Appeared in *Discrete Applied Mathematics*, Vol. 37, pp. 421-436, 1992.

Asymptotically-Tight Bounds on the Number of Cycles in Generalized de Bruijn-Good Graphs

Ueli M. Maurer

Institute for Signal and Information Processing
Swiss Federal Institute of Technology
CH-8092 Zürich, Switzerland

Abstract. The number of cycles of length k that can be generated by q-ary n-stage feedback shift-registers is studied. This problem is equivalent to finding the number of cycles of length k in the natural generalization, from binary to q-ary digits, of the so-called de Bruijn-Good graphs [2, 7]. The number of cycles of length k in the q-ary graph $G_n^{(q)}$ of order n is denoted by $\beta^{(q)}(n,k)$. Known results about $\beta^{(2)}(n,k)$ are summarized and extensive new numerical data is presented. Lower and upper bounds on $\beta^{(q)}(n,k)$ are derived showing that, for large k, virtually all q-ary cycles of length k are contained in $G_n^{(q)}$ for $n > 2\log_q k$, but virtually none of these cycles is contained in $G_n^{(q)}$ for $n < 2\log_q k - 2\log_q \log_q k$. More precisely, if $\nu_k^{(q)}$ denotes the total number of q-ary length k cycles, then for any function $f(k)$ that grows without bounds as $k \to \infty$ (e.g. $f(k) = \log_q \log_q \log_q k$), the bounds obtained on $\beta^{(q)}(n,k)$ are asymptotically tight in the sense that they imply

$$\begin{aligned}
\lim_{k\to\infty} \frac{\beta^{(q)}(n(k),k)}{\nu_k^{(q)}} &= 0 \quad \text{for } n(k) = \lfloor 2\log_q k - 2\log_q \log_q k - f(k) \rfloor, \quad \text{and} \\
\lim_{k\to\infty} \frac{\beta^{(q)}(n(k),k)}{\nu_k^{(q)}} &= 1 \quad \text{for } n(k) = \lfloor 2\log_q k + f(k) \rfloor,
\end{aligned}$$

where $\lfloor . \rfloor$ denotes the integer part of the enclosed number. Finally, some approximations for $\beta^{(q)}(n,k)$ are given that make the global behavior of $\beta^{(q)}(n,k)$ more transparent.

1. Introduction

The n-th order q-ary de Bruijn-Good graph $G_n^{(q)}$ [7, 11], sometimes in the literature also simply called de Bruijn graph, is the graph of all states and all possible state transitions of a q-ary n-stage feedback shift-register [12]. $G_n^{(q)}$ is hence a directed graph on q^n vertices labelled with q-ary n-tuples $(b_0, b_1, \ldots, b_{n-1})$, $b_i \in \{0, \ldots, q-1\}$, with q^{n+1} directed edges such that each vertex $(b_0, b_1, \ldots, b_{n-1})$ has q edges directed out to $(b_1, \ldots, b_{n-1}, x)$, for $x \in \{0, \ldots, q-1\}$, and q edges directed in from $(y, b_0, \ldots, b_{n-2})$, for $y \in \{0, \ldots, q-1\}$. The de Bruijn-Good graphs are not only of theoretical interest in graph theory and combinatorics, but investigating them seems to be of practical use as well. The property that the number of edges leaving and entering a vertex is constant for all vertices is of interest when the de Bruijn-Good graphs are considered as interconnection networks. Another interesting property is that the diameter of $G_n^{(q)}$ is minimal and equal to n, i.e., there exists a (directed) path of length at most n from every vertex to every other vertex. The powerful structure of the de Bruijn-Good graphs, together with the fact that they admit simple routing strategies, suggest that they might be useful for the solution of certain interconnection problems arising in communication networks and multiprocessor systems [15]. Moreover, de Bruijn-Good graphs are of interest in other research areas, for instance in cryptography [6, 10, 14, 24] where one of the problems considered is the generation of random-looking pseudo-random sequences to be used as the keystream in so-called additive stream ciphers. One popular way of generating these sequences is by the use of nonlinear feedback shift-registers whose state transition diagram is for every particular feedback function a subgraph of the de Bruijn-Good graph. This aim of this paper is to treat some structural aspects of these de Bruijn-Good graphs. We remark that de Bruijn graphs are sometimes further generalized [8, 15] to graphs having arbitrarily many vertices (not only powers of q), but such generalizations will not be considered in this paper.

For the special case of binary de Bruijn-Good graphs (i.e., $q = 2$), previous authors have treated the problems of counting the number of Hamiltonian cycles of maximal length 2^n [7], generating these maximal length cycles (called de Bruijn sequences) [10], counting the number $\beta^{(2)}(n,k)$ of cycles of a certain length k [2, 3, 23] and determining the maximal number of disjoint cycles into which $G_n^{(2)}$ can be decomposed [12, 17, 21]. The binary case is of special interest because it is best suited for implementations in digital electronics. When dealing with binary graphs we will omit the superscript $^{(2)}$ and use the notations G_n, $\beta(n,k)$ and ν_k consistent with [2] instead. The graphs $G_1, G_2, G_3, G_4, G_1^{(3)}$ and $G_2^{(3)}$ are shown in Figure 1.

In Section 2 we summarize some known results about G_n, generalize some of them to the q-ary graphs $G_n^{(q)}$, and present some extensive tables of $\beta^{(q)}(n,k)$ and, in particular, of $\beta(n,k)$. [We hope that these tables, which extend much beyond the tables previously available in the literature, will be of use to researchers interested in the combinatorial problem of finding exact expressions for $\beta^{(q)}(n,k)$.] Section 3 introduces the main results of the paper, namely the asymptotically tight, but for every k and n valid, lower and upper bounds on $\beta^{(q)}(n,k)$. In Section 4 we give some approximations for $\beta^{(q)}(n,k)$.

2. Known Results and New Numerical Data

A q-ary semi-infinite sequence $\underline{s} = s_0, s_1, \ldots$, with $s_i \in \{0, \ldots, q-1\}$, is said to be periodic with period k if k is the smallest positive integer for which $s_i = s_{i+k}$ for $i \geq 0$. In the following, we shall call a semi-infinite sequence simply a "sequence" when no confusion is possible. With every sequence $\underline{s}$ with period k, one can associate the set $\{(s_0, s_1, \ldots, s_{k-1}), (s_1, s_2, \ldots, s_{k-1}, s_0), \ldots, (s_{k-1}, s_0, \ldots, s_{k-2})\}$ of k distinct k-tuples obtained by shifting a window of length k along the sequence. This set is a q-ary cycle of length k and will be denoted by any one of its elements written in square brackets, e.g., by $[s_0, \ldots, s_{k-1}]$. Note that every cyclic shift of $[s_0, \ldots, s_{k-1}]$ denotes the same cycle, i.e., for example, $[01011] = [10110]$. The total number $\nu_k^{(q)}$ of q-ary length k cycles is given recursively by

$$\nu_k^{(q)} = \frac{1}{k}\left[q^k - \sum_{d|k, d\neq k} d\nu_d^{(q)}\right] \tag{1}$$

since the set of $k\nu_k^{(q)}$ k-tuples associated with all cycles of length k is the set of all q^k k-tuples reduced by those that have a divisor of k as subperiod. The first ten terms of the binary so-called "necklace sequence" ν_k are $2, 1, 2, 3, 6, 18, 30, 56, 99, 186$. The recursive equation (1) can be transformed by the Moebius transform (see for example [22], chapter 20) into the non-recursive form

$$\nu_k^{(q)} = \frac{1}{k}\sum_{d|k} \mu(k/d)\; q^d \tag{2}$$

where $\mu(.)$ is the Moebius function and is defined by $\mu(1) = 1$, $\mu(n) = 0$ if n is divisible by a square, and $\mu(n) = (-1)^k$ if n is the product of k distinct primes.

In the following we denote the length of a cycle c by $T(c)$. We define the *recursive complexity* $D(\underline{s})$ of a q-ary sequence (finite or infinite periodic) $\underline{s} = s_0, s_1, s_2, \ldots$ as the length of the shortest feedback shift-register that can generate it [16, 24], or, more precisely, as the smallest integer d such that there exists a function $f : \{0, \ldots, q-1\}^d \to \{0, \ldots, q-1\}$ such that

$$s_i = f(s_{i-d}, \ldots, s_{i-1}) \qquad \text{for} \quad i \geq d. \tag{3}$$

Similarly, we define the recursive complexity of a q-ary cycle c, denoted by $D(c)$, as the recursive complexity of any one of the k corresponding sequences with period k or, equivalently, as the smallest integer n such that c is contained in $G_n^{(q)}$. Obviously, the sequence $\underline{s}$ has recursive complexity greater than d if and only if $\underline{s}$ contains two identical d-tuples with distinct successor digits since then there exists no function f that generates both digits according to (3). The recursive complexity of $\underline{s}$ is hence the least integer d such that $\underline{s}$ contains no two $(d+1)$-tuples that agree in the first d digits but disagree in the last, i.e., that disagree in the last digit only. The recursive complexity of a cycle $c = [s_0, \ldots, s_{k-1}]$ of length k is thus the least integer d such that there exist no two integers u and v with $0 \leq u < v \leq k-1$ such that $s_{u+i} = s_{v+i}$ for $i = 0, 1, \ldots, d-1$ but $s_{u+d} \neq s_{v+d}$, where all indices are reduced modulo k. For example $D([0001011]) = 3$ because the 3-tuples 010 and 011 differ in the last digit only and there exist no two 4-tuples in 00010110001011000... that differ in the last digit only. Other examples are $D([0010011]) = 5$ and $D([0101011]) = 6$. $\beta^{(q)}(n, k)$ can thus be defined as

$$\beta^{(q)}(n, k) = \#\{c : T(c) = k, D(c) \leq n\}. \tag{4}$$

The basic problem treated in de Bruijn's 1946 paper [7], which introduced the binary graphs G_n, is the determination of the number of Hamiltonian cycles of length 2^n. de Bruijn proved that $\beta(n, 2^n) = 2^{2^{n-1}-n}$. His result can be generalized [10] to

$$\beta^{(q)}(n, q^n) \;=\; [(q-1)!]^{q^{n-1}} \cdot q^{q^{n-1}-n}. \tag{5}$$

One can further show that

$$\beta^{(q)}(n, q^n-1) \;=\; \frac{q}{q-1}\beta^{(q)}(n, q^n).$$

It is obvious that $\beta^{(q)}(n,k) = 0$ for $k > q^n$ and that $\beta^{(q)}(n,k) \le \beta^{(q)}(n+1,k)$, i.e., that every cycle contained in $G_n^{(q)}$ also appears in $G_{n+1}^{(q)}$. One can easily show that $\beta^{(q)}(n,k) = \nu_k^{(q)}$ for $n \ge k-1$, i.e., that $G_n^{(q)}$ contains all cycles of length $n+1$ or less, and that $\beta^{(q)}(n,k) < \nu_k^{(q)}$ for $n < k-1$. Bryant and Everett [3] proved that $\beta(k-2,k) = \nu_k - \phi_k$, where ϕ_k is Euler's totient function and equals the number of positive integers less than or equal to k and relatively prime to k. Bryant and Christensen [2] also proved that, for $k > 5$, $\beta(k-3,k) = \nu_k - 2\phi_{k,2} + 2$, where $\phi_{k,r}$ is defined to be the number of integers $l < k$ satisfying $(k,l) \le r$, where $(.,.)$ denotes the greatest common divisor of the two enclosed integers. They further conjectured that $\beta(k-4,k) = \nu_k - 4\phi_{k,3} - 2(k,2) + 10$ for $k \ge 8$, $\beta(k-5,k) = \nu_k - 8\phi_{k,4} - (k,3) + 19$ for $k \ge 11$, and $\beta(k-6,k) = \nu_k - 16\phi_{k,5} - 4(k,2) - 2(k,3) + 48$ for $k \ge 15$. These conjectures were proved correct by Wan, Xiong and Yu [23] who proved a more general result characterizing the number of cycles of length $k \le \frac{4}{3}(n+1)$ in G_n, namely, if $n < k \le \frac{4}{3}(n+1)$, then

$$\beta(n,k) = \nu_k - 2^{k-n-2}\phi_{k,m-1} - \sum_{i=1}^{k-n-2}\sum_{j=0}^{k-n-i-2}\sum_{2\le q\le 1+(k-n-(k,i)-j)/i}\mu(q)\cdot 2^{(k,i)+e_j}, \tag{6}$$

where $\mu(.)$ is the Moebius function defined earlier and where $e_j = 0$ if $j = 0$ and $e_j = j-1$ if $j > 0$. The fact that this expression is quite complicated already for cycle lengths close to the order n of the de Bruijn-Good graph suggests that for $\frac{4}{3}(n+1) < k \le 2^n$ the expression would be even more complicated. Therefore, and because we are interested in the global behavior of $\beta(n,k)$ and $\beta^{(q)}(n,k)$ more than in the exact numbers, we derive lower and upper bounds on $\beta^{(q)}(n,k)$ in the following section.

The fact that there exists a primitive polynomial of every positive degree over every finite field (see [19]) allows one to prove, by arguments similar to those used in [4], that for q a prime power there exist cycles of every length $k \le q^n$ in $G_n^{(q)}$, i.e., that

$$\beta^{(q)}(n,k) > 0 \quad \text{for} \;\; 1 \le k \le q^n. \tag{7}$$

Essentially, one has to prove that the cycle of (for linear feedback maximal) length $q^n - 1$ generated by a linear feedback shift register with a primitive feedback polynomial [19] can be shortened to any length between 1 and $q^n - 1$ by modifying only 2 entries in the function table of the linear feedback function. This follows from the fact that the digitwise sum of any two phases of such a linear maximal-period sequence is again another phase of the same sequence. We conjecture, but yet are unable to prove, that (7) holds for every $q \ge 2$, not only for prime powers.

In order to assist those working on the difficult combinatorial problem of counting cycles in the de Bruijn-Good graphs $G_n^{(q)}$, and as a reference for later work on the subject, we present some tables of $\beta(n,k)$ and $\beta^{(q)}(n,k)$ which extend much beyond the tables given in [2]. The compilation of these tables required several dozen hours of computation time by an optimized program on a VAX-8600 computer. In particular, the complete lists of the number of cycles of all lengths are given for $G_1 - G_6$ (tables Ia-Id), $G_1^{(3)} - G_3^{(3)}$ (tables IIa and IIb), $G_2^{(4)}$ (table III) and $G_2^{(5)}$ (table IV). It seems to be computationally completely infeasible to compute (and mathematically too difficult to derive) the complete list for any other de Bruijn-Good graph. Some of the entries in table Ib are taken from [5], where $\beta(n,k)$ is tabulated for $k \leq 26$ and $n \leq 26$. Table Ia is given below, all remaining tables are summarized in Appendix B. Horizontal (vertical) arrows indicate that all remaining entries in that row (column) are constant and take on the value of the table entry the arrow points away from.

n \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	2	1	0	0	$\rightarrow$													
2	$\downarrow$	1	2	1	0	0	0	0	$\rightarrow$									
3		$\downarrow$	2	3	2	3	4	2	0	0	0	0	0	0	0	0	$\rightarrow$	
4			$\downarrow$	3	6	7	8	12	14	17	14	13	12	20	32	16	0	0
5				$\downarrow$	6	9	12	20	32	57	78	113	154	208	300	406	538	703
6					$\downarrow$	9	18	26	46	73	124	217	348	574	944	1528	2456	4000
7						$\downarrow$	18	30	50	85	154	271	482	877	1502	2638	4618	8105
8							$\downarrow$	30	56	95	168	309	552	1009	1826	3370	6066	11071
9								$\downarrow$	56	99	176	325	590	1083	1996	3718	6872	12874
10									$\downarrow$	99	186	331	608	1119	2102	3894	7282	13690
11										$\downarrow$	186	335	618	1139	2142	3986	7496	14106
12											$\downarrow$	335	630	1155	2168	4038	7600	14342
13												$\downarrow$	630	1161	2174	4058	7654	14436
14													$\downarrow$	1161	2182	4072	7680	14482
15														$\downarrow$	2182	4080	7694	14510
16															$\downarrow$	4080	7710	14526
17																$\downarrow$	7710	14532
18																	$\downarrow$	14532
19																		$\downarrow$

Table Ia: $\beta(n,k)$ for $1 \leq k \leq 18$ and $1 \leq n \leq 19$.

3. Asymptotically-Tight Lower and Upper Bounds on $\beta^{(q)}(n,k)$

As mentioned earlier, every q-ary cycle of length k corresponds to a set of k q-ary k-tuples. According to the definition, a necessary and sufficient condition for a k-tuple $(s_0, s_1, \ldots, s_{k-1})$ to correspond to a cycle $c = [s_0, s_1, \ldots, s_{k-1}]$ with recursive complexity $D(c) > n$ is that there exist two integers u and v with $0 \leq u < v \leq k-1$ such that $s_{u+i} = s_{v+i}$ for $0 \leq i \leq n-1$ but $s_{u+n} \neq s_{v+n}$, where all indices are reduced modulo

k. The above condition, with $s_{u+n} \neq s_{v+n}$ removed, is a necessary but not sufficient condition for a k-tuple to correspond to a cycle c with recursive complexity $D(c) > n$. For each of the $\binom{k}{2}$ choices for u and v there exist exactly q^{k-n} k-tuples satisfying this condition since $k-n$ digits of the k-tuples can be chosen arbitrarily while the remaining n digits are then completely determined. Hence there exist at most $\binom{k}{2}q^{k-n}$ k-tuples that can possibly correspond to a cycle of recursive complexity greater than n and thus the total number of cycles with recursive complexity greater than n, $\nu_k^{(q)} - \beta^{(q)}(n,k)$, is upperbounded by $(\binom{k}{2}q^{k-n})/k$. Using $\binom{k}{2} < k^2/2$ now gives the following theorem.

Theorem 1: *For every $k \geq 1$ and $n \geq 1$,*

$$\beta^{(q)}(n,k) \ > \ \nu_k^{(q)} - \tfrac{1}{2}kq^{k-n}. \tag{8}$$

For $n = 8$, this bound is shown in Figure 2 as a dotted line. The following corollary illustrates that virtually all cycles of length k have recursive complexity less than $2\log_q k + x$ where x is a small constant.

Corollary to Theorem 1: *For every $k \geq 1$ and for every real number $x > 0$,*

$$\frac{\beta^{(q)}(\lceil 2\log_q k + x\rceil, k)}{\nu_k^{(q)}} \ > \ 1 - \frac{q^{-x}}{2(1-kq^{-k/2})} \tag{9}$$

Here $\lceil . \rceil$ denotes the smallest integer greater or equal to the enclosed number. Note that $kq^{-k/2} \approx 0$ already for moderate k. The left side of (9) approaches 1 exponentially fast with increasing x.

Proof: By letting $n = \lceil 2\log_q k + x\rceil$ in (8) we obtain

$$\beta^{(q)}(\lceil 2\log_q k + x\rceil, k) \ > \ \nu_k^{(q)} - \tfrac{1}{2}kq^{k-\lceil 2\log_q k + x\rceil} \ \geq \ \nu_k^{(q)} - \tfrac{1}{2}kq^{k-2\log_q k - x} \ = \ \nu_k^{(q)} - \frac{q^{k-x}}{2k}. \tag{10}$$

Furthermore, by using (2), $\mu(1) = 1$ and $\mu(n) \geq -1$ for $n \geq 1$, $k\nu_k^{(q)}$ can be lower bounded by

$$k\nu_k^{(q)} \ = \ \sum_{d|k} \mu(k/d)q^d \ \geq \ q^k - \sum_{d|k,d<k} q^d \ \geq \ q^k - \sum_{d|k,d<k} q^{k/2} \ \geq \ q^k - kq^{k/2} \tag{11}$$

The corollary follows by dividing both sides of (10) by $\nu_k^{(q)}$, replacing the resulting term $k\nu_k^{(q)}$ in the denominator by $q^k - kq^{k/2}$ and dividing numerator and denominator by q^k. □

We now turn to the problem of upper bounding $\beta^{(q)}(n,k)$. A necessary condition for a k-tuple $\underline{s} = s_0, \ldots, s_{k-1}$ to correspond to a cycle with recursive complexity less than n, i.e., $D([s_0, \ldots, s_{k-1}]) < n$, is that no two of the $\lfloor k/n \rfloor$ n-tuples, resulting by cutting $\underline{s}$ into non-overlapping n-tuples $(s_0, \ldots, s_{n-1})$ up to $(s_{(\lfloor k/n\rfloor -1)n}, \ldots, s_{\lfloor k/n\rfloor n-1})$ and a "tail" $s_{\lfloor k/n\rfloor n}, \ldots, s_{k-1}$, disagree in the last digit only. Note that the n-tuples need not necessarily be distinct. Here $\lfloor . \rfloor$ denotes the greatest integer smaller or equal to the enclosed number. Let $R^{(q)}(n,t)$ be defined as the number of sequences of t q-ary n-tuples that satisfy the condition (in the sequel called condition C) that no two n-tuples disagree in the last digit only. Then the number of q-ary k-tuples satisfying the above condition is given by $q^{k-tn}R^{(q)}(n,\lfloor k/n\rfloor)$ since the length of the tail is $k-tn$ and the tail digits can be chosen arbitrarily. Hence the number of cycles with recursive complexity less than n, i.e., less or equal to $n-1$, is upper bounded by

$$\beta^{(q)}(n-1,k) \;\le\; \frac{1}{k}q^{k-tn}R^{(q)}(n,\lfloor k/n\rfloor). \tag{12}$$

In order to derive an upper bound on $R^{(q)}(n,t)$, we partition the set of sequences of t n-tuples satisfying condition C into t subsets according to the number of distinct n-tuples they contain. Let $\overline{R}^{(q)}(n,t,r)$ be the number of sequences of t q-ary n-tuples that satisfy condition C and that contain exactly r distinct n-tuples. Then, obviously,

$$R^{(q)}(n,t) \;=\; \sum_{r=1}^{t}\overline{R}^{(q)}(n,t,r). \tag{13}$$

In the sequel a recursive equation for $\overline{R}^{(q)}(n,t,r)$ is derived that will be used later to derive an upper bound on $R^{(q)}(n,t)$.

Given a sequence of $t-1$ n-tuples satisfying condition C and containing exactly i distinct n-tuples$(1\le i\le t-1)$, a sequence of t n-tuples still satisfying condition C can be obtained either by adding an n-tuple that already occurred (i possibilities), or by adding an n-tuple that does not agree with any of the $t-1$ n-tuples in the first $n-1$ digits. For this second case there are $(q^{n-1}-i)q$ possibilities since the number of choices for the first $n-1$ digits and for the last digit are $q^{n-1}-i$ and q, respectively. In the first case the number of distinct n-tuples remains constant, but in the second case it is increased by 1 to $i+1$. $\overline{R}^{(q)}(n,t,r)$ is thus given recursively as

$$\overline{R}^{(q)}(n,t,r) \;=\; r\cdot\overline{R}^{(q)}(n,t-1,r)+[q^n-q(r-1)]\cdot\overline{R}^{(q)}(n,t-1,r-1) \tag{14}$$

with the trivial initial condition $\overline{R}^{(q)}(n,t,1)=q^n$ for $n\ge 1$ and $t\ge 1$ and with the convention that $\overline{R}^{(q)}(n,t,r)=0$ for $r\le 0$ and $r>t$. It is proved in appendix A that the solution of this recursion satisfies the following upper bound.

Lemma 1: *For $1\le r\le t$ and $n\ge 1$, $\overline{R}^{(q)}(n,t,r)$ is upper bounded by*

$$\overline{R}^{(q)}(n,t,r) \;\le\; \frac{t^{2(t-r)}q^{rn}}{2^{t-r}(t-r)!}\,\exp\Big\{-\tfrac{1}{2}r(r-1)q^{-n+1}\Big\}. \tag{15}$$

The following lemma is also proved in appendix A by application of Lemma 1 and equation (13).

Lemma 2: *For $n\ge 1$ and $t\ge 1$,*

$$R^{(q)}(n,t) \;\le\; q^{tn}\,\exp\Big\{-\tfrac{1}{2}q^{-n+1}t(t-1)+\tfrac{1}{2}t^2q^{-n}e^{-tq^{-n+1}}\Big\}. \tag{16}$$

Theorem 2 is now an immediate consequence of inequality (12), applied to $\beta^{(q)}(n,k)$ instead of $\beta^{(q)}(n-1,k)$, and of Lemma 2, for which the terms in the exponent expression are reordered.

Theorem 2: *For every $k \geq 1$ and $n \geq 1$,*

$$\beta^{(q)}(n,k) \;<\; \frac{q^k}{k}\,\exp\left\{-\tfrac{1}{2}t^2 q^{-n}\left[1-e^{-tq^{-n}}/q\right] \;+\tfrac{1}{2}tq^{-n}\right\} \tag{17}$$

where $t = \lfloor k/(n+1)\rfloor$.

The aim in applying Theorem 2 is to choose n such that $tq^{-n} \approx 0$ but t^2q^{-n} is substantially greater than 0. The choice $n = \lceil 2\log_q k - 2\log_q\log_q k - x\rceil$ guarantees that

$$\frac{(\log_q k)^2}{k^2}q^{x-1} \;<\; q^{-n} \;\leq\; \frac{(\log_q k)^2}{k^2}q^{x}\;, \tag{18}$$

$$\text{and with}\quad \frac{k-n-1}{n+1} \;=\; \frac{k}{n+1}-1 \;<\; t \;=\; \left\lfloor\frac{k}{n+1}\right\rfloor \;\leq\; \frac{k}{n+1} \quad\text{that} \tag{19}$$

$$\begin{aligned} tq^{-n} \;&<\; \frac{k}{2\log_q k - 2\log_q\log_q k - x + 1}\,\frac{(\log_q k)^2}{k^2}q^x \\ &=\; \frac{(\log_q k)^2 q^x}{k(2\log_q k - 2\log_q\log_q k - x + 1)} \end{aligned} \tag{20}$$

$$\begin{aligned} \text{and}\quad t^2q^{-n} \;&>\; \left(\frac{k - 2\log_q k - 2\log_q\log_q k - x - 1}{2\log_q k - 2\log_q\log_q k - x + 2}\right)^2 \frac{(\log_q k)^2}{k^2}q^{x-1} \\ &=\; \frac{1}{4}q^{x-1}\left(\frac{1-(2\log_q k - 2\log_q\log_q k - x - 1)/k}{1-(2\log_q\log_q k + x - 2)/(2\log_q k)}\right)^2. \end{aligned} \tag{21}$$

Note that for a fixed x (and fixed q), if $k \to \infty$ (which implies that $n \to \infty$ as well), then $tq^{-n} \to 0$, $e^{-tq^{-n}} \to 1$ and $t^2q^{-n} \to \frac{1}{4}q^{x-1}$ which by Theorem 2 together with $\lim_{k\to\infty} q^k/(k\nu_k^{(q)}) = 1$ (which follows from (11) and $k\nu_k^{(q)} \leq q^k$) implies the first inequality of the following theorem. The second inequality is a direct consequence of the Corollary to Theorem 1 and the fact that $\lim_{k\to\infty} kq^{-k/2} = 0$.

Theorem 3: *For every positive real number x,*

$$\limsup_{k\to\infty}\frac{\beta^{(q)}(\lceil 2\log_q k - 2\log_q\log_q k - x\rceil, k)}{\nu_k^{(q)}} \;\leq\; \exp\left\{-\tfrac{1}{8}q^{x-1}(1-1/q)\right\} \tag{22}$$

$$\textit{and}\quad \liminf_{k\to\infty}\frac{\beta^{(q)}(\lceil 2\log_q k + x\rceil, k)}{\nu_k^{(q)}} \;\geq\; 1-\tfrac{1}{2}q^{-x} \tag{23}$$

In particular, for every positive-valued function $f(k)$ with $\lim_{k\to\infty} f(k) = \infty$ (e.g. $f(k) = \log_q\log_q\log_q k$),

$$\lim_{k\to\infty}\frac{\beta^{(q)}(n(k),k)}{\nu_k^{(q)}} \;=\; 0 \quad \textit{for}\quad n(k) = \lfloor 2\log_q k - 2\log_q\log_q k - f(k)\rfloor, \quad \textit{and}$$

$$\lim_{k\to\infty}\frac{\beta^{(q)}(n(k),k)}{\nu_k^{(q)}} \;=\; 1 \quad \textit{for}\quad n(k) = \lfloor 2\log_q k + f(k)\rfloor.$$

Theorem 3 demonstrates that the recursive complexity of virtually all of the $\nu_k^{(q)}$ q-ary cycles of length k is within a very small interval of width roughly $2\log_q \log_q k$ and located close to $2\log_q k$.

4. Approximations for $\beta^{(q)}(n,k)$

For investigating the global behaviour of $\beta(n,k)$ it is advantageous to use a logarithmic scale. The plots of $\log_2 \nu_k$ and $\log_2 \beta(n,k)$ for $3 \le n \le 8$ are shown in Figure 2. From the fact that the curves in Figure 2 are close to linear in certain ranges it is obvious that for $n \ge 5$ the following approximation holds for wide ranges of k-values: $\log_2 \beta(n,k) \approx 2\log_2 \beta(n,k-1) - \log_2 \beta(n,k-2)$ which is equivalent to

$$\beta(n,k) \;\approx\; \frac{\beta(n,k-1)^2}{\beta(n,k-2)}. \tag{24}$$

For almost all values of k, this approximation is slightly greater than the actual value, which means that the curves $k \to \log_2 \beta(n,k)$ are almost everywhere concave, with exceptions at both ends of the interval $[1, 2^n]$.

Another good approximation for $n \ge 5$ for the range $2n \le k \le 2^{n/4}$ is $\log_2 \beta(n,k) \approx \frac{1}{2}[\log_2 \beta(n-1,k) + \nu_k]$ which is equivalent to

$$\beta(n,k) \;\approx\; \sqrt{\beta(n-1,k)\cdot \nu_k}\;. \tag{25}$$

These approximations for $\beta(n,k)$ also hold for $\beta^{(q)}(n,k)$ as can be verified by inspection of tables IIa, IIb and III.

5. Conclusions

The problem of enumerating the number $\beta^{(q)}(n,k)$ of cycles of length k in the generalized de Bruijn-Good graphs $G_n^{(q)}$ has been treated by an approach that is more statistical than combinatorial. Asymptotically-tight lower and upper bounds on $\beta^{(q)}(n,k)$ have been derived that imply that the recursive complexity of virtually all cycles of length k is very close to $2\log_q k$, where the recursive complexity of a cycle is defined as the length of the shortest feedback shift-register that can generate it or, equivalently, as the smallest order n of a de Bruijn-Good graph $G_n^{(q)}$ in which it occurs. Similar asymptotically-tight lower and upper bounds as given in Theorem 3 can be derived [20] for the number of finite q-ary sequences of length k having recursive complexity less than $\lceil 2\log_q k - 2\log_q \log_q k - x\rceil$ or greater than $\lceil 2\log_q k + x\rceil$, respectively.

Acknowledgement

It is a pleasure to acknowledge the contributions of Felix Tarköy to this paper; major parts of Appendix A are taken from our joint work [20]. One of the referees suggested

an improvement in the proof of Lemma 1 that we gratefully acknowledge. Results for a different but related problem that are of a form similar to the second part of Theorem 3 have been found independently by Arratia and Waterman [1].

Appendix A: Proofs of Lemmas 1 and 2

Proof of Lemma 1: We first prove by induction that for $t \geq r$,

$$\overline{R}^{(q)}(n,t,r) \;\leq\; \frac{t^{2(t-r)}}{2^{t-r}(t-r)!} \prod_{i=0}^{r-1}(q^n - iq)\;. \tag{26}$$

As the basis of the induction we note that for $r = 1$ inequality (26) holds for all t, namely

$$\overline{R}^{(q)}(n,t,1) \;=\; q^n \;\leq\; \frac{t^{2(t-1)}}{2^{t-1}(t-1)!}q^n \;=\; \underbrace{\frac{(t^2/2)^{t-1}}{(t-1)!}}_{\geq 1 \text{ for } t\geq 1} q^n. \tag{27}$$

Assuming that (26) holds for $\overline{R}^{(q)}(n,t{-}1,r)$ and $\overline{R}^{(q)}(n,t{-}1,r{-}1)$ we show by application of (14) that (26) holds for $\overline{R}^{(q)}(n,t,r)$:

$$\begin{aligned}
\overline{R}^{(q)}(n,t,r) &= r\cdot\overline{R}^{(q)}(n,t{-}1,r) \;+\; [q^n - q(r{-}1)]\cdot\overline{R}^{(q)}(n,t{-}1,r{-}1) \\
&\leq r\cdot\frac{(t{-}1)^{2(t-r-1)}}{2^{t-r-1}(t{-}r{-}1)!}\prod_{i=0}^{r-1}(q^n - iq) + [q^n - q(r{-}1)]\frac{(t{-}1)^{2(t-r)}}{2^{t-r}(t{-}r)!}\prod_{i=0}^{r-2}(q^n - iq) \\
&= \underbrace{\left[2r(t{-}r)(t{-}1)^{2(t-r-1)} + (t{-}1)^{2(t-r)}\right]}_{T}\frac{1}{2^{t-r}(t-r)!}\prod_{i=0}^{r-1}(q^n - iq)\;.
\end{aligned}\tag{28}$$

It remains to show that $T \leq t^{2(t-r)}$. For $r = t$, $T = 1 = t^{2(t-r)}$. For $r \leq t{-}1$,

$$\begin{aligned}
t^{2(t-r)} &= \Big[(t-1)+1\Big]^{2(t-r)} \\
&\geq (t-1)^{2(t-r)} + \tbinom{2(t-r)}{1}(t-1)^{2(t-r)-1} + \tbinom{2(t-r)}{2}(t-1)^{2(t-r)-2} \\
&= (t-1)^{2(t-r)} + \Big[2(t-r)(t-1) + (t-r)(2t-2r-1)\Big]\,(t-1)^{2(t-r-1)} \\
&= (t-1)^{2(t-r)} + \underbrace{(4t-2r-3)}_{>2r \text{ for } r\leq t-1}(t-r)(t-1)^{2(t-r-1)}\;.
\end{aligned}$$

For $r \leq t-1$ (i.e., $t \geq r+1$), $4t-2r-3 \geq 4(r+1)-2r-3 = 2r+1 > 2r$, which, by comparison with T, implies $T \leq t^{2(t-r)}$ for $r \leq t-1$ and together with (28) proves (26).

The second step of the proof of Lemma 1 is to show that

$$\prod_{i=0}^{r-1}(q^n - iq) \;=\; q^{rn}\prod_{i=1}^{r-1}\left(1 - iq^{-n+1}\right) \;\leq\; q^{rn}\exp\left\{-\tfrac{1}{2}r(r{-}1)q^{-n+1}\right\}. \tag{29}$$

For $r > q^{n-1}$, the product on the left is zero and hence the inequality is trivially satisfied. For $r \leq q^{n-1}$, all the terms in the product are positive. Let $\alpha = q^{-n+1}$. Using the fact that $\ln(1-x) \le -x$ for $x < 1$ we obtain

$$\ln \prod_{i=1}^{r-1} (1 - iq^{-n+1}) \;=\; \sum_{i=1}^{r-1} \ln(1 - i\alpha) \;\le\; -\sum_{i=1}^{r-1} i\alpha \;=\; -\frac{1}{2} r(r-1)\alpha.$$

The inequality in (29) follows immediately. $\square$

Proof of Lemma 2: Application of Lemma 1 and equation (13) yields

$$R^{(q)}(n,t) \;\le\; \sum_{r=1}^{t} \frac{t^{2(t-r)}}{2^{t-r}(t-r)!}\, q^{rn} \exp\left\{-\tfrac{1}{2} r(r-1) q^{-n+1}\right\} \; .$$

We now multiply the sum by q^{tn} and compensate the effect by replacing q^{rn} in the sum by $q^{(r-t)n}$. Using the index transformation $s = t-r$ such that $r(r-1) = (t-s)(t-s-1) = t(t-1) - 2ts + s(s+1) \ge t(t-1) - 2ts$ we obtain

$$\begin{aligned} R^{(q)}(n,t) \;&\le\; q^{tn} \sum_{s=0}^{t-1} \frac{1}{s!} \left(\frac{t^2 q^{-n}}{2}\right)^{s} \exp\left\{\tfrac{1}{2} q^{-n+1} [t(t-1) - 2ts]\right\} \\ &=\; q^{tn} \exp\left\{\tfrac{1}{2} q^{-n+1} t(t-1)\right\} \sum_{s=0}^{t-1} \frac{1}{s!} \left(\tfrac{1}{2} t^2 q^{-n} e^{-tq^{-n+1}}\right)^{s} \; . \end{aligned}$$

Extending the summation from $s = 0$ to ∞ cannot reduce the sum (because the terms are positive) and transforms it into the power series expansion of $\exp\{\frac{1}{2} t^2 q^{-n} e^{tq^{-n+1}}\}$, which is hence an upper bound on the sum. This completes the proof of Lemma 2. $\square$

Appendix B: Tables of $\beta^{(q)}(n,k)$

n \ k	19	20	21	22	23	24	25	26	27	28
5	842	1085	1310	1465	1544	1570	1968	2132	2000	2480
6	6348	10131	15970	24625	37972	57802	86608	128355	188602	272634
7	14262	24931	43912	76236	132632	229990	397260	684130	1173028	2006754
8	20222	37001	67748	123807	226764	415004	758616	1385771	2531084	4618229
9	23782	44341	82880	154876	290268	543880	1020356	1914402	3595934	6751951
10	25662	48517	91182	172256	326162	618302	1173910	2230341	4243134	8077453
11	26620	50433	95494	181839	345218	658042	1256436	2401716	4597790	8810393
12	27102	51385	97652	186169	354918	679278	1298946	2490427		
13	27348	51879	98748	188357	359808	689052	1320506	2537529		
14	27468	52127	99366	189451	362258	693946	1331322	2559125		
15	27530	52267	99608	190001	363488	696398	1336726	2569951		
16	27560	52335	99730	190275	364102	697752	1339446	2575359		
17	27576	52355	99794	190415	364412	698310	1340810	2578075		
18	27594	52368	99836	190483	364564	698616	1341490	2579431		
19	$\downarrow$	52377	99846	190519	364642	698742	1341834	2580111		
20		$\downarrow$	99858	190547	364680	698812	1342034	2580449		
21			$\downarrow$	190557	364700	698848	1342104	2580621	4970662	
22				$\downarrow$	364722	698862	1342138	2580705	4970832	9586049
23					$\downarrow$	698870	1342156	2580749	4970920	9586221
24						$\downarrow$	1342178	2580783	4970974	9586329
25							$\downarrow$	2580795	4970990	9586361
26								$\downarrow$	4971008	9586381
27									$\downarrow$	9586395

Table Ib: $\beta(n,k)$ for $19 \le k \le 28$ and $5 \le n \le 27$.

n \ k	29	30	31	32	33	34	35	36
4	0	0	0	0	$\rightarrow$			
5	2176	2816	4096	2048	0	0	0	0
6	390190	552724	768844	1060280	1443260	1930641	2559256	3348409
7	3410476	5777696	9741000	16361136	27357028			
8	8414038	15317619	27845580	50567566	91678382			

Table Ic: $\beta(n,k)$ for $29 \le k \le 36$ and $4 \le n \le 8$.

k	$\beta(6,k)$	k	$\beta(6,k)$	k	$\beta(6,k)$
35	2559256	45	21235540	55	59083776
36	3348409	46	24504208	56	63380992
37	4311450	47	28452128	57	61390848
38	5492251	48	32129328	58	60764160
39	6896304	49	35951488	59	62619648
40	8593846	50	40066592	60	70057984
41	10507554	51	44494144	61	59768832
42	12800269	52	48144432	62	88080384
43	15264574	53	51336384	63	134217728
44	18044775	54	54675776	64	67108864

Table Id: $\beta(6,k)$ for $35 \leq k \leq 64$.

n \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	3	3	2	0	0	0	0	0	0	$\rightarrow$					
2	$\downarrow$	3	8	12	18	20	24	36	24	0	0	0	0	0	0
3		$\downarrow$	8	18	36	86	186	372	792	1596	3108	6002	11088	19152	31752
4			$\downarrow$	18	48	110	264	672	1644	4071	10158	25335	63006	155802	383286
5				$\downarrow$	48	116	294	762	1998	5340	14010	37305	100002	268554	723806
6					$\downarrow$	116	312	798	2136	5688	15390	42090	114870	316122	874352
7						$\downarrow$	312	810	2166	5814	15864	43500	120036		
8							$\downarrow$	810	2184	5868	16020				

Table IIa: $\beta^{(3)}(n,k)$ for $1 \leq k \leq 15$ and $1 \leq n \leq 8$.

n \ k	16	17	18	19	20	21	22 (25)	23 (26)	24 (27)
3	51216	77952	113712	160608	212160	259648	317952 435456	369792 559872	376704 373248
4	937272	2274078	5481078	13102062					
5	1950894	5265714							

Table IIb: $\beta^{(3)}(n,k)$ for $16 \le k \le 27$ and $3 \le n \le 5$.

n \ k	1	2	3	4	5	6	7	8	9	10	11	12	13 (15)	14 (16)
1	4	6	8	6	0	0	0	0	0	0	0	0	0	0
2	$\downarrow$	6	20	48	120	280	672	1500	2976	5472	9216	13824	17280 27648	20736 20736
3		$\downarrow$	20	60	180	586	1848	5844	18872	60408	192696	612158	1925088	5996376
4			$\downarrow$	60	204	658	2208	7644	26372	92334	326124	1157560	4124832	
5				$\downarrow$	204	670	2304	8028	28424	102006	367236	1335580		
6					$\downarrow$	670	2340	8136	28988	104046	377760			

Table III: $\beta^{(4)}(n,k)$ for $1 \le k \le 16$ and $1 \le n \le 6$.

k	$\beta^{(5)}(2,k)$	k	$\beta^{(5)}(2,k)$	k	$\beta^{(5)}(2,k)$
1	5	10	167808	19	306892800
2	10	11	495360	20	479499264
3	40	12	1392240	21	678481920
4	130	13	3692160	22	846028800
5	444	14	9241920	23	995328000
6	1500	15	21747456	24	1244160000
7	5160	16	47678400	25	995328000
8	17130	17	96595200		
9	54600	18	179781120		

Table IV: $\beta^{(5)}(2,k)$ for $1 \le k \le 25$.

References

[1] R. Arratia and M.S. Waterman, *An Erdös-Rényi law with shifts*, Adv. in Math., Vol. 55, pp. 13-23, 1985.

[2] P.R. Bryant and J. Christensen, *The enumeration of shift-register sequences*, J. Combinatorial Theory (A), Vol. 35, pp. 154-172, 1983.

[3] P.R. Bryant and D. Everett, *Cycles from feedback shift registers: a counting problem*, in Kyoto International Conference on Circuit and System Theory, Kyoto, Japan, 1970.

[4] P.R. Bryant, F.G. Heath and R.D. Killick, *Counting with feedback shift registers by means of a jump technique*, IRE Trans. Electron. Comput., Vol. EC-19, pp. 1204-1209, 1970.

[5] T. Burger and F. Tarköy, *Der de Bruijn-Good Graph und die nichtlineare Komplexitaet binaerer Sequenzen*, diploma thesis, Inst. for Signal and Info. Proc., ETH Zurich, 1987.

[6] A.H. Chan, R.A. Games and E.L. Key, *On the complexity of de Bruijn sequences*, J. Combinatorial Theory (A), Vol. 33, pp. 233-246, 1982.

[7] N.G. de Bruijn, *A combinatorial problem*, Proc. K. Ned. Akad. Wet. Ser. A, Vol. 49, pp. 758-764, 1946.

[8] D.Z. Du and F.K. Hwang, *Generalized de Bruijn digraphs*, Networks, Vol. 18, No. 1, pp. 27-38, 1988.

[9] H.M. Fredricksen, *Disjoint cycles from the de Bruijn Graph*, Ph.D. Thesis, University of Southern California, 1968.

[10] H. Fredricksen, *A survey of full length nonlinear shift register cycle algorithms*, SIAM Review, Vol. 24, No. 2, pp. 195-221, 1982.

[11] I.J. Good, *Normal recurring decimals*, J. London Math. Soc., Vol. 21, pp. 169-172, 1946.

[12] S.W. Golomb, *Shift Register Sequences*, revised edition, Aegean Park Press, Laguna Hills, CA, 1982.

[13] S.W. Golomb, L.R. Welch and R.M. Goldstein, *Cycles from nonlinear shift registers*, Jet Propulsion Lab., California Institute of Technology, Pasadena, CA, Progress Report 20-389, August 1959.

[14] C.G. Günther, *Alternating step generators controlled by de Bruijn sequences*, Proc. EUROCRYPT'87, Lecture Notes in Comp. Sc., Vol. 304, pp. 5-14, Berlin, Heidelberg: Springer, 1988.

[15] M. Imase and M. Itoh, *A Design for directed graphs with minimum diameter*, IEEE Trans. Comput., Vol. C-32, pp. 782-784, 1983.

[16] C.J.A. Jansen, *Investigations on nonlinear streamcipher systems: construction and evaluation methods*, Ph. D. Thesis, Technical University Delft, The Netherlands, 1989.

[17] A. Lempel, *On extremal factors of the de Bruijn Graph*, J. Combinatorial Theory (B), Vol. 11, pp. 17-27, 1971.

[18] A. Lempel, *On a homomorphism of the de Bruijn Graph and its application to the design of feedback shift registers*, IEEE Trans. on Computers, Vol. C-19, No. 12, pp. 1204-1209, 1970.

[19] R. Lidl and H. Niederreiter, *Finite fields*, Encyclopedia of Math. and its Appl., Vol. 20, Addison-Wesley, 1983.

[20] U.M. Maurer and F. Tarköy, *The nonlinear shift-register complexity of random sequences*, unpublished report, Inst. for Signal and Info. Proc., ETH Zurich, March 1987.

[21] J. Mykkeltveit, *A proof of Golomb's conjecture for the de Bruijn Graph*, J. Combinatorial Theory (B), Vol. 13, pp. 40-45, 1972.

[22] M.R. Schroeder, *Number Theory in Science and Communication*, 2nd edition, Springer Verlag, Series in Information Sciences, 1986.

[23] Z. Wan, R. Xiong and M. Yu, *On the number of cycles of short length in the de Bruijn-Good Graph G_n*, Preprint, Graduate School, Academia Sinica, Beijing, 1985.

[24] M. Wang, *Cryptographic aspects of sequence complexity measures*, Ph. D. Thesis No. 8723, Swiss Federal Institute of Technology, Zurich, 1988.