The theorem is proved. □

Now we are in a position to determine the number of clusters. For, we first verify that

$$\operatorname{rank}(S_p - R_p^T) = \omega_{p-1} - 1. \tag{3.30}$$

Recall that $S_p - R_p^T$ has dimension $\omega_{p-1} \times \omega_p$. Every column of S_p and R_p^T has only one element equal to 1, whereas the others equal 0. Summing all the rows of $S_p - R_p^T$ gives a null row. This yields $\operatorname{rank}(S_p - R_p^T) \le \omega_{p-1} - 1$. In order to get the equality, we consider the square matrix E_p which is formed from the last ω_{p-1} columns of $S_p - R_p^T$ and has the form

$$E_{p-1} = I_{\omega_{p-1}} - T_{p-1},$$

where $T_{p-1} = (a_{ij})_{\omega_{p-1} \times \omega_{p-1}}$ is the matrix formed from the last ω_{p-1} columns of R_p^T. Since $a_{ij} = 0$ for all $1 \le i \le j \le \omega_{p-1} - 1$, it follows that the $(\omega_{p-1} - 1)$th main subdeterminant of E_{p-1} is 1, and so (3.30) holds. This implies that system (3.29) contains $\omega_{p-1} - 1$ linearly independent equations with ω_p variables. By choosing $\omega_p - \omega_{p-1} =: \kappa_p$ appropriate variables freely, the remaining ω_{p-1} ones are uniquely fixed by equations (3.28)–(3.29). Furthermore, the constraint $n_a \ge 0$ for $a \in X_p$ must be satisfied. They define a κ_p-polytope $\mathcal{V}_p$ in $\mathbb{R}^{\kappa_p}$. For instance, the polytope $\mathcal{V}_p$ is given by

$$\begin{cases} n - (a_{11}x_1 + \cdots + a_{1\kappa_p}x_{\kappa_p}) \ge 0 \\ a_{21}x_1 + \cdots + a_{2\kappa_p}x_{\kappa_p} \ge 0 \\ \cdots \\ a_{\omega_{p-1}1}x_1 + \cdots + a_{\omega_{p-1}\kappa_p}x_{\kappa_p} \ge 0 \\ x_i \ge 0, i = 1, \ldots, \kappa_p \end{cases}$$

for some integers $a_{ij}, i = 1, \ldots, \omega_p - 1, j = 1, \ldots, \kappa_p$. Consider the polytope $\mathcal{W}_p$ given by

$$\begin{cases} a_{11}x_1 \quad + \cdots + a_{1\kappa_p}x_{\kappa_p} \le 1 \\ a_{21}x_1 \quad + \cdots + a_{2\kappa_p}x_{\kappa_p} \ge 0 \\ \cdots \\ a_{\omega_{p-1}1}x_1 + \cdots + a_{\omega_{p-1}\kappa_p}x_{\kappa_p} \ge 0 \\ x_i \ge 0, i = 1, \ldots, \kappa_p. \end{cases}$$

Then $\mathcal{V}_p = n\mathcal{W}_p$ and the number of solutions of systems (3.28)–(3.29) with non-negative integer components is $\mathrm{card}(n\mathcal{W}_p \cap \mathbb{Z}^{\kappa_p})$. According to Ehrhart's theorem (see [5, Section 3.4]), this number is a polynomial in n with the leading term $\mathrm{vol}(\mathcal{W}_p)n^{\kappa_p}$, where $\mathrm{vol}(\mathcal{W}_p)$ is the volume of polytope $\mathcal{W}_p$; see [13] for a calculation of $\mathrm{vol}(\mathcal{W}_p)$. We have shown the following result.

Theorem 3.14 *The number of p-clusters of n-periodic sequences in $\mathcal{P}_n$ satisfies the asymptotic formula*

$$N(n, p) = w_p n^{\kappa_p}(1 + O(1/n)) \quad as \quad n \to \infty,$$

where w_p is a constant depending on p.

Note that $\kappa_p = \omega_p - \omega_{p-1}$ is obtained from (2.3):

$$\kappa_p = \frac{\sqrt{5}}{10}\Big[\Big(\frac{3+\sqrt{5}}{2}\Big)^{p-1} - \Big(\frac{3-\sqrt{5}}{2}\Big)^{p-1}\Big],$$

where w_p can be explicitly computed for small value p.

Remark 3.15 Condition (iii) in Theorem 3.13 can be rewritten as follows: For every $a \in \{[0], [1], [4]\}$, $n_a(n - n_a)(\sum_{b \in X_{k;p}\setminus\{a\}} n_b - 1) \geq 0$, where k is the first symbol of a. For $p = 2$, this becomes

$$n_{00}(n-n_{00})(n_{01}+n_{03}-1) \geq 0, n_{11}(n-n_{11})(n_{10}+n_{13}-1) \geq 0, n_{44}(n-n_{44})(n_{42}-1) \geq 0. \tag{3.31}$$

Example 3.16 (a) The number of 2-clusters in $\mathcal{P}_n$ is represented by the number of vectors

$$\boldsymbol{n}_2 = (n_{00}, n_{01}, n_{03}, n_{10}, n_{11}, n_{13}, n_{32}, n_{34}, n_{20}, n_{21}, n_{23}, n_{42}, n_{44}) \in \mathbb{Z}_+^{13}$$

satisfying the system

$$\begin{cases} n_{00} + n_{01} + n_{03} + n_{10} + n_{11} + n_{13} + n_{32} & \\ +n_{34} + n_{20} + n_{21} + n_{23} + n_{42} + n_{44} & = n \\ n_{01} + n_{03} - n_{10} - n_{20} & = 0 \\ n_{03} + n_{13} - n_{20} - n_{21} & = 0 \\ n_{03} + n_{13} + n_{23} - n_{32} - n_{34} & = 0 \\ n_{34} - n_{42} & = 0 \end{cases} \tag{3.32}$$

with constraint (3.31). For $n = 7$, the above system has 98 solutions with non-negative integer coordinates, so the number of 2-clusters is 98, which agrees the one in Table 5. Note that without constraint (3.31), i.e, without conditions (iii), this number is 114.

In order to know how periodic orbits contribute to each cluster, we can use algorithms in Sect. 3.3. The total 123 periodic orbits are grouped in 98 2-clusters. There are 77 clusters in which each cluster has one element, 17 clusters with two elements and four clusters with three elements. The 2-clusters with three elements are {[0001011], [0010011], [0001101]}, {[0010111], [0011011], [0011101]}, {[0001032], [0003201], [0010032]}, {[0111321], [0113 211], [0132111]}.

(b) For arbitrary n, choose independent integers $n_{01} = x_1, n_{03} = x_2, n_{13} = x_3, n_{20} = x_4, n_{23} = x_5, n_{34} = x_6, n_{00} = x_7, n_{11} = x_8$. Then $n_{10} = x_1 + x_2 - x_4$, $n_{21} = x_2 + x_3 - x_4$, $n_{32} = x_2 + x_3 + x_5 - x_6$, $n_{42} = x_6$, $n_{44} = n - (2x_1 + 4x_2 + 3x_3 - x_4 + 2x_5 + x_6 + x_7 + x_8)$. The number of 2-clusters is the number of points with non-negative integer components in the polytope $\mathcal{V}_8$ consisting of $(x_1, \ldots, x_8) \in \mathbb{R}^8$ satisfying the constraint

$$\begin{cases} n-(2x_1+4x_2+3x_3-x_4+2x_5+x_6+x_7+x_8)\geq 0,\\ x_1+x_2-x_4\geq 0,\\ x_2+x_3-x_4\geq 0,\\ x_2+x_3+x_5-x_6\geq 0,\\ x_i\geq 0, i=1,\ldots,8. \end{cases} \tag{3.33}$$

Consider the polytope P defined by

$$\begin{cases} 2x_1+4x_2+3x_3-x_4+2x_5+x_6+x_7+x_8\leq 1,\\ -x_1-x_2+x_4\leq 0,\\ -x_2-x_3+x_4\leq 0,\\ -x_2-x_3-x_5+x_6\leq 0,\\ -x_i\leq 0, i=1,\ldots,8. \end{cases} \tag{3.34}$$

The volume of polytope P is

$$\mathrm{vol}(P)=\frac{1}{414720}.$$

Then, the number of solutions of system (3.31)–(3.32) with non-negative integer components is

$$N(n,2)=\frac{n^8}{414720}+O(n^7)\quad\text{as}\quad n\to\infty.$$

It is worth noting that the leading term of $N(n,2)$ does not change in the case of constraint (3.31) applied.

Remark 3.17 (a) In [10, Appendix C], the authors did not require constraint (iii) in Theorem 3.13. As a result, this leads to a larger number of clusters than the one it must be. In Example, they considered $p=2$ and $n=7$. Constraints (i) and (ii) in Theorem 3.13 take the form

$$n_{01}=n_{10},\quad n_{11}+n_{10}+n_{00}=7,$$

which has 20 non-negative integer solutions $(n_{00},n_{01},n_{10},n_{11})$: (0,0,0,7), (0,1,1,5), (0,2,2,3), (0,3,3,1), (1,0,0,6), (1,1,1,4), (1,2,2,2), (2,0,0,5), (2,1,1,3), (2,2,2,1), (3,0,0,4), (3,1,1,2), (3,2,2,0), (4,0,0,3), (4,1,1,1), (5,0,0,2) (5,1,1,0), (6,0,0,1), (7,0,0,0). However, according to [10, Figure 2], there are only 14 2-clusters because the solutions (1,0,0,6), (2,0,0,5), (3,0,0,4), (4,0,0,3), (5,0,0,2), (6,0,0,1) do not give us any periodic orbits. The matter of fact is that if the symbols 00 and 11 both appear in the same sequence then there must be the presences of 01 and 10. This means that the constraints in [10, Appendix C] are not enough and constraint (iii) in Theorem 3.13 is necessary. Nevertheless, this does not affect the leading asymptotic formula for the number of p-clusters for arbitrary p, in particular, for $p=2$ (see [10, (C.3)]).

(b) Theorem 3.13 gives the number of p-clusters but it does not show any information about which sequences are in the same cluster. Algorithms in the next subsection will help us overcome this disadvantage. ◇

3.3 Algorithms that List All Clusters

In the remaining part of this paper we present algorithms that list all periodic orbits and group them into p-clusters for all $p\leq n$. In what follows, $|X|$ denoted the cardinality of a given set X.

Algorithm 1: Find X_n, the set of all possible sequences of length n

Input: Matrix $\mathcal{A}$, length n
Output: X_n - the set of all possible sequences of length n

/* Create the set X_2 */
$X_2 = \varnothing$ // Initialize the set X_2
for $i = 1 : 5$ **do**
  for $j = 1 : 5$ **do**
    if $\mathcal{A}_{ij} \neq 0$ **then**
      $X_2 = X_2 \cup [i - 1, j - 1]$

/* Partition the set X_2 into basic subsets */
for $i = 0 : 4$ **do**
  $X_{i;2} = \varnothing$ // Initialize the set $X_{i;2}$
for $i = 1 : |X_2|$ **do**
  Let $j = X_2(i, 1)$ // Get the index of the basic subset
  $X_{j;2} = X_{j;2} \cup X_2(i, :)$

/* Create the set X_n */
for $k = 3 : n$ **do**
  $X_k = \varnothing$ // Initialize the set X_k
  for $i = 1 : |X_{k-1}|$ **do**
    Set $j = X_{k-1}(i, k - 1)$
    $X_k = X_k \cup [X_{k-1}(i, 1 : k - 2), X_{j;2}]$

Algorithm 2: Find the set consisting of all periodic sequences of period n

Input: X_n - the set of all possible sequences of length n
Output: $\mathcal{P}_n$ - the set consisting of all periodic sequences in $\Lambda_{\mathcal{A}}$ of period n after eliminating permutations

// Initialize the set $\mathcal{P}_n$
$\mathcal{P}_n = \varnothing$
for $i = 1 : |X_n|$ **do**
  // Take the last and the first elements of the ith sequence
  Set $x_{\text{last}} = X_n(i, n)$, $x_{\text{first}} = X_n(i, 1)$
  if $[x_{last}, x_{first}] \in X_2$ **then**
    // Create the set of all permutations of the ith sequence
    $X_{\text{permutation}} = \{x \mid x \text{ is a permutation of } X_n(i, :)\}$
    // Check that no permutation of i-th (periodic) sequence in $\mathcal{P}_n$
    if $X_{permutation} \cap \mathcal{P}_n = \varnothing$ **then**
      $\mathcal{P}_n = \mathcal{P}_n \cup X_n(i, :)$

Algorithm 3: Create equivalence classes with respect to the relation $\overset{p}{\sim}$

Input: $\mathcal{P}_n$ - the set of n-periodic sequences, $p \in \mathbb{N}$, X_p - the set of all possible sequences of length p

Output: $X_{\text{branch2cluster}}$: set of equivalence classes of $\mathcal{P}_n$ with respect to the relation $\overset{p}{\sim}$

Set $m = |\mathcal{P}_n|$
/* Create a matrix X_{count} whose columns are the number of occurrences of an element of $\mathcal{P}_n$ in X_p */
Set $X_{\text{count}} = O_{m \times |X_p|}$ // zero matrix of size $m \times |X_p|$
Set $X_{\text{countU}} = \varnothing$ // the set of unique elements in X_{count} (eliminate equivalence)
for $i = 1 : m$ **do**
    for $j = 1 : n$ **do**
        if $j \leq n + 1 - p$ **then**
            // Take p elements in $\mathcal{P}_n(i)$ from the j-element
            Set $X_{\text{window}} = \mathcal{P}_n(i, j : j + p - 1)$
        else
            // Take the last $(n - j)$ elements and the first $(p - (n - j))$ elements of $\mathcal{P}_n(i)$
            Set $X_{\text{window}} = [\mathcal{P}_n(i, j : n), \mathcal{P}_n(1 : j + p - n - 1)]$
        // Save the number of occurences of X_{window} in X_p
        for $k = 1 : |X_p|$ **do**
            if $X_{window} = X_p(k)$ **then**
                $X_{\text{count}}(i, k) = X_{\text{count}}(i, k) + 1$
    // Save this element to X_{countU} if it has not yet been appeared in this set
    Set $X_{\text{countU}} = X_{\text{countU}} \cup X_{\text{count}}(i, :)$
Set $N = |X_{\text{countU}}|$
for $k = 1 : N$ **do**
    // The indices of equivalet elements in X_{count}
    $I = \{i \in \{1, \ldots, m\} \mid X_{\text{count}}(i, :) = X_{\text{countU}}(k)\}$
    Set $X_{\text{branch2cluster}}(k) = \mathcal{P}_n(I)$

Algorithm 4: Find all p-close sequences in $\mathcal{P}_n$

```
Input: P_n, X_n
Output: p-close sequences
Set m = |P_n|.
Set p = 0.
// The set of clusters will be branched in the next iteration
Set BranchNext = {P_n}.
repeat
    // The set of clusters is branched in this iteration
    Set BranchCurrent = BranchNext.
    Set BranchNext = ∅
    p = p + 1.
    Set Cluster(p) = ∅                    // the set of all p-close sequences
    for i = 1 : |BranchCurrent| do
        Perform Algorithm 3 to get the set of clusters X_branch2cluster, where the input is
        P_n = BranchCurrent(i), p and X_p.
        Set Cluster(p) = Cluster(p) ∪ X_branch2cluster.
        for k = 1 : |X_branch2cluster| do
            if |X_branch2cluster(k)| ≥ 2 // Check if each cluster has at least 2
                sequences
            then
                BranchNext = BranchNext ∪ X_branch2cluster(k)
until BranchNext = ∅
```

Running the above algorithms, we get the result in Table 5, in which the symbol $\star$ indicates the respective number of periodic orbits. Note that for a fixed n, the number of p-clusters is monotonically increasing on p since one p-cluster may be decomposed into several $(p+1)$-clusters..

The more number of p-clusters, the less number of that containing many periodic orbits. Based on the result above, the maximum of p, denoted by $p_{\max}$, for which there are at least one p-cluster having more than one element is illustrated in the table below.

Appendix: Complexity of Algorithms

A1 Algorithm 1

First, we will prove that the number of elements in the set X_n is bounded by $13 \cdot 3^{n-2}$, or equivalently, $|X_n| = O(3^{n-2})$ for all $n \geq 2$ by induction. For, recall from Corollary 2.8 (b) that $|X_n| < 3|X_{n-1}|$ for $n \geq 2$. Then $|X_n| < 13 \cdot 3^{n-2}$ for $n \geq 2$. From lines 1 to 10 of Algorithm 1, the complexity is $O(1)$ (independent of n). At line 13, for each $k \in \{3, \dots, n\}$, the algorithm performs $|X_{k-1}|$ iterations. Therefore, the complexity of the algorithm is

$$|X_2| + |X_3| + \cdots + |X_{n-1}| \leq 13 + 13 \cdot 3 + \cdots + 13 \cdot 3^{n-3} = \frac{13}{2}(3^{n-2} - 1).$$

Or equivalently, the complexity of Algorithm 1 is $O(3^{n-2})$.

Table 5 Number of p-clusters

n	$p(n)$	Number of p-clusters									
		$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 8$	$p = 9$	$p = 10$
3	8	**8**	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$
4	15	13	**15**	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$
5	27	20	**27**	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$
6	60	30	53	59	**60**	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$
7	123	43	98	121	**123**	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$
8	285	60	182	267	283	**285**	$\star$	$\star$	$\star$	$\star$	$\star$
9	648	81	325	578	639	646	**648**	$\star$	$\star$	$\star$	$\star$
10	1529	107	565	1250	1486	1523	**1529**	$\star$	$\star$	$\star$	$\star$
11	3603	138	950	2664	3452	3586	3601	**3603**	$\star$	$\star$	$\star$
12	8680	175	1558	5631	8096	8592	8665	8678	**8680**	$\star$	$\star$
13	20,883	218	2482	11,670	18,890	20,601	20,852	20,877	**20,883**	$\star$	$\star$
14	50,825	268	3864	23,805	44,112	49,788	50,692	50,803	50,823	**50,825**	$\star$
15	124,056	325	5878	47,635	102,621	120,541	123,585	123,980	124,036	124,052	**124,056**
16	304,575	390	8764	93,562	237,965	292,742	302,957	304,352	304,539	304,571	**304,575**

Table 6 The largest p for which there is at least one p-cluster containing at least two elements

n	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$p_{\max}$	$\nexists$	1	1	3	3	4	5	5	6	7	7	8	9	9

A2 Algorithm 2

The main computational complexity of Algorithm 2 lies in generating permutations of a sequence with n elements (line 5 of this algorithm), which has a complexity of $O(n!)$. The algorithm needs to perform $|X_n|$ iterations (line 2). Recall that $|X_n| \leq 13 \cdot 3^{n-2}$ for all $n \geq 2$. Therefore, combining these above results, the overall complexity of Algorithm 2 is $O(3^{n-2} \cdot n!)$.

A3 Algorithm 3

We first recall that $|P_n| = \lambda^n + \lambda^{-n}$. Since $\mathcal{P}_n$ is the set of elements in P_n which are not related by a cyclic shift and the shift map σ has three fixed points, one has

$$m := |\mathcal{P}_n| = \frac{\lambda^n + \lambda^{-n} - 3}{n} + 3.$$

From lines 6 to 9 of Algorithm 3, extracting p elements from $\mathcal{P}_n(i)$ has a complexity of $O(p)$. Comparing two sequences with p elements in line 11 also has a complexity of $O(p)$. Thus, the complexity of lines 10 to 12 is $O(p|X_p|) = O(p \cdot 3^{p-2})$ (according to Sect. 1). Consequently, the complexity of lines 5 to 12 is $O(n(p+p\cdot 3^{p-2})) = O(np\cdot 3^{p-2})$. Updating elements in line 13 involves comparing with elements already present in X_{countU} up to m times, and each comparison has a complexity of $O(3^{p-2})$ (comparing two sequences with $|X_p|$ elements). Therefore, line 13 has a complexity of $O(m \cdot 3^{p-2})$. Thus, the complexity of lines 4 to 13 is $O((np+m)\cdot 3^{p-2})$. Note that $np < m$ (see Table 1). Hence, the complexity of lines 4 to 13 is $O(m \cdot 3^{p-2})$.

We note that $N = |X_{\text{countU}}| = O(m)$. In line 16, the complexity of creating set I is $O(m \cdot 3^{p-2})$ (perform m comparisons of sequences with length $|X_p|$). Line 17 has a complexity of $O(m)$ (since it involves creating a list of sequences from indices I with at most m elements). Therefore, the complexity of lines 15 to 17 is $O(m(m \cdot 3^{p-2} + m)) = O(3^{p-2} \cdot m^2)$.

In summary, Algorithm 3 has a complexity of $O(m^2 \cdot 3^{p-2}) = O(\frac{\lambda^{2n}}{n^2} \cdot 3^{p-2})$.

A4 Algorithm 4

According to Appendix A3, line 10 of Algorithm 4 has a complexity of $O(|\text{BranchCurrent}(i)|^2 \cdot 3^{p-2})$ for all $p \in \mathbb{N}$. The condition in line 13 and updating the set BranchNext in line 15 only have a complexity of $O(1)$ (no sequence comparisons are required as in Algorithm 3). Therefore, lines 12 to 15 have a complexity of

$$O(|X_{\text{branch2cluster}}|) = O(N(n, p)) = O(m),$$

since $|X_{\text{branch2cluster}}| = N(n, p)$ and the set $\mathcal{P}_n$ can be partitioned into at most m equivalence classes. Note that after each p-th iteration (the "repeat" loop), the BranchNext set only adds at most $|X_{\text{branch2cluster}}|/2$ elements (see lines 13–15). Therefore, in the $(p+1)$-th iteration, the number of elements in the BranchCurrent set is no more than $|X_{\text{branch2cluster}}|/2$, or equivalently $|\text{BranchCurrent}| \leq N(n,p)/2$. Thus, in the p-th iteration for $p \geq 2$, line 9 will perform at most $N(n,p-1)/2$ iterations. Additionally, in the p-th iteration for $p \geq 2$, we have $|\text{BranchCurrent}(i)| \leq N(n,p-1)$, for all $i = 1,\ldots,|\text{BranchCurrent}|$. Hence, for $p \geq 2$, lines 9 to 15 have a complexity of

$$\begin{aligned}&O(N(n,p-1)\cdot(|\text{BranchCurrent}(i)|^2\cdot 3^{p-2}+N(n,p)\cdot N(n,p-1))\\ &\quad = O(N(n,p-1)^2\cdot 3^{p-2}+N(n,p)\cdot N(n,p-1)).\end{aligned}$$

In particular, for $p = 1$ (the first "repeat" loop), line 10 has a complexity of $O(m^2)$ (according to Sect. A3), and lines 12–15 have a complexity of $O(N(n,1))$. Therefore, for $p = 1$, lines 9 to 15 have a complexity of $O(m^2 + N(n,1))$. By setting $N(n,0) = 1$, we can conclude that the complexity of lines 9 to 15 at the p-th iteration is $O(N(n,p-1)^2\cdot 3^{p-2}+N(n,p)\cdot N(n,p-1))$. Moreover, since $N(n,p) \leq m$, this implies that the complexity of lines 9 to 15 is $O(m^2\cdot 3^{p-2}+m^2) = O(m^2\cdot 3^{p-2})$ for all $p \geq 1$.

Since $p \leq n$, there are at most n "repeat" loops. Therefore, the complexity of Algorithm 4 is

$$O(m^2\cdot 3^{1-2}) + O(m^2\cdot 3^{2-2}) + \cdots + O(m^2\cdot 3^{n-2}) = O(m^2\cdot 3^n) = O\Big(\frac{\lambda^{2n}}{n^2}\cdot 3^n\Big).$$

In summary, from the above development, the total computational complexity of the proposed algorithm is

$$O((m^2+n!)3^n) = O\left(\Big(\frac{\lambda^{2n}}{n^2}+n!\Big)3^n\right),$$

where n is the period of sequences and $\lambda = \frac{3-\sqrt{5}}{2}$.

Acknowledgements The authors would like to thank the anonymous referees for careful reading the manuscript and for giving us suggestions which improve the quality of this paper. HMH thanks Prof. Markus Kunze for hospitality during his stay at Cologne University and for suggesting Gutkin/Osipov's method in [10]. A part of this paper was done during the stay of HMH and NBT at Vietnam Institute for Advanced Study in Mathematics (VIASM). They would like to thank VIASM for its wonderful working condition. This work is supported by Vietnam National Foundation for Science and Technology Development (Grant No. 101.01.2023.47).

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Altland, A., Braun, P., Haake, F., Heusler, S., Knieper, G., Müller, S.: Near action-degenerate periodic-orbit bunches: a skeleton of chaos. In: Janke, W., Pelster, A. (eds.) Path Integrals—New Trends and Perspectives: Proc 9th Int Conf (Dresden). World Scientific, Singapore (2008)
2. Apostol, T.M.: Introduction to Analytic Number Theory. Springer, New York (1976)
3. Arnold, V., Avez, A.: Ergodic Problems in Classical Mechanics. W.A. Benjamin Inc, New York (1968)
4. Barreira, L.: Ergodic Theory, Hyperbolic Dynamics and Dimension Theory. Springer, Berlin (2012)
5. Beck, M., Robins, S.: Computing the Continuous Discretely. Integer-Point Enumeration in Polyhedra, 2nd edn. Springer, New York (2015)
6. Berkolaiko, G.: Quantum Star Graphs and Related Systems. PhD Thesis, University of Bristol (2000)
7. Bohigas, O., Giannoni, M.J., Schmit, C.: Characterization of chaotic quantum spectra and universality of level fluctuation laws. Phys. Rev. Lett. **52**, 1–5 (1984)
8. Bowen, R.: Markov partitions for Axiom A diffeomorphisms. Am. J. Math. **95**(3), 725–747 (1970)

9. Gavish, U., Smilansky, U.: Degeneracies in the length spectra in metric graph. J. Phys. A Math. Theor. **40**, 10009–10020 (2007)
10. Gutkin, B., Osipov, V.A.: Clustering of periodic orbits in chaotic systems. Nonlinearity **2**, 177–199 (2012)
11. Gutkin, B., Osipov, V.A.: Clustering of periodic orbits and ensembles of truncated unitary matrices. J. Stat. Phys. **153**, 1049–1064 (2013)
12. Katok, A., Hasselblatt, B.: Introduction to Modern Theory of Dynamical Systems. Cambridge University Press, Cambridge (1995)
13. Lawrence, J.: Polytope volume computation. Math. Comput. **195**(57), 259–217 (1991)
14. Lind, D., Marcus, B.: An Introduction to Symbolic Dynamics and Coding. Cambridge University Press, Cambridge (1995)
15. Pollicott, M., Yuri, M.: Dynamical Systems and Ergodic Theory. Cambridge University Press, Cambridge (1998)
16. Sharp, R.: Degeneracy for length spectrum of metric graphs. Geom. Dedicata **149**, 177–188 (2010)
17. Sieber, M., Richter, K.: Correlations between periodic orbits and their rôle in spectral statistics. Phys. Scr. T **90**, 128–133 (2001)
18. Tanner, G.: Spectral statistics for unitary transfer matrices of binary graphs. J. Phys. A Math. Gen. **33**, 3567–3585 (2000)
19. Varvoglis, H.: Chaos, random walks and diffusion in Hamiltonian systems. In: Lega, E., Benest, Froeschle, C. (eds.) Hamiltonian Systems and Fourier Analysis, pp. 247–287. Cambridge Scientific Publishers, Cambridge (2005)