The maximum independent sets of de Bruijn graphs of diameter 3

Dustin A. Cartwright
Department of Mathematics
Yale University
P.O. Box 208283
New Haven, CT 06520-8283, USA
dustin.cartwright@yale.edu

María Angélica Cueto
Department of Mathematics
Columbia University
MC 4403
New York, NY 10027, USA
macueto@math.columbia.edu

Enrique A. Tobis †
Department of Ophthalmology, Children's Hospital Boston
Center for Brain Science and Swartz Center for Theoretical Neuroscience
Harvard University, 300 Longwood Avenue
Boston, MA 02115, USA
Enrique.Tobis@childrens.harvard.edu

Submitted: Oct 5, 2010; Accepted: Sep 23, 2011; Published: Oct 3, 2011
Mathematics Subject Classification: 05C69,05A15

Abstract

The nodes of the de Bruijn graph $B(d, 3)$ consist of all strings of length 3, taken from an alphabet of size d, with edges between words which are distinct substrings of a word of length 4. We give an inductive characterization of the maximum independent sets of the de Bruijn graphs $B(d, 3)$ and for the de Bruijn graph of diameter three with loops removed, for arbitrary alphabet size. We derive a recurrence relation and an exponential generating function for their number. This recurrence allows us to construct exponentially many comma-free codes of length 3 with maximal cardinality.

†Corresponding author.
D. A. Cartwright was supported by the National Science Foundation grants DMS-0354321, DMS-0456960, and DMS-0757236. M. A. Cueto was supported by a UC Berkeley Chancellor's Fellowship. E. A. Tobis was supported by a CONICET doctoral fellowship, CONICET PIP 5617, ANPCyT PICT 20569 and UBACyT X042 and X064 grants.

1 Introduction

For any positive integers d and D, the *de Bruijn* graph $B(d, D)$ is the directed graph whose d^D nodes consist of all the D-digit words from the alphabet $\{0, \dots, d-1\}$. There is a directed edge from a word $x = x_1 \dots x_D$ to $y = y_1 \dots y_D$ if and only if $x_2 \dots x_D = y_1 \dots y_{D-1}$. These graphs were introduced in [7], under the name of *T-nets*. Since then, de Bruijn graphs have been used in several contexts, notably as a network topology [2, 6, 13], and for building protein-binding microarrays [1].

We concern ourselves with the maximum independent sets of these graphs, previously studied in [11, 12]. The graph $B(d, D)$ contains d nodes of the form $x \dots x$, which have an edge to themselves. In a slight abuse of notation, we will refer to such a node as *the loop x*. Notice that a loop cannot be in any independent set of $B(d, D)$, and therefore we call these sets *loop-less maximum independent sets* (LMISs). The maximum independent sets of the subgraph of $B(d, D)$ obtained by removing the edges $x \dots x \to x \dots x$ are called *maximum independent sets* (MISs). Figure 1 depicts $B(3, 3)$ with an MIS highlighted.

A natural question to ask is what is the *stable size* of $B(d, D)$ for arbitrary d and D, i.e. the sizes of an MIS and a loop-less MIS. This question was studied in [12]. Lichiardopol defined $\alpha(d, D)$ to be the size of an MIS with loops and $\alpha^*(d, D)$ to be the size of a loop-less MIS [12]. For D a prime at least 3, he proved the inequalities

$$\alpha(d, D) \leqslant \frac{(D-1)(d^D - d)}{2D} + 1 \quad \text{and} \quad \alpha^*(d, D) \leqslant \frac{(D-1)(d^D - d)}{2D}. \tag{1}$$

He then showed that in fact, equality holds for D equal to 3, 5 or 7 and conjectured that the same is true for all odd primes D. He furthermore showed that, given a prime D, if equality holds in (1) for $d = 2$, then it holds for any d. As a byproduct of his work, we conclude that any MIS of $B(d, D)$ has at most two loops.

In the case of $D = 3$, we give a complete recursive characterization of the maximum independent sets of $B(d, 3)$. To do so, we give four functions which extend an MIS in $B(d, 3)$ to an MIS in $B(d+1, 3)$ or $B(d+2, 3)$ (Definitions 2.6, 2.8, 2.10 and 2.11). Our main result is that every maximum independent set in $B(d, 3)$ can be formed by beginning with an MIS in $B(1, 3)$ or $B(2, 3)$ and successively applying our four functions and permuting the alphabet. Moreover, since the sequence of functions and permutations is unique up to certain transpositions, we can compute the number of MISs, which corresponds to the Sloane sequence `A052608` [14]:

Theorem 4.5. If we let a_d be the number of maximum independent sets of $B(d, 3)$, then a_d has exponential generating function

$$\sum_{d=1}^{\infty} \frac{a_d t^d}{d!} = \frac{t + t^2}{1 - 2t - t^2}.$$

In addition, we prove that the number of loop-less maximum independent sets has the same generating function (Theorem 5.6).

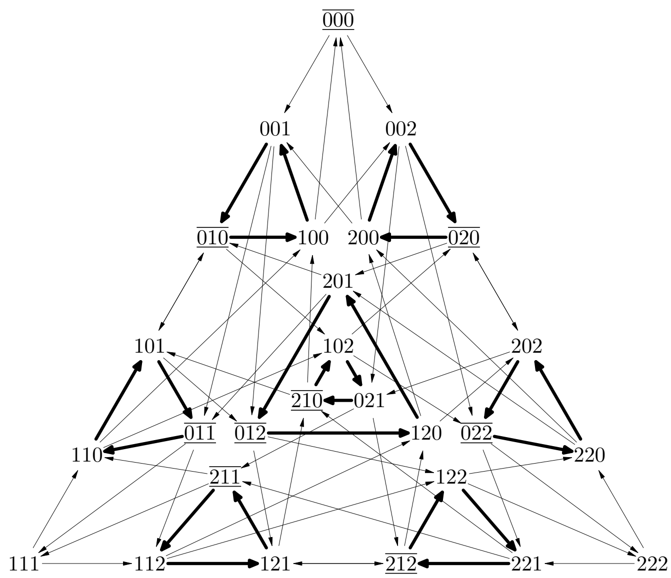

Figure 1: The de Bruijn graph $B(3,3)$, with the loops on 000, 111 and 222 removed. The highlighted nodes belong to one of the 42 possible MISs in $B(3,3)$. Bold arrows indicate edges under the shift function θ defined in (2).

A loop-less maximum independent set in $B(d,3)$ is a maximum comma-free code of length 3. Comma-free codes were introduced by Crick, Griffith, and Orgel as a hypothetical encoding of amino acid sequences in DNA [5], and further generalized in [9, 10, 15, 3]. A *comma-free code* is a set S of D-digit words such that if $x_1 \ldots x_D$ and $y_1 \ldots y_D$ are in S, then no substring of $x_2 \ldots x_D y_1 \ldots y_{D-1}$ is in S. In [9] and [8], it was shown that a comma-free code of length $D = 3$ could have as many as $(d^3 - d)/3$ elements by giving the same example (up to permuting the alphabet), namely the code consisting of all words $x_1x_2x_3$ such that $x_1 < x_2 \geqslant x_3$. In contrast to this single example, our results give an explicit construction of exponentially many equivalence classes of maximum comma-free codes (Theorem 5.7).

For $D = 2$ and $d \geqslant 4$, the maximum independent sets have size $\alpha(d,2) = \alpha^*(d,2) = \lfloor d^2/4 \rfloor$ [12, Prop. 5.1], and the same analysis as in that proof shows that number of maximum independent sets of $B(d,2)$ is $\binom{d}{d/2}$ if d is even and $2\binom{d}{(d-1)/2}$ if d is odd.

On the other hand, for $D > 3$, even small values of d yield de Bruijn graphs with a large number of maximum independent sets in $B(d,D)$. For example, using the computer algebra system CoCoA [4], we found out that there are 1 and 44 maximum independents sets of $B(1,5)$ and $B(2,5)$, respectively. However, we know that there are at least 210492 maximum independent sets of $B(3,5)$. This rapid growth means that the maximum independent sets in $B(3,5)$ cannot be produced from smaller independent sets using only permutations of the alphabet and a handful of functions.

We conjecture that an analogue of Theorem 4.5(1) for $D > 3$ would require starting with MISs in $B(d,D)$ for all $d < D$. Moreover, as it occurs for diameter three, we would also need functions taking a maximum independent set in $B(d,D)$ to one in $B(d+k,D)$

for all $k < D$. Finding explicit formulas for these functions would require knowledge of the sets $B(k, D)$ for $k < D$. To summarize:

Conjecture 1.1. *Let D be a fixed odd prime number. Then the exponential generating function of the number of maximum independent sets of $B(d, D)$ is the ratio of two polynomials, each of degree $D - 1$.*

The rest of this paper is organized as follows. In Section 2, we define two functions f and f' that take a maximum independent set of $B(d, 3)$ to a maximum independent sets of $B(d + 1, 3)$. Likewise, we construct another two functions g and g' that take a maximum independent set of $B(d, 3)$ to a maximum independent set of $B(d + 2, 3)$. In Section 3 we compute the stabilizers of the maximum independent sets produced by f, f', g and g' under the action of the symmetric group $\mathbb{S}_d$, and show that the functions take disjoint orbits to disjoint orbits. In Section 4, we prove our main theorems. In Section 5, we give a bijection between the maximum independent sets and the loop-less maximum independent sets of $B(d, 3)$, from which we conclude that their numbers coincide.

2 Inductive Construction of Maximum Independent Sets

In this section, we present two pairs of combinatorial operations that transform a maximum independent set in the de Bruijn graph $B(d, 3)$ into a maximum independent set in either $B(d + 1, 3)$ or $B(d + 2, 3)$.

Convention 2.1. Throughout this paper, we let $[\mathbf{d}]$ stand for the set $\{0, \ldots, d - 1\}$.

Essential to the structure of the de Bruijn graph $B(d, 3)$ are the cycles under the shift function θ, defined as

$$\theta\colon V(B(d, 3)) \to V(B(d, 3)) \qquad \theta(xyz) = yzx, \tag{2}$$

where $V(B(d, 3))$ denotes the set of nodes of the graph $B(d, 3)$ [12]. Note that the fixed points of θ are exactly the loops of $B(d, 3)$. On the other hand, if xyz is not a loop, then xyz, $\theta(xyz)$, and $\theta^2(xyz)$ form a directed 3-cycle. In Figure 1, the θ-cycles are indicated by bold edges.

Convention 2.2. Whenever we speak of *cycles*, we mean the cycles induced by θ.

The action of θ induces a decomposition of the nodes of $B(d, 3)$ into $(d^3 - d)/3$ cycles of length 3, and d cycles of length 1 (i.e. the loops). Each of these disjoint cycles contributes at most one node to any independent set of $B(d, 3)$.

The following proposition explains the role played by a loop in a maximum independent set of $B(d, D)$ and it shows that such a set can have at most two loops.

Proposition 2.3. *Let D be an odd prime number, and let S be a maximum independent set of $B(d,D)$ achieving the maximum possible size $\frac{(D-1)(d^D-d)}{2D}+1$. Then S contains one or two loops. Moreover, if a is a loop in S and x is any digit which is not a loop of S, then the node $(ax)^{\frac{D-1}{2}}a$ is in S. If S has two loops a,b, then, possibly after swapping a and b, $(ab)^{\frac{D-1}{2}}a$ is in S. Moreover, each cycle contributes exactly $(D-1)/2$ nodes to S, except for one of the form $b^{D-2i}(ab)^i$, for some $1 \leqslant i \leqslant (D-1)/2$, which only contributes $(D-1)/2-1$ nodes.*

Proof. The proof is contained in the proof of [12, Proposition 4.3]. □

From the previous result we see that the loops of an MIS play a special role. More precisely, if S is a maximum independent set then all the cycles of $B(d,D)$ of length D contribute at most $(D-1)/2$ elements to S. If D is an odd prime, and S has only one loop, then Lichiardopol's conjecture says that equality holds [12]. If, on the other hand, S contains two loops a,b, then (up to swapping a and b) we can assume that $(ab)^{\frac{D-1}{2}}a \in S$. Hence, all cycles of $B(d,D)$ of length D contribute at most $(D-1)/2$ nodes, except for one cycle with more b's than a's, which contributes at most $(D-1)/2-1$. Again, Lichiardopol's conjecture states that these maximal contributions are achieved [12].

Since the conjecture holds for $D=3$, every cycle (with the possible exception of the cycle of bab) contributes one element to any maximum independent set. This motivates the following definition, which will play an essential role in our inductive construction of maximum independent sets of $B(d,3)$.

Definition 2.4. Let A be a set of nodes from $B(d,3)$. Let x and y be two digits in $[\mathbf{d}]$. We say that *y appears between x in A* if the node xyx belongs to A. We define $\mathcal{M}_x(A)$ as the set of digits which *do not appear* between x in A. We define $m_x(A)$ as the number of digits which do not appear between x in A, i.e. $m_x(A) = |\mathcal{M}_x(A)|$.

Notation 2.5. If w is a node in $B(d,3)$, we will denote by $w[x \to y]$ the node that results from replacing every occurrence of the digit x by the digit y in w. We write $x \in w$ to mean that x is one of the digits that appear in w.

We denote by $L(S)$ the set of loops of a maximum independent set S. We denote by a the element of $L(S)$ such that $m_a(S)=0$. We will refer to it as *the distinguished loop*. If S has another loop we denote it b. This distinction will be extremely important for the construction of our four operations on $B(d,3)$.

We now define our first operation, sending a maximum independent set of $B(d,3)$ to a subset of $B(d+1,3)$. Proposition 2.7 will show that this subset is a maximum independent set.

Definition 2.6. Let S be a maximum independent set of $B(d,3)$. Following Notation 2.5, we define $f(S) \subset B(d+1,3)$ as the set $S \cup \bigcup_{i=1}^{5} U_i(S)$, where

$$U_1(S) = \{w[a \to d] \mid w \in S, a \in w, w \neq aaa, w \neq aba\},$$
$$U_2(S) = \{axd \mid x \in [\mathbf{d}]\backslash L(S)\}, \qquad U_3(S) = \{dxa \mid x \in [\mathbf{d}]\backslash L(S)\},$$
$$U_4(S) = \{udv \mid u,v \in L(S)\}, \qquad U_5(S) = \{udd \mid u \in L(S)\}.$$

Proposition 2.7. *If S is a maximum independent set of $B(d,3)$, then $f(S)$ is a maximum independent set of $B(d+1,3)$.*

Proof. By definition, $f(S)$ is made up of six disjoint sets. We will see that $f(S)$ is an independent set and that it has the right cardinality, as in (1). We start by showing that $f(S)$ is an independent set. This amounts to noticing that there are no arrows between the six sets defining $f(S)$. The only remark to bear in mind is that axa is in S for all x, and that bxb is also in S, except for $x = a$. We leave the details to the reader.

We now compute the cardinality of $f(S)$. Let l be the number of loops of S. We have $|S| = 1 + (d^3 - d)/3$, and

$$|U_1(S)| = (d-1)^2 - (l-1) + (d-l) = d^2 - d + 2 - 2l$$
$$|U_2(S)| = |U_3(S)| = (d-l), \qquad |U_4(S)| = l^2, \qquad |U_5(S)| = l.$$

Only the cardinality of $U_1(S)$ requires explanation. Notice that $B(d,3)$ has $(d-1)^2$ cycles whose nodes contain the digit a once. Each of these contributes one element to S and thus to $U_1(S)$, with the exception of $abb \to bba \to bab$ in the case that $l = 2$, that contributes no node to S nor $U_1(S)$. Likewise, S and $U_1(S)$ contain one element from each of the $d - l$ cycles of the form $aax \to axa \to xaa$, where x is not a loop. Hence, $|U_1(S)| = d^2 - d + 2 - 2l$.

We add the sizes of our six constituents, to obtain

$$|f(S)| = \frac{(d+1)^3 - (d+1)}{3} + 1 + (l-1)(l-2).$$

Since l is either 1 or 2 by Proposition 2.3, $f(S)$ has the size of an MIS in $B(d+1,3)$. □

We next define another function very similar to f and prove that it has analogous properties.

Definition 2.8. Let S be a maximum independent set of $B(d,3)$. We define $f'(S) \subset B(d+1,3)$ as the union of S, the sets $U_1(S)$, $U_2(S)$, $U_3(S)$, $U_4(S)$ from Definition 2.6, and $U_5'(S) = \{ddu \mid u \in L(S)\}$, which is the reverse of $U_5(S)$.

Proposition 2.9. *If S is a maximum independent set of $B(d,3)$, then $f'(S)$ is a maximum independent set of $B(d+1,3)$.*

Proof. This proposition is proved analogously to Proposition 2.7. □

We now define another pair of operators g and g'. These will send a maximum independent set of $B(d,3)$ to a maximum independent set of $B(d+2,3)$. As before, we follow the convention of Notation 2.5.

Definition 2.10. Let S be a maximum independent set of $B(d,3)$. We define $g(S) \subset B(d+2,3)$ to be the union $S \cup \bigcup_{i=1}^{8} V_i(S)$, where

$$\begin{aligned}
V_1(S) &= \{w[a \to y] \mid y \in \{d, d+1\}, w \in S, a \in w, w \neq aaa, w \neq aba\},\\
V_2(S) &= \{axy \mid x \in [\mathbf{d}] \backslash L(S), y \in \{d, d+1\}\},\\
V_3(S) &= \{yxa \mid x \in [\mathbf{d}] \backslash L(S), y \in \{d, d+1\}\},\\
V_4(S) &= \{yxz \mid y, z \in \{d, d+1\}, y \neq z, x \in [\mathbf{d}] \backslash L(S)\},\\
V_5(S) &= \{uyv \mid u, v \in L(S), y \in \{d, d+1\}\},\\
V_6(S) &= \{uyy \mid u \in L(S)\}, y \in \{d, d+1\}\},\\
V_7(S) &= \{yzu \mid y, z \in \{d, d+1\}, y \neq z, u \in L(S)\},\\
V_8(S) &= \{d(d+1)(d+1), (d+1)dd\}.
\end{aligned}$$

Definition 2.11. Let S be a maximum independent set of $B(d,3)$. We define $g'(S) \subset B(d+2,3)$ to be the union of S, the sets $V_1(S)$, $V_2(S)$, $V_3(S)$, $V_4(S)$, $V_5(S)$ from Definition 2.10, and the sets

$$\begin{aligned}
V_6'(S) &= \{yyu,\ u \in L(S), y \in \{d, d+1\}\},\\
V_7'(S) &= \{uyz,\ y, z \in \{d, d+1\}, y \neq z, u \in L(S)\},\\
V_8'(S) &= \{(d+1)(d+1)d, dd(d+1)\},
\end{aligned}$$

which are the reverses of $V_6(S)$, $V_7(S)$, and $V_8(S)$ respectively.

Proposition 2.12. *If S is a maximum independent set of $B(d,3)$, then $g(S)$ and $g'(S)$ are maximum independent sets of $B(d+2,3)$.*

Proof. We will prove the statement for the set $g(S)$. The result for $g'(S)$ can be proven analogously. The set $g(S)$ is made up of nine disjoint sets. By definition, it is easy to see that $g(S)$ is an independent set. We now show that it has the desired cardinality. We have $|S| = 1 + (d^3 - d)/3$. If l is the number of loops of S, then

$$\begin{aligned}
&|V_1(S)| = 2|U_1(S)| = 2(d^2 - d + 2 - 2l),\\
&|V_2(S)| = 2|U_2(S)| = 2|U_3(S)| = |V_3(S)| = 2(d - l),\\
&|V_4(S)| = 2|U_4(S)| = 2l^2, \qquad |V_5(S)| = 2(d - l),\\
&|V_6(S)| = 2|U_5(S)| = 2l, \qquad |V_7(S)| = 2l, \qquad |V_8(S)| = 2.
\end{aligned}$$

The sum of these sizes is $|g(S)| = \big((d+2)^3 - (d+2)\big)/3 + 1 + 2(l-1)(l-2)$. Since $l = 1$ or 2, the result follows. □

3 Action of the Symmetric Group on $B(d,3)$

In this section, we study the interaction between $\mathbb{S}_d$, the group of permutations of $[\mathbf{d}]$, and the four functions we defined in the previous section. In particular, we show that, up to a permutation of the digits, every maximum independent set in $B(d,3)$ can be obtained uniquely by successively composing our four operators and evaluating this new function at a maximum independent set of $B(1,3)$ or $B(2,3)$.

The group $\mathbb{S}_d$ acts on the nodes of $B(d,D)$ by $\sigma(x_1 \cdots x_D) = \sigma(x_1) \cdots \sigma(x_D)$ for $\sigma \in \mathbb{S}_d$. This action preserves the graph structure, and therefore permutes the maximum independent sets. We will write $A \sim B$ to mean A and B are two sets in the same orbit under the action of $\mathbb{S}_d$. Note that the functions f, f', g, and g' are defined so that if $A \sim B$, then $f(A) \sim f(B)$, etc. Therefore, each of these functions takes an $\mathbb{S}_d$-orbit of MISs to an $\mathbb{S}_{d+1}$- or $\mathbb{S}_{d+2}$-orbit of MISs.

Proposition 3.1. *Let S be a maximum independent set of $B(d,3)$. Let $H \subset \mathbb{S}_d$ and $H', H'' \subset \mathbb{S}_{d+1}$ be the stabilizers of S, $f(S)$ and $f'(S)$, respectively. Then $H = H' = H''$, where we identify H with its image under the inclusion $\mathbb{S}_d \hookrightarrow \mathbb{S}_{d+1}$.*

Proof. We only show the equality $H = H'$. The result for H and H'' will follow in much the same way. We know that $H \subseteq H'$, and we must prove the other inclusion. Let $\sigma \in H'$, and let $L(S)$ be the loops of S, with a the distinguished loop with $m_a(S) = 0$. The set of loops must be preserved by σ and moreover, by Proposition 2.3, σ fixes each loop. We want to show that $\sigma(d) = d$. Suppose that $\sigma(d) = z \neq d$ and $\sigma(x) = d$, for some $x \neq d$. Since x is not a loop, the node axd then belongs to the set $U_2(S)$ from Definition 2.6, and so to $f(S)$. That means that $\sigma(axd) = adz$ must be in $f(S)$. Since this word begins with a, and has d in the middle, it could only be in $U_4(S)$. But $z \notin L(S)$, and so $adz \notin U_4(S)$. Therefore, $\sigma(d) = d$.

Now, since $\sigma(d) = d$, σ is also an element of $\mathbb{S}_d$. Furthermore, it must be in the stabilizer of S. Otherwise, it should map a node of S into a node having a d. Since this is not possible, $\sigma \in H$. □

Proposition 3.2. *Let S be a maximum independent set of $B(d,3)$. Let $H \subset \mathbb{S}_d$ and $H', H'' \subset \mathbb{S}_{d+2}$ be the stabilizers of S, $g(S)$ and $g'(S)$, respectively. Let $\tau \in \mathbb{S}_{d+2}$ be the transposition interchanging d and $d+1$. Then*

$$H' = H'' = \langle \tau, H \rangle,$$

where, again, we identify H with its image in $\mathbb{S}_{d+2}$. Note that τ commutes with every element of H.

Proof. Again, we only show the equality $H' = \langle \tau, H \rangle$, since the statement involving H'' is analogous.

As in the proof of Proposition 3.1, we know that $\langle \tau, H \rangle \subseteq H'$. Now, let $\sigma \in H'$. Again, σ must preserve the set $L(S)$ of loops in $g(S)$, and by Proposition 2.3, σ in fact fixes each loop. We will show that either σ or $\tau\sigma$ fixes d and $d+1$. Let x, y, z and v be such that

$$x \overset{\sigma}{\longmapsto} d \overset{\sigma}{\longmapsto} y \qquad \text{and} \qquad z \overset{\sigma}{\longmapsto} d+1 \overset{\sigma}{\longmapsto} v.$$

We know that $x, y, z, v \notin L(S)$. Suppose that x is neither d nor $d+1$. Then we must have $dxa \in V_3(S)$ from Definition 2.10. The node $\sigma(dxa) = yda$ has to be in $g(S)$, but it can only be in $V_7(S)$. That means that $y = d+1$. Likewise, considering

$$\sigma((d+1)za) = v(d+1)a,$$

we have $v = d$. So $\sigma(d) = d+1$ and $\sigma(d+1) = d$. This contradicts our assumption about x, and implies that $x = d$ or $d+1$. Analogously, $z = d+1$ or d. That means that σ fixes d and $d+1$ or that it transposes them. Therefore, either σ or $\tau\sigma$ is in H, and so $\sigma \in \langle \tau, H \rangle$. □

We now show the precise way in which our functions and $\mathbb{S}_d$ interact.

Lemma 3.3. *Let S and S' be maximum independent sets of $B(d,3)$. Then $f(S) \not\sim f'(S')$ and $g(S) \not\sim g'(S')$.*

Proof. We first prove the result for f and f'. For contradiction, suppose that there is $\sigma \in \mathbb{S}_{d+1}$ such that $f(S) = \sigma f'(S')$. Let $L(S)$ and $L(S')$ be the loops of S and S'. By construction, we have $\sigma L(S') = \sigma L(f'(S')) = L(f(S)) = L(S)$. Call a and a' the distinguished loops of S and S'. By Proposition 2.3, we know that $\sigma(a') = a$.

Let $x \notin L(S')$ and $y \notin L(S)$ be such that $x \xmapsto{\sigma} d \xmapsto{\sigma} y$. Suppose that $y \neq d$. Then the node ayd is in $U_2(S)$, and hence in $f(S)$. Therefore, $\sigma^{-1}(ayd)$ must be in $f'(S')$. But $\sigma^{-1}(ayd) = a'dx$, which cannot be in any of the sets that make up $f'(S')$. This implies that $y = d$, hence $\sigma(d) = d$. In other words, σ lies in the image of $\mathbb{S}_d$ in $\mathbb{S}_{d+1}$, and so $\sigma f'(S') = f'(\sigma S')$. However, $f(S)$ has at least one element of the form udd, and $f'(\sigma S')$ has none, so $f(S) \not\sim f'(S')$.

The proof for g and g' is similar. Namely, suppose that there exists $\sigma \in \mathbb{S}_{d+2}$ such that $g(S) = \sigma g'(S')$. Let $x, z \notin L(S')$, $y, v \notin L(S)$ be such that

$$x \xmapsto{\sigma} d \xmapsto{\sigma} y \qquad \text{and} \qquad z \xmapsto{\sigma} d+1 \xmapsto{\sigma} v.$$

Suppose that $y \neq d, d+1$. Then the node ayd is in $V_2(S)$, and therefore in $g(S)$. That means that $\sigma^{-1}(ayd) = a'dx$ must be in $g'(S')$. But such a node does not belong to any of the sets that make up $g'(S')$. This implies that either $\sigma(d) = d$ or $\sigma(d) = d+1$. Analogously, we can prove that $\sigma(d+1) = d+1$ or $\sigma(d+1) = d$.

Therefore, σ transposes d and $d+1$ or leaves them fixed. By Proposition 3.2, the transposition $(d, d+1)$ is in the stabilizer of $g'(S')$ and so by possibly multiplying σ on the right by this transposition, we can assume that σ fixes d and $d+1$ and so it lies in $\mathbb{S}_d$. Therefore, $\sigma g'(S') = g'(\sigma S')$, but $g(S)$ has at least one node of the form udd, and $g'(\sigma S')$ has none, so $g(S) \not\sim g'(S')$. □

We now state two invariants that completely characterize maximum independent sets of $B(d,3)$. This is useful to prove that our functions f, f', g, and g', together with the action of $\mathbb{S}_d$, allow us to construct all maximum independent sets of $B(d,3)$. In order to reverse these functions, we make the following observation, which also holds for loop-less maximum independent sets. Since we will use it in Section 5 we state it in full generality.

Proposition 3.4. *Let S be a (possibly loop-less) maximum independent set of $B(d,3)$, with loops $L(S)$. Let d' be any integer such that $c < d' < d$ for all $c \in L(S)$. Then, $S' = S \cap B(d',3)$ is a maximum independent set of $B(d',3)$ with loops $L(S)$.*

Proof. Since $B(d',3)$ is a subgraph of $B(d,3)$, S' is clearly an independent set. Furthermore, since S has one element from each cycle except possibly a cycle that only uses the digits a and b, then S' has the same property. Therefore, S' has the cardinality of a maximum independent set. □

Proposition 3.5. *Let S be a maximum independent set of $B(d,3)$ with l loops, where d is at least 3. There exists a digit x such that $m_x(S) = l+1$* if and only if *there exist $\sigma \in \mathbb{S}_d$ and S' a maximum independent set of $B(d-1,3)$ such that $S = \sigma f(S')$ or $S = \sigma f'(S')$.*

Proof. The reverse implication follows from the definitions of f and f', taking $x = \sigma(d-1)$. Conversely, suppose that there is an x with $m_x(S) = l+1$. We know it is not a loop by Proposition 2.3. We define the transposition $\sigma = (d-1, x)$ and the set $S' = \sigma S \cap B(d-1,3)$, which is a maximum independent set of $B(d-1,3)$ by Proposition 3.4.

Let a denote the distinguished loop of S. We know that the node $xax \notin S$. Therefore, either xxa or axx must be in S. Suppose that $axx \in S$, in which case we claim that $S = \sigma f(S')$.

We now consider each of the sets that make up $\sigma f(S')$, and show that they are included in S. The nodes of $\sigma S'$ belong to S, by definition of S'. Let us consider the nodes of $\sigma U_1(S')$. The nodes of this set are of the form xyx, xyy, yyx, xyz or yzx, for y and z distinct from x and $y, z \notin L(S)$.

- The nodes of the form xyx are all in S by the hypothesis on x.
- If $xyy \in \sigma U_1(S')$, then $ayy \in S'$. This means that $ayy \in S$, and so yyx cannot be in S. The node yxy cannot be in S either, since xyx is. So, $xyy \in S$. Analogously, if $yyx \in \sigma U_1(S')$, then $yyx \in S$.
- If $xyz \in \sigma U_1(S')$, then $ayz \in S'$ and $ayz \in S$. Since neither zxy (adjacent to xyx) nor yzx (adjacent to ayz) can be in S, xyz must be in S. The same reasoning applies to yzx.

Let us consider the nodes of $\sigma U_2(S')$. These have the form ayx. The nodes yxa (adjacent to xyx) and xay (adjacent to aya) cannot be in S, which implies that $ayx \in S$. The same reasoning shows that $\sigma U_3(S') \subset S$.

A node from $\sigma U_4(S')$ is of the form uxv, with u and v loops. The nodes xuv (adjacent to uxu) and uvx (adjacent to vxv) cannot be in S. Therefore, $uxv \in S$, and $\sigma U_4(S') \subset S$.

Finally, we know that $axx \in S$ or $xxa \in S$. Assume the first case. If S has a single loop, we have that $\sigma U_5(S') \subset S$. If S has an extra loop b, the nodes xbx (adjacent to bxb) and xxb (adjacent to axx) cannot be in S. That implies that $bxx \in S$, which means $\sigma U_5(S') \subset S$. This proves that $S \supseteq \sigma f(S')$. Since both sets have the same cardinality, we conclude that equality holds.

On the other hand, if $xxa \in S$, an analogous procedure shows that $S = \sigma f'(S')$. □

The following lemma is used in the proof of Proposition 3.7, which is the analogue of Proposition 3.5 for the operators g and g'. Note that in preparation for our study of loop-less maximum independent sets in Section 5, we prove Lemma 3.6 for loop-less maximum independent sets as well.

Lemma 3.6. *Let S be a (possibly loop-less) maximum independent set of $B(d,3)$, with $d \geqslant 3$. If there exist two different digits y and z, which are not loops, such that*

$$m_y(S) = m_z(S) = l + 2,$$

then $yzy \notin S$ and $zyz \notin S$.

Proof. Suppose that $yzy \in S$. Then, by the assumptions on $m_y(S)$, there must be some $v \neq y$ such that $yvy \notin S$. Suppose that $vyy \in S$. The node zyz cannot be in S, and by the assumption on $m_z(S)$, $zvz \in S$. Therefore, the nodes zvy (adjacent to vyy), vyz (adjacent to yzy) and yzv (adjacent to zvz) are not in S. But then the cycle $zvy \to vyz \to yzv$ contributes no nodes to S, which contradicts the fact that S has maximum cardinality. If we assume that $yyv \in S$, then the cycle $yvz \to vzy \to zyv$ cannot contribute any node to S, a contradiction.

In conclusion, our assumption that yzy is in S is inconsistent with S being a maximum independent set. By symmetry, the same holds if we assume $zyz \in S$. □

Proposition 3.7. *Let S be a maximum independent set of $B(d,3)$, $d \geqslant 3$, with l loops ($l = 1$ or 2). Then, there are two different digits y and z such that*

$$m_y(S) = m_z(S) = l + 2$$

and no *digit x such that $m_x(S) = l+1$,* if and only if *there exist $\sigma \in \mathbb{S}_d$ and S' a maximum independent set of $B(d-2,3)$ such that*

$$S = \sigma g(S') \quad or \quad S = \sigma g'(S').$$

Proof. One implication follows from the construction of g and g' taking $y = \sigma(d-1)$ and $z = \sigma(d-2)$. The proof in the other direction is analogous to the proof of Proposition 3.5. We can safely assume that $y = d-1$ and $z = d-2$. By Lemma 3.6, either the pair $(d-1)(d-2)(d-2)$ and $(d-2)(d-1)(d-1)$ are in S, or the pair $(d-1)(d-1)(d-2)$ and $(d-2)(d-2)(d-1)$ are in S. In the former case, we find that there is an S' such that $S = \sigma g(S')$. In the latter case, we find that $S = \sigma g'(S')$. □

Corollary 3.8. *Let S and S' be maximum independent sets of $B(d-1,3)$ and $B(d-2,3)$ with $d \geqslant 3$. Then for $\mathcal{F} = f, f'$ and $\mathcal{G} = g, g'$, we have $\mathcal{F}(S) \not\sim \mathcal{G}(S')$.*

Proof. This result follows from the invariants of $\mathcal{F}(S)$ and $\mathcal{G}(S')$ that are stated in Propositions 3.5 and 3.7. □

This corollary, together with Lemmas 3.3 and 3.3, shows that all four functions produce essentially different (i.e. in different $\mathbb{S}_d$-orbits) maximum independent sets.

4 Characterization of Maximum Independent Sets

In this section, we show that the functions f, f', g, and g', together with the action of $\mathbb{S}_d$ are sufficient to construct every maximum independent set of $B(d,3)$. For the rest of this section, L will denote the set of loops of S, and l will denote the cardinality of L. In Section 5, we will work with loop-less maximum independent sets. For that reason, we prove some of the results of this section in that context too.

As we mentioned in Section 2, the sets $\mathcal{M}_x(S)$ from Definition 2.4 play a key role. We start our discussion with two technical lemmas about them.

Lemma 4.1. *Let S be a (possibly loop-less) maximum independent set of $B(d,3)$. There cannot be three different digits x, y, and z, with $x,y,z \notin L$, such that*

$$\begin{aligned} \mathcal{M}_x(S) &= \mathcal{M}_y(S) = L \cup \{x,y,z\}, \\ \mathcal{M}_z(S) &= L \cup \{x,z\} \text{ or } L \cup \{x,y,z\}. \end{aligned} \tag{3}$$

Proof. Suppose that S is a maximum independent set and x, y, and z satisfy (3). Without loss of generality, we can assume that x, y, z, and the loops are smaller than $l+3$. Then $S' = S \cap B(l+3,3)$ is a maximum independent set in $B(l+3,3)$ by Proposition 3.4 with $\mathcal{M}_x(S') = \mathcal{M}_x(S)$, $\mathcal{M}_y(S') = \mathcal{M}_y(S)$, and $\mathcal{M}_z(S') = \mathcal{M}_z(S)$.

Without loss of generality we may assume that $xyy, yxx, xzz, zyy, zxx \in S'$ since yxy, $xyx, zxz, yzy, xzx \notin S'$. But this implies that there is no element of the cycle containing zyx in S', a contradiction. Therefore, no such S exists. □

Lemma 4.2. *Let S be a (possibly loop-less) maximum independent set of $B(d,3)$. There cannot be three different digits x, y, and z, none of which are loops, such that*

$$m_x(S) = m_y(S) = m_z(S) = l+2.$$

Proof. We prove the result by contradiction. Suppose there are such x, y and z. We know that $L \cup \{x\} \subset \mathcal{M}_x(S)$ and $|\mathcal{M}_x(S)| = l+2$. Therefore, at least one of y and z must appear between x. An analogous statement holds for y and z. Without loss of generality, suppose that y appears between x. Then xyx (adjacent to xyx) is not in S, which forces z to appear between y. That, in turn, forces x to appear between z. That is, the nodes xyx, yzy and zxz are in S. But then, none of the nodes $xyz \to yzx \to zxy$ are in S, contradicting the maximality of S. □

Remark 4.3. Note that a maximum independent set S of $B(d,3)$ with l loops can have at most one digit satisfying $m_x(S) = l+1$. If there were two, say x and y, then xyx and yxy would have to be in S, a contradiction.

The next proposition shows that, up to permutation, any maximum independent set lies in the image of one of our four operations.

Proposition 4.4. *Let S be a (possibly loop-less) maximum independent set of $B(d,3)$ with $d \geqslant 3$. Suppose there is no digit z such that $m_z(S) = l+1$. Then, there must be exactly two digits x and y such that $m_x(S) = m_y(S) = l+2$. Moreover, $\mathcal{M}_x(S) = \mathcal{M}_y(S) = L \cup \{x,y\}$.*

Proof. We just need to show that $m_x(S) = m_y(S) = l + 2$. Lemma 3.6 implies that $\mathcal{M}_x(S) = \mathcal{M}_y(S) = L \cup \{x, y\}$. By reordering the digits, we can assume that $m_{d-1}(S) \leqslant m_{d-2}(S) \leqslant m_i(S)$ for all $i < d-2$. By hypothesis, we know that $m_{d-1}(S) \geqslant l+2$ and we want to prove that $m_{d-2}(S) = l+2$. Lemma 4.2 will then imply that $d-2$ and $d-1$ are the only digits with this property.

We prove that $m_{d-2}(S) = l + 2$ by induction on d. Our base cases are $d \leqslant l + 3$. If $d = l + 1$, then the unique z not in L satisfies $m_z(S) = l + 1$, which contradicts our hypothesis. If $d = l+2$, and x and y are not in L, then $m_x(S) \geqslant l+2$ implies $m_x(S) = l+2$, and likewise for y. If $d = l + 3$, then Lemma 4.1 gives us the result.

Now, let d be greater than $l + 3$ and consider $S' = S \cap B(d-1, 3)$. By the inductive hypothesis, we must have one of two possibilities:

Case 1: S' has exactly one digit z with $m_z(S') = l + 1$. If $z = d - 2$, we are done. Suppose that $z \neq d-2$. By Remark 4.3, $m_{d-2}(S') > m_z(S')$, $m_{d-2}(S) \leqslant m_z(S)$ and $m_{d-2}(S') \leqslant m_{d-2}(S)$. Thus, we must have $m_z(S') = m_z(S) - 1$ and $m_{d-2}(S') = m_{d-2}(S)$. This means that $z(d-1)z$ is not in S and $(d-2)(d-1)(d-2) \in S$, which implies that $m_{d-2}(S) = l + 2$, as we wanted to show.

Case 2: S' has exactly two digits x and y with $m_x(S') = m_y(S') = l + 2$. We split this situation in two subcases.

Case 2.1: We suppose $x, y \neq d-2$. By an argument similar to that of Case 1, we know that $\mathcal{M}_x(S) = \mathcal{M}_y(S) = L \cup \{x, y, d-1\}$ and

$$m_{d-2}(S) = l + 3, \qquad \mathcal{M}_{d-2}(S) \supseteq L \cup \{d-2, d-1, x, y\},$$

which is a contradiction.

Case 2.2: Either x or y equals $d - 2$. Suppose $y = d - 2$. Since $m_{d-2}(S') = l + 2$, then $m_{d-2}(S) = l + 2$ (and we are done) or $m_{d-2}(S) = m_x(S) = l + 3$. Hence,

$$\mathcal{M}_x(S) = \mathcal{M}_{d-2}(S) = L \cup \{x, d-2, d-1\}. \tag{4}$$

Since $m_{d-1}(S) \leqslant m_{d-2}(S) = l + 3$, we have that

$$\mathcal{M}_{d-1}(S) = L \cup \{d-1, u, v\} \text{ or } L \cup \{d-1, u\}.$$

Case 2.2.1: Suppose $m_{d-1}(S) = l + 3$. We will show that $\mathcal{M}_{d-1}(S) = L \cup \{x, d-2, d-1\}$, which, together with (4), contradicts Lemma 4.1.

Assume $u, v \neq d-2$. That means that $(d-1)(d-2)(d-1) \in S$. Since $u \neq v$, we can assume without loss of generality that $u \neq x$. Then $xux \in S$ and $(d-2)u(d-2) \in S$. The nodes $(d-1)u(d-2)$ and $(d-2)u(d-1)$ must be in S, because the rest of the nodes in their cycles are adjacent to something just shown to be in S. We know that $(d-1)u(d-1)$ is not in S, because of the definition of u. Additionally, the nodes $(d-1)(d-1)u$ and $u(d-1)(d-1)$ are adjacent to the nodes we just showed are in S. Therefore, neither of them belong to S, a contradiction. Hence, one of u and v must equal $d - 2$, and so we have

$$\mathcal{M}_{d-1}(S) = L \cup \{u, d-2, d-1\}.$$

To finish, we need to prove that $u = x$. Assume the contrary. Then xux and $(d-1)x(d-1)$ are in S. Therefore, by inspecting their cycles we see that both $xu(d-1)$ and $(d-1)ux$ must be in S. On the other hand, either $u(d-1)(d-1) \in S$ or $(d-1)(d-1)u \in S$. However, $u(d-1)(d-1) \in S$ implies $xu(d-1) \notin S$, and $(d-1)(d-1)u \in S$ implies $(d-1)ux \notin S$. Therefore, $u = x$.

Case 2.2.2 Suppose $m_{d-1}(S) = l + 2$. If we assume x and $d-2$ are not in $\mathcal{M}_{d-1}(S)$ and proceed as in the previous case, we get a contradiction. Therefore, Lemma 4.1 applied to x, $d-1$ and $d-2$ leads to a contradiction. □

We now state our main result.

Theorem 4.5 (Characterization of the Maximum Independent Sets of $\mathbf{B(d,3)}$). *For all positive d we have:*

1. *Any orbit of independent sets of $B(d,3)$ under the action of $\mathbb{S}_d$ is obtained from the $\{000\}$ and the orbit of $\{000, 010, 111\}$ under $\mathbb{S}_2$ by a* unique *sequence of applications of f, f', g, and g'.*

2. *Let S be an MIS of $B(d,3)$. Then the subgroup of $\mathbb{S}_d$ stabilizing S is generated by disjoint transpositions. In particular, the cardinality of the stabilizer of S is a power of 2.*

3. *Let $b_{d,k}$ be the number of orbits of MISs in $B(d,3)$ whose elements have stabilizers of size 2^k. Then we have the recurrence relation*

$$\begin{cases} b_{1,0} = 1,\ b_{2,0} = 3, \\ b_{d,k} = 2b_{d-1,k} + 2b_{d-2,k-1} \quad \text{for } d \geqslant 3, \end{cases}$$

and the generating function

$$\sum_{d=1}^{\infty} \sum_{k=0}^{\infty} b_{d,k} t^d s^k = \frac{t + t^2}{1 - 2t - 2t^2 s}.$$

4. *The number a_d of maximum independent sets of $B(d,3)$ satisfies*

$$\begin{cases} a_1 = 1,\ a_2 = 6, \\ a_d = 2da_{d-1} + d(d-1)a_{d-2} \quad \text{for } d \geqslant 3, \end{cases}$$

and has exponential generating function

$$\sum_{d=1}^{\infty} \frac{a_d t^d}{d!} = \frac{t + t^2}{1 - 2t - t^2}.$$

Proof. For $d = 1$, the only maximum independent set of $B(1, 3)$ consists of the unique node $\{000\}$. For the case of $d = 2$, it can be checked manually that the three orbits of maximum independent sets under $\mathbb{S}_2$ are the orbits of $\{000, 010, 011\}$, $\{000, 010, 110\}$, and $\{000, 010, 111\}$. Note that the first two of these are $f(\{000\})$ and $f'(\{000\})$ respectively. Thus, the existence statement in (1) follows from Propositions 3.5, 3.7 and 4.4. The uniqueness comes from Lemma 3.3 and Corollary 3.8.

The statements in (2) and (3) follow from the previous result and the description of the stabilizers in Propositions 3.1 and 3.2. Finally, the generating function in (4) is obtained by substituting $s = 1/2$ into the previous generating function, because

$$a_d = \sum_{k=0}^{\infty} \frac{d! b_{d,k}}{2^k}.$$

The recurrence follows immediately. □

The following table lists the values of $b_{d,k}$, for all $d \leqslant 6$.

$k \backslash d$	1	2	3	4	5	6
0	1 (1,0)	3 (2,1)	6 (4,2)	12 (8,4)	24 (16,8)	48 (32,16)
1			2 (2,0)	10 (8,2)	32 (24,8)	88 (64,24)
2					4 (4,0)	28 (24,4)

In each entry, the first number indicates the number of orbits whose elements have only one loop. The second one is the number of orbits with two loops.

5 Loop-less Maximum Independent Sets

In this section, we analyze the number of *loop-less* maximum independent sets (LMISs) of $B(d, 3)$, for all d. Recall from the introduction that the size of an LMIS of $B(d, 3)$ is

$$\alpha^*(d, 3) = \frac{d^3 - d}{3} = \alpha(d, 3) - 1.$$

By MIS, we will continue to mean a maximum independent set *with* loops. As in previous sections, we let a be the loop of S such that $m_a(S) = 0$, and the other loop (if there is one) is denoted by b.

In what follows, we provide an explicit bijection between LMISs and MISs of $B(d, 3)$.

Definition 5.1. Let S be a maximum independent set of $B(d, 3)$, $d \geqslant 3$. We define

$$h(S) = \begin{cases} S \backslash \{aaa\} & \text{if } S \text{ has only one loop,} \\ S \backslash \{aaa, bbb, aba\} \cup \{aab, bba\} & \text{if } S \text{ has two loops } a < b, \\ S \backslash \{aaa, bbb, aba\} \cup \{baa, abb\} & \text{if } S \text{ has two loops } a > b. \end{cases} \tag{5}$$

Proposition 5.2. *Let S be a maximum independent set of $B(d,3)$. Then $h(S)$ is an LMIS of $B(d,3)$.*

Proof. Let S be an MIS of $B(d,3)$. If S has only one loop, then eliminating it leaves us with an independent set of the correct size.

If S has two loops, say a and b, then $h(S)$ is a set of the correct size, since the nodes we added were not already present in S. However, we must see that $h(S)$ is an independent set. Assume $a < b$. Suppose we have a node adjacent to aab. Then it is of the form abx or xaa. Since bxb and aaa are in S, then abx and xaa cannot be in S. A similar argument show that adding bba to S preserves independence. Therefore, the nodes we add are not adjacent to any other nodes in the construction, and the result follows. The case $a > b$ is proved analogously. □

Proposition 5.3. *The function h is injective.*

Proof. Let S and S' be two different MISs of $B(d,3)$. Then showing that $h(S) \neq h(S')$ is just a matter of analyzing all the possible combinations of loops and their relative order in S and S'. We leave the details to the reader. □

Lemma 5.4. *Let S be a maximum independent set with two loops a and b. Let τ be the transposition of a and b. Let $S' = S \backslash \{aaa, bbb, aba\}$. Then $S' = \tau S'$.*

Proof. We must show that for every node $w \in S'$ such that $a \in w$, we have $w[a \to b] \in S'$ and vice versa. Notice that any node of S' cannot contain a and b simultaneously. The nodes that contain two a's or two b's are axa and bxb, and they are in S' for all $x \neq a, b$. Thus, $xay \notin S'$ for all $x, y \neq a$.

The nodes that contain only one a are xya or axy for $x, y \neq a$. If $xya \in S'$, then $bxy \notin S'$, and so xyb must be in S' in order to have one element from its cycle. We can prove that $axy \in S'$ implies $bxy \in S'$ in a similar way. □

Proposition 5.5. *The function h is surjective.*

Proof. Let S be an LMIS of $B(d,3)$. By Proposition 4.4, we have two possibilities:

First, if there is a digit x such that $m_x(S) = 1$, then there is no node of the form xxy or yxx. Therefore, $S' = S \cup \{xxx\}$ is an MIS of $B(d,3)$, and $S = h(S')$.

Second, if there are two digits x and y such that $m_x(S) = m_y(S) = 2$, then we have either $xxy, yyx \in S$ or $yxx, xyy \in S$. In the first case, we construct

$$S' = S \cup \{xxx, yyy, xyx\} \backslash \{xxy, yyx\}.$$

If $x < y$, then $S = h(S')$. If $x > y$, then by Lemma 5.4, $S = h(\tau S')$, where τ is the transposition of x and y. The remaining case is dealt with analogously. □

Theorem 5.6. *Let a_d^* be the number of loop-less maximum independent sets of $B(d,3)$. Then $a_d^* = a_d$.*

Proof. This follows from Propositions 5.3 and 5.5. □

We conclude with a result that links comma-free codes and loop-less maximum independent sets.

Theorem 5.7. *Every loop-less maximum independent set is a maximum comma-free code of length* 3*. In particular, the number of equivalence classes of comma-free codes in an alphabet of size d is at least* 2^d*, where equivalence means equivalence under the action of* $\mathbb{S}_d$*.*

Proof. If S is an LMIS with $x_1x_2x_3$ and $y_1y_2y_3$ elements of S, then $x_2x_3y_1$ cannot be an element of S because it is adjacent to $x_1x_2x_3$ in $B(d,3)$. Likewise, $x_3y_1y_2$ cannot be in S because it is adjacent to $y_1y_2y_3$. Therefore, S is a comma-free code.

For the second statement, by considering only the first term of the recurrence relation in Theorem 4.5(4), we see that $a_d \geqslant 2^d d!$. Therefore, the number of maximum comma-free codes is at least $2^d d!$, so the number of equivalence classes under the action of $\mathbb{S}_d$ must be at least 2^d. □

The set $S = \{100, 110\}$ is an example of a maximum comma-free code which is not an independent set for $d = 2$.

Acknowledgements

We would like to thank Verónica Becher for introducing de Bruijn graphs to us. We also thank Alicia Dickenstein for her useful comments. We are indebted to Yukio Shibata, for kindly sharing the work of his group with us. We used the free computer algebra system CoCoA [4] for the initial computations which suggested our results.

References

[1] M. Berger and M. Bulyk, *Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors*, Nat Prot **4** (2009), 393–411.

[2] J.-C. Bermond and C. Peyrat, *De Bruijn and Kautz networks: a competitor for the hypercube?*, Proceedings of the 1st European Workshop on Hypercubes and Distributed Computers, Rennes (F. André and J-P. Verjus, eds.), North Holland, 1989, pp. 279–293.

[3] Alexander L. Churchill, *Restrictions and generalizations on comma-free codes*, Electron. J. Combin. **16** (2009), no. 1, Research Paper 25, 15. MR 2482093 (2010h:94156)

[4] CoCoATeam, CoCoA*: a system for doing Computations in Commutative Algebra*, Available at `http://cocoa.dima.unige.it`.

[5] F. H. C. Crick, J. S. Griffth, and L. E. Orgel, *Codes without commas*, P Natl Acad Sci USA **43** (1957), no. 5, 416–421.

[6] A. Datta, S. Girdzijauskas, and K. Aberer, *On de Bruijn routing in distributed hash tables: There and back again*, IEEE International Conference on Peer-to-Peer Computing (2004), 159–166.

[7] N. G. de Bruijn, *A combinatorial problem*, K Ned Akad Van Wet **49** (1946), 758–764.

[8] W. L. Eastman, *On the construction of comma-free codes*, IEEE Trans. on Information Theory **11** (1965), no. 2, 263–266.

[9] S. W. Golomb, B. Gordon, and L. R. Welch, *Comma-free codes*, Canadian J Math **10** (1958), no. 2, 202–209.

[10] B. H. Jiggs, *Recent results in comma-free codes*, Canadian J Math **15** (1963), 178–187.

[11] Y. Kikuchi and Y. Shibata, *On the independent set of de Bruijn graphs*, Topics in Applied and Theoretical Mathematics and Computer Science, WSEAS Press, 2001, pp. 117–128.

[12] N. Lichiardopol, *Independence number of de Bruijn graphs*, Discrete Math **306** (2006), no. 12, 1145–1160.

[13] B. Mukherjee, *Optical communication networks*, Series on Computer Communications, McGraw-Hill, New York, 1997.

[14] N. J. A. Sloane, `http://oeis.org/A052608`.

[15] B. Tang, S. W. Golomb, and R. L. Graham, *A new result on comma-free codes of even word-length*, Canadian J Math **39** (1987), no. 3, 513–526.