Abstract. De Bruijn cycles of some order $k$ are sequences that contain every substring of length $k$ over a finite alphabet exactly once. It is well known that the family of de Bruijn sequences has a super-exponential size in terms of the order of the sequences. A de Bruijn sequence can be transformed into another de Bruijn sequence by applying the cross-join operation. In this paper we show that any de Bruijn sequence can be cross-joined several times to reach any other de Bruijn sequence. This was done for binary sequences by Mykkeltveit et al. [14] using the theory of Boolean functions. Using graph theoretic methods, we generalize this result to Hamiltonian cycles of the so-called generalized de Bruijn digraph. This already includes regular de Bruijn sequences, binary and non-binary, as a special case.

1. Introduction

Consider an alphabet $A$ of size $d$. A de Bruijn cycle of order $k$ over the alphabet $A$ is a periodic sequence of characters of $A$ such that, within one period, every possible string of size $k$ occurs exactly once as a substring. For example 00110 and 0022120110 are de Bruijn sequences of order 2 over the alphabets $\{0, 1\}$ and $\{0, 1, 2\}$ respectively. These sequences were popularized by de Bruijn [1] and Good [9] even though their existence was established much earlier, see Flye-Sainte Marie [6] and Martin [12]. Beyond the mere proof of existence for any $d$ and $k$, de Bruijn [1] and Good [9] established that the number of de Bruijn cycles is $(d!)^{d^{k-1}}/d^k$.

de Bruijn sequences play a pivotal role in coding theory and cryptography as they are the main building blocks of many stream ciphers. The binary case is especially useful, although there has been much research on sequences with non-binary alphabets. Linear sequences, i.e. linear feedback shift register sequences of maximal length, are known not to be safe and hence not useful for such applications, see Massey [13], for example. de Bruijn sequences based on nonlinear feedback functions are far more useful and more numerous than linear ones. They

*Key words and phrases.* de Bruijn Sequence, generalized de Bruijn graph, Hamiltonian cycle, universal cycle, nonlinear feedback shift register, generation of combinatorial objects.
are, however, far from being mathematically understood as a class. There are many references in the literature that introduce and study properties of subclasses of de Bruijn sequences that are nonlinearly constructed. Golomb [8] initiated this study in his pioneering work, but there are newer methods that have been introduced ever since Golomb first published his book in 1967. Such methods are typically combinatorial or graph-theoretic producing one de Bruijn sequences or a collection of similar sequences. Fredricksen [7] is an excellent review article that outlines, and often details, many of these known methods and algorithms. Much more research on nonlinear sequences has been done in the past two decades. Some recent publications are [5], [18], [19]. Another breed of nonlinear sequences are those produced by efficient successor rules for both binary and non-binary alphabets as in [16] and [17].

Among the methods of generating de Bruijn sequences is the cross-join technique, which starts with a de Bruijn sequence and interchanges two appropriately chosen pairs of vertices to obtain another de Bruijn sequence. The aim of this paper is to show that any de Bruijn sequence of any alphabet size can be transformed via a sequence of cross-joins into any other de Bruijn sequence of the same order. In the binary case, this result was recently proven by Mikkeiltveit and Szmidt [14]. Our method applies in the binary case as well. Even though our main objective was to generalize the result to non-binary de Bruijn sequences, it turns out that our method of proof applies without modification to a generalized version of de Bruijn sequences where the number of vertices need not be a pure power of some alphabet size $d$. So we chose to give our proof in the language of the latter, in order to widen the scope of the result as well as to pin-point to the actual assumption that is sufficient for the proof. We explain the set up of generalized de Bruijn digraphs in Section 3. In Section 4 we formulate and prove our main results and we round up in Section 5. In the next section we give some background material.

2. Preliminaries

There are several known methods to generate de Bruijn sequences. The one method that is most understood is the algebraic method that we outline next. First we assume that the alphabet $A$ is residue ring $\mathbb{Z}_d$. A construction is achieved if we have an initial state $(s_1, \ldots, s_k)$ and a rule that provides the next symbol $a_{k+1}$ of the alphabet $A$ when the current state is $(a_1, \ldots, a_k)$. In this case the next state is $(a_2, \ldots, a_k, a_{k+1})$. Thus, the next state is determined by a recurrence relation of order $k$ and of the form

$$a_i = f(a_{i-k}, \ldots, a_{i-1})$$

where $i \geq k + 1$ and the feedback function $f$ maps $A^k$ to $A$. Since the number of possible states is finite, the recurrence equation gives a periodic sequence. de Bruijn sequences correspond to feedback functions with the maximal period of $d^k$. The sequence is called linear when the recurrence function is linear and homogeneous.
When a linear homogeneous recurrence function is applied to an initial string of zeros we obtain the constant zero sequence. Therefore, the maximal possible period of a linear feedback function is \( d^k - 1 \). A sequence constructed this way is one that misses the all-zero string of size \( k \). Such a sequence is called a maximal length linear feedback shift register (LFSR), or sometimes a “punctured” linear de Bruijn sequence, because a full de Bruijn sequence can then be obtained by appending a zero next to any of the \( d - 1 \) occurrences of \( k - 1 \) consecutive zeros.

We will concern ourselves here with a method called the cross-join method, which starts with a de Bruijn cycle of some order and produces another de Bruijn cycle of the same order. This is done by first dividing the first cycle into two disjoint cycles by swapping the successors of two appropriately chosen vertices, and then rejoining these two cycles by swapping the successors of another pair of vertices, one residing on each cycle. Clearly, the success of this method relies on locating a pair of vertices whose successors can be swapped. Once this cross step is done, one needs to locate another pair of vertices, one on each disjoint cycle, and swap their successors into order to connect the two cycles into a new de Bruijn cycle. Such pairs of vertices are called conjugate vertices. One can immediately see that two vertices are conjugate if they have the form \((x_1, x_2, \ldots, x_k)\) and \((\hat{x}_1, x_2, \ldots, x_k)\) for \( x_1 \neq \hat{x}_1 \). Similarly, the predecessors of two vertices can be swapped if they have the form \((x_1, \ldots, x_{k-1}, x_k)\) and \((x_1, \ldots, x_{k-1}, \hat{x}_k)\) for \( x_k \neq \hat{x}_k \). These vertices are referred to as companion vertices.

Two pairs of vertices that allow to transform a de Bruijn cycle \( u \) to another de Bruijn cycle \( v \) are called cross-join pairs. In the binary case, each vertex has exactly one conjugate, so it is enough to determine one vertex from each pair and the resulting pair is called a cross-join pair. The idea of cross-joining a binary de Bruijn sequence into another has been appealing to many investigators, especially because it is always possible to generate a linear de Bruijn sequence (an LFSR sequence) and cross-join it to another punctured nonlinear de Bruijn sequence. Chang et al. [2] conjectured that any two binary maximal length LFSR sequences of the same span \( k \) have the same number of cross-join pairs that only depends on \( k \), and therefore the same number of punctured de Bruijn sequences can be made from each LFSR by using a single cross-join operation. This common number is \( \frac{(2^k-1)(2^{k-1}-2)}{6} \). The conjecture was proven by Helleseth and Klöve [10]. The author of this paper was not able to find any information in the literature on the number of direct cross-join neighbors of a de Bruijn sequence when it is not based on a maximal period LFSR. Using computation to inspect this number of neighbors for low order binary de Bruijn sequences, we find that many distinct values exist. Indeed, the 16 binary de Bruijn sequences of order 4 are equally divided between sequences with 7 cross-join neighbors and sequences with 10 cross-join neighbors. The case of binary sequences of order 5 is quite different. Table 2 reports the
possible number of neighbors along with the frequency of vertices that have this number of neighbors. Notice that the formula of Chang et al [2] is 7 and 35 respectively for \( k = 4 \) and \( 5 \), while the corresponding numbers of maximal LFSR sequences are 2 and 15 so that many nonlinear de Bruijn sequences share this number of neighbors with the LFSR sequences.

Recently, Mykkeltveit and Szmidt [14] settled the following question in the affirmative. “Is it possible to obtain any binary de Bruijn sequence by applying a sequence of cross-join pair operations to a given binary de Bruijn sequence?” They mentioned that this is a several-decade-old question that was recently asked at the International Workshop on Coding and Cryptography 2013 in Bergen. They even claim this result as an explanation of the origins of nonlinear Boolean functions that yield de Bruijn cycles. This is indeed the case because once a cross-join pair is identified, the feedback function of the initial de Bruijn sequence can easily be altered to give the feedback function of the new sequence.

As mentioned in the introduction, the main theme of this paper is to give a generalization of the main result of [14]. Rather than generalizing to regular de Bruijn cycles with non-binary alphabets, we formulate and prove a more general result for the set of Hamiltonian cycles in the so-called \((d, N)\) generalized de Bruijn digraphs, in which the number of vertices is any positive integer \( N \geq 2 \) and which boils down to a regular de Bruijn digraph of order \( k \) when \( N = d^k \).

| \( n \) | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| \( f \) | 188 | 152 | 240 | 272 | 136 | 208 | 16 | 176 | 64 | 48 | 16 | 40 | 0 | 112 | 0 | 136 |

**Table 1.** The frequency \( f \) of binary de Bruijn sequences of order 5 with \( n \) cross-join neighbors for all possible values of \( n \)

3. Generalized de Bruijn digraphs

A de Bruijn sequence is interchangeably called a de Bruijn cycle because it can be seen as a Hamiltonian cycle on the de Bruijn digraph. A de Bruijn digraph with alphabet \( A = \{0, 1, \ldots, d-1\} \) has \( d^k \) vertices which can be taken as the set of \( d^k \) vectors \((x_1, \ldots, x_k)\). For two vertices \( \mathbf{a} = (a_1, \ldots, a_k) \) and \( \mathbf{b} = (b_1, \ldots, b_k) \), there exists an edge connecting \( \mathbf{a} \) to \( \mathbf{b} \) if and only if \( b_i = a_{i+1} \) for \( i = 1, \ldots, k-1 \). When the vertices are regarded as decimal numbers represented in base \( d \), the set of edges consists of all pairs \((x, y)\) (or interchangeably \( x \rightarrow y \)) where \( x \) and \( y \) are integers in \( \{0, \ldots, d^k - 1\} \) and \( y = dx + r \) (mod \( d^k \)), \( r = 0, 1, \ldots, d-1 \).

In a generalized de Bruijn digraph, \( d^k \) is replaced by any integer \( N > d \). This digraph was introduced to dispense with the restrictive number of vertices in ordinary de Bruijn digraphs. Formally, a generalized de Bruijn digraph \( G_B(d, N) = (V, E) \)
where the vertex set is $V = \{0, 1, \ldots, N - 1\}$ and $(x, y)$ is contained in the edge set $E$ if and only if $y = dx + r \mod N$ for some $r \in \{0, \ldots, d - 1\}$.

This digraph preserves many of the properties of ordinary de Bruijn digraphs. As shown below, it is a regular digraph. Imase and Itoh \[11\], Reddy, Pradhan and Kuhl \[15\] prove that $G_B(d, N)$ has a very short diameter just like de Bruijn digraphs. They also show that $G_B(d, N)$ is strongly connected. Indeed, Du and Hwang \[4\] show that when $\gcd(d, N) > 1$, $G_B(d, N)$ is Hamiltonian. In the rest of the paper the following notation will be used. If $(x, y) \in E$, we say that $y$ is a successor of $x$ and that $x$ is a predecessor of $y$. The set of possible successors of a vertex $x$ is denoted by $\Gamma^+_x$ while the set of predecessors is $\Gamma^-_x$. Two vertices $x_1$ and $x_2$ are said to be conjugate vertices if there exist two vertices $y_1$ and $y_2$ such that $(x_1, y_1)$, $(x_1, y_2)$, $(x_2, y_1)$ and $(x_2, y_2)$ are all edges in $G_B(d, N)$. In this case we also say that $y_1$ and $y_2$ are companion vertices. A path is a sequence of vertices $x_1, \ldots, x_k$ such that $(x_i, x_{i+1})$ is an edge for all $i = 1, \ldots, k - 1$. A path is simple if all of its vertices are distinct. A cycle is a path in which the first and last vertices coincide. A cycle is simple if, except for the first and last, all of its vertices are distinct. A Hamiltonian cycle is a simple cycle that includes all the vertices of the digraph. Since $G_B(d, N)$ generalizes de Bruijn digraphs, we are going to refer to a Hamiltonian cycle of $G_B(d, N)$ simply as a de Bruijn cycle. When $N = d^k$ for some integers $d > 1$ and $k \geq 1$, $G_B(d, N)$ reduces to the regular $d$-ary de Bruijn digraph of order $k$ which we denote by $B(d, k)$. Figures (1) and (2) illustrate regular and generalized de Bruijn digraphs.

**Definition 3.1.** Two de Bruijn cycles are called adjacent if it is possible to cross-join one into another via a single cross-join operation.

This allows for the following definition.

**Definition 3.2.** The cross-join graph $C(d, N)$ of $G_B(d, N)$ has the set of all de Bruijn sequences as the set of vertices. There is an edge between two vertices if they are adjacent in the sense of definition 3.1. In particular, $C(d, d^k)$ is the cross-join graph of a regular de Bruijn digraph of order $k$.

**Lemma 3.3.** $G_B(d, N)$ is $d$-regular. That is, each vertex has $d$ successors and $d$ predecessors.

**Proof.** We only need to show that each vertex in $G_B(d, N)$ has exactly $d$ predecessors. Let $\gcd(d, N) = \delta$. Given a vertex $y$ between 0 and $N - 1$, we need to count the number of solutions to the equations

$$y \equiv dx + r \mod N,$$

where $r$ can be $0, 1, \ldots, d - 1$. Equivalently, $dx \equiv (y - r) \mod N$. When $y - r$ is not a multiple of $\delta$ there obviously is no solution. Suppose that $y - r = \delta t$ for some $t$ in $\{0, 1, \ldots, N/\delta - 1\}$. The equation implies that $(d/\delta)x \equiv t \mod N/\delta$. 


Figure 1. de Bruijn digraphs $B(3,2)$ (left) and $B(2,3)$ (right).

Since $\gcd(d/\delta, N/\delta) = 1$, the last equation admits a unique solution modulo $N/\delta$ and therefore it has $\delta$ solutions modulo $N$. The lemma now follows, as there are exactly $d/\delta$ multiples of $\delta$ in the $d$ consecutive integers $y, y-1, \ldots, y-d+1$ for each fixed $y$. \hfill \Box

**Lemma 3.4.** Suppose that $d$ divides $N$, then for any vertex $y \in V$, $\Gamma^-_y = \{ \lfloor y/d \rfloor + t \cdot N/d : t = 0, \ldots, d-1 \}$.

**Proof.** Since $y/d - 1 < \lfloor y/d \rfloor \leq y/d$, we have

$$y - d < d\lfloor y/d \rfloor \leq y,$$

so that $y = d\lfloor y/d \rfloor + r$ for some $r \in \{0, \ldots, d-1\}$. Hence, $\lfloor y/d \rfloor$ is a predecessor of $y$, and clearly so is $\lfloor y/d \rfloor + t \cdot N/d$ for each $t = 1, \ldots, d-1$. By lemma 3.3, these are the only predecessors of $y$. \hfill \Box

If follows that, when $d$ divides $N$, for each vertex $x \in \{0, \ldots, N/d - 1\}$ and each $t = 1, \ldots, d-1$, $\Gamma^+_x$. We formulate this in the following form that is more usable below.
Figure 2. generalized de Bruijn digraphs $G_B(2, 6)$ (left), $G_B(2, 12)$ (middle), and $G_B(3, 10)$ (right).

Lemma 3.5. Suppose that $d$ divides $N$. Let $y_1$ and $y_2$ be two successors of $x_1$. Suppose that $y_1$ is the successor of some $x_2 \neq x_1$. Then $y_2$ is also a successor of $x_2$.

The reason we require that $d$ divides $N$ in the above two lemmas is precisely to require any two conjugate vertices $x_1$ and $x_2$ to be completely conjugates, in the sense that $\Gamma^+_{x_1} = \Gamma^+_{x_2}$. In fact, when $d$ does not divide $N$, the conjugacy relation is not transitive. That is, it is possible to find vertices $x_1, x_2, x_3$ such that $x_1, x_2$ is a conjugate pair, $x_2, x_3$ is a conjugate pair but $x_1, x_3$ is not a conjugate pair. As an illustration, we list in the table below all the edges of $G_B(d, N)$ for $d = 4$ and three values of $N$ with different divisibility conditions. The edges are listed in teh form $x \rightarrow y_1, \ldots, y_{d-1}$, meaning that $(x, y_i)$ is an edge for each $i$. In Case (a)
note that 0 and 2 are conjugates (as they both have an edge to 0 and 1), 2 and 4 are conjugates but 0 and 4 are not conjugates. Similar intransitive vertices can be found in (b). Only in (c), where \(d|N\), conjugate vertices form an equivalent class and so they are completely conjugate.

To find the set of conjugates of a vertex \(x_0\) let \(x\) be such a conjugate. There must exist some \(y\) in \(\{0, \ldots, N-2\}\) such that \((x_0, y), (x_0, y+1), (x, y), (x, y+1)\) are all edges in \(G_B(d, N)\). Therefore, \(y = dx + r_1 \equiv dx_0 + r_2 \mod N\), or

\[
d(x-x_0) \equiv (r_2 - r_1) \mod N,
\]

for some \(r_1\) and \(r_2\) in \(\{0, \ldots, d-2\}\). It is evident that this equation has a solution in \(x\) only if \(r_2 - r_1\) is a multiple of \(\delta = \gcd(d, N)\), say \(r_2 - r_1 = \delta t\). The equation in the display above becomes \(x = x_0 + (d/\delta)^{-1}t \mod N/\delta\), where the inverse is taken modulo \(N/\delta\). We have thus proved

**Lemma 3.6.** The conjugates of \(x_0\) are of the form

\[
x_0 + (d/\delta)^{-1}t + q \cdot N/\delta,
\]

where \(q \in \{0, \ldots, \delta-1\}\) and \(t\) is any integer such that \(-(d-2)/\delta \leq t \leq (d-2)/\delta\).

Note that when \(d\) divides \(N\), \(t\) can only be zero, so that the conjugates of \(x_0\) are of the form \(x_0 + q \cdot N/d\), for \(q = 0, \cdots, d-1\).

4. **Connectedness**

In this section we formulate and prove our main result. We will first define a metric distance between de Bruijn sequences. To this end let us align de Bruijn sequences so that they all start with the same vertex. Without loss of generality we choose this initial vertex to be 0. In the following, vertices of the cross-join graph, i.e., de Bruijn sequences, are denoted by \(u, v\), etc. while vertices of \(G_B(d, N)\) are denoted by \(x, y\), etc.
Definition 4.1. Let $u$ and $v$ be two vertices of $C(d, N)$ aligned as in the previous paragraph. The distance $D(u, v)$ is defined as $N - L$, where $L$ is the length of the longest initial path that is common to $u$ and $v$.

It is straightforward to verify that $D(u, v)$ satisfies the three axioms of a distance. Also, it is essential to keep in mind that this distance is not related to the distance defined by the graph adjacency on $C(d, N)$. As an example, for $d = 3$ and $N = 10$, the following are two de Bruijn sequences aligned to start at 0, with the last 0 repeated to stress that the sequences cycle back to the initial vertex.

$$u = (0, 2, 7, 1, 5, 6, 9, 8, 4, 3, 0)$$
$$v = (0, 2, 7, 1, 4, 3, 9, 8, 5, 6, 0)$$

The maximum common initial path is $(0, 2, 7, 1)$ so $D(u, v) = 10 - 4 = 6$. The following lemma is fundamental for the rest of the paper.

Lemma 4.2. Let $u$ and $v$ be two distinct de Bruijn sequences in $G_B(d, N)$ where $d$ divides $N$. Then there exists a de Bruijn sequence $u_1$ which is a neighbor of $u$ in $C(d, N)$ such that $D(u_1, v) < D(u, v)$.

We prove this lemma after we state and prove our main result.

Theorem 4.3. When $d$ divides $N$, the cross-join graph $C(d, N)$ is connected.

Proof. Let $u$ and $v$ be two distinct vertices in $C(d, N)$. By Lemma 4.2 $u$ has a neighbor $u_1$ on $C(d, N)$ such that $D(u_1, v) < D(u, v)$. If $u_1 = v$ then we are done, otherwise the same argument can be iterated to get a vertex $u_2$, which is a neighbor of $u_1$, with $D(u_2, v) < D(u_1, v)$. Due to the strict inequality, and since the number of vertices of $C(d, N)$ is finite, it is evident that this iterative process must end at $v$ after a finite number $m$, leading to the desired path $u_0 = u, u_1, \ldots, u_m = v$. □

Proof of Lemma 4.2. As a road map, we are going to state and prove three claims within the proof, the main one is Claim 1, while Claim 2 and Claim 3 are stated and proved within the proof of Claim 1.

Let $M_0$ be the maximal common initial sequence of $u$ and $v$. That is, suppose that the sequence

$$M_0 : 0 = x_1 \rightarrow x_2 \rightarrow \cdots \rightarrow x_{L_0}$$

is common to $u$ and $v$ and $L_0$ is maximal. Since $u \neq v$, $L_0 < N$ and so the successors of $x_{L_0}$ in $u$ and $v$ are both distinct from 0. Let us refer to these successors respectively as $x^{(1)}$ and $x_{L_0+1}$. Since $u$ is a de Bruijn sequence, it contains every vertex of $G_B(d, N)$ so it must contain $x_{L_0+1}$. The latter is evidently one of the vertices of $M_0$, the complement of $M_0$ in $u$; that is, the sub-path of $u$ that starts with $x^{(1)}$ and ends with the vertex just before 0. Let $x_0$ be the predecessor of $x_{L_0+1}$ in $u$. Since $x_{L_0+1}$ belongs to $M_0$, the vertex $x_0$ is either in $M_0$ or it is $x_{L_0}$ itself. But the latter is not possible because otherwise the common initial path of $u$ and $v$ would extend to $x_{L_0+1}$, defying the maximality of $M_0$. Now $x_{L_0}$ and
\(x_0\) are predecessors of the same vertex. They must be conjugate by Lemma 3.5. Swapping their successors we split \(u\) into two cycles, a cycle \(C_1\) that includes the vertex 0 and another cycle \(\tilde{C}_1\) that includes the edge \(x_0 \to x^{(1)}\).

The cycle \(C_1\), aligned to start at 0, and the de Bruijn cycle \(v\) have a maximal common initial sequence

\[ M_1 : 0 = x_1 \to x_2 \to \cdots \to x_L \to \cdots \to x_{L_1} \]

where \(L_1 \geq L_0 + 1\). The rest of the proof depends on establishing the following

Claim 1: It is possible to join \(C_1\) and \(\tilde{C}_1\) by using vertices in \(\complement M_1\), the complement of \(M_1\) in \(C_1\).

To show this suppose we cannot. Then let the successors of \(x_{L_1}\) in \(v\) and \(C_1\) be \(x_{L_1+1}\) and \(x^{(2)}\) respectively. Obviously, \(x_{L_1+1}\) is not on the path \(M_1\). Since every possible vertex is either on \(C_1\) or on \(\tilde{C}_1\), it follows that \(x_{L_1+1}\) is on on \(M_1\), the complement of \(M_1\) in \(C_1\), as it cannot be on \(\tilde{C}_1\), by our assumption. Let \(x_1\) be the predecessor of \(x_{L_1+1}\) in \(C_1\). Similarly to the previous paragraph, we can argue that \(x_1\) is in \(\complement M_1\).

Interchanging the successors of \(x_{L_1}\) and \(x_1\) we further split the cycle \(C_1\) into two cycles \(C_2\) and \(\tilde{C}_2\) with \(C_2\) being the cycle that includes 0 and shares a larger still initial path with \(v\), say,

\[ M_2 : 0 = x_1 \to x_2 \to \cdots \to x_{L_2}, L_2 > L_1. \]

In essence, this process can be iterated, without using vertices from \(\tilde{C}_1\), only a finite number of times. Let \(k\) be the maximal number of iterations and let \(C_k\) be the resulting cycle that includes the vertex 0 with maximal initial path

\[ M_k : 0 = x_1 \to x_2 \to \cdots \to x_{L_k}, L_k > L_{k-1} \]

that is common with the de Bruijn sequence \(v\).

Claim 2: The sub-path of the cycle \(C_k\) that begins with \(x_{L_k}\) and ends with 0 is simply an edge \((x_{L_k}, 0)\). That is, there is no vertex in \(C_k\) between \(x_{L_k}\) and 0.

To see this, suppose that \(x^{(k+1)} \neq 0\) is the successor of \(x_{L_k}\) in \(C_k\). Let \(x_{L_k+1}\) be the successor of \(x_{L_k}\) in \(v\), so that \(x_{L_k+1}\) and \(x^{(k+1)}\) are companion vertices. We then see that \(x_{L_k+1} \neq 0\) since otherwise the de Bruijn cycle \(v\) would be shorter than \(C_k\). Since \(C_1\) and \(\tilde{C}_1\) include all vertices, \(x_{L_k+1}\) is either in \(\tilde{C}_1\) or it is in the part \(\complement M_1\) of \(C_1\). If the first case is true, swapping the predecessor of \(x_{L_k+1}\) in \(\tilde{C}_1\) with the predecessor of \(x^{(k+1)}\) (which is evidently one of the vertices of \(\complement M_1\)) shows that \(C_1\) and \(\tilde{C}_1\) can be joined into a de Bruijn sequence using a vertex outside \(M_1\), contradicting the original assumption of Claim 1.

If the second case is true, that is, if \(x_{L_k+1}\) belongs to \(\complement M_k\) or any of the cycles made by the previous iteration and that are at most \(C_2, \ldots, C_{k-1}, \tilde{C}_2, \ldots, \tilde{C}_{k-1}\) (equivalently, if it is one of the vertices of \(\complement M_1\)), then we can swap the predecessors...
of $x^{(k+1)}$ and $x_{L_k+1}$ to get yet another cycle $C_{k+1}$ that shares a longer initial segment with $v$, contradicting the maximality of $C_k$.

It follows that $x_{L_k}$ is a predecessor of 0. We next prove

Claim 3: $C_k$ includes all predecessors of 0.

We prove this claim in a way similar to the proof of the previous claim. In effect, suppose that $y$ is a predecessor of 0 that is not on $C_k$. If $y$ belongs to $\bar{C}_1$ we get a contradiction because we could have joined $C_1$ and $\bar{C}_1$ by swapping the successors of $y$ and $x_{L_k}$ (which is on $\bar{M}_1$). Likewise, the presence of $y$ on any of the intermediate cycles $C_2, \ldots, C_{k-1}, \bar{C}_2, \ldots, \bar{C}_{k-1}$ contradicts the maximality of $k$.

The validity of this last claim means that the sequence $M_k$ cannot be continued into a de Bruijn sequence as it cannot cycle back to 0 without using one of the predecessors of 0 for a second time. This of course is not true because $M_k$ is already an initial path of the de Bruijn sequence $v$.

We have thus proven that $C_1$ and $\bar{C}_1$ can be joined by swapping the successor of a vertex in $\bar{M}_1$ with that of a conjugate vertex in $\bar{C}_1$. This makes a new de Bruijn sequence $u_1$ which is a neighbor of $u$ on $C(d, N)$. Since $L_0 < L_1, N - L_1 < N - L_0$, and $u_1$ satisfies the inequality

$$D(u_1, v) < D(u,v)$$

as desired.

5. Conclusion

We have proven that the cross-join graph that corresponds to the generalized de Bruijn digraph $G_B(d, N)$ with an arbitrary number of vertices $N$ is connected when $d$ divides $N$. This of course includes the class of all regular de Bruijn sequences of arbitrary alphabet size $d$. The usefulness of this result lies in the fact that a linearly generated sequence can be cross joined several times to obtain virtually any nonlinear sequence.

As the proof of the fundamental Lemma 4.2 depends on the fact that the conjugacy of vertices of $G_B(d, N)$ is an equivalence relation, the results of this paper are still open when $d$ does not divide $N$.

References

[1] N. G. de Bruijn, A Combinatorial Problem, Koninklijke Nederlandse Akademie v. Wetenschappen 49, pp. 758764, (1946).
[2] T. Chang, I. Song, and S. H. Cho, Some properties of cross-join pairs in maximum length linear sequences. Proc. ISZTA 90, Honolulu, Hawaii, pp. 1077- 1079, (1990).
[3] P. Dabrowski, G. Labuzek, T. Rachwalik, J. Szmidt, Searching for nonlinear feedback shift registers with parallel computing. IACR Cryptology ePrint Archive 2013/542. 2013 Military Communications and Information Systems Conference. MCC 2013, Malto, France.
[4] D.Z. Du, F.K. Hwang, Generalized de Bruijn Digraphs, Networks, vol. 18, pp. 27-38, (1988).
[5] E. Dubrova, A scalable method for constructing Galois NLFSRs with period $2^n-1$ using cross-join pairs. IEEE Trans. on Inform. Theory, 59(1), pp. 703-709, (2013).
[6] C. Flye-Sainte Marie, Solution to problem number 58, l’Intermédiaire des Mathématiciens, vol. 1, pp. 107-110, (1894).
[7] H. Fredricksen, A Survey of Full Length Nonlinear Shift Register Cycle Algorithms. SIAM Review, Vol. 24, No. 2, pp. 195-221, (1982).
[8] S. Golomb, Shift register sequences. San Fransisco, Holden-Day, (1967), revised edition, Laguna Hills, CA, Aegean Park Press, 1982.
[9] I. J. Good, Normal recurring decimals. Journal of the London Mathematical Society 21 (3) pp. 167-169, (1946).
[10] T. Helleseth, T. Klöve, The Number of Cross-join pairs in maximum length linear sequences. IEEE Trans. on Inform. Theory, 31, pp. 1731-1733, (1991).
[11] M. Imase, M. Itoh, A Design for Directed Graphs with Minimum Diameter. IEEE Transactions on Computers. Vol. C-32, No. 8, August 1983.
[12] M. H. Martin, A Problem in Arrangements. Bulletin of the American Mathematical Society, 40, 859-864, (1934).
[13] J. L. Massey, Shift-register synthesis and BCH decoding. IEEE Trans. Information Theory, IT-15 (1): 122-127 (1969).
[14] J. Mykkeltveit, J. Szmidt, On Cross Joining de Bruijn Sequences. Contemporary Mathematics, 63, pp.335-346, (2015).
[15] S. M. Reddy, D. K. Pradhan and J. G. Kuhl, Direct Graphs with Minimum Diameter and Maximal Connectivity. School of Engineering, Oakland University Tech. Rep., July 1980.
[16] J. Sawada, A. Williams, and D. Wong. A Surprisingly Simple de Bruijn Sequence Construction. Discrete Math., 339, pp 127-131, (2016).
[17] J. Sawada, A. Williams, and D. Wong. A Simple Shift Rule for k-ary de Bruijn sequences. Discrete Math., 340, pp 524-531, (2017).
[18] J. Szmidt, Nonlinear feedback shift registers and Zech logarithms, arXiv:1710.09556v2, (2017).
[19] M.S. Turan, On the nonlinearity of maximum-length NFSR feedbacks. Cryptography and Communications, 4(3-4), pp. 233-243, (2012).