Kragujevac J. Math. 24 (2002) 19–41.

ENUMERATING KAUTZ SEQUENCES

Vladimir Raphael Rosenfeld

Institute of Evolution, University of Haifa, Mount Carmel, Haifa 31905, Israel
(e-mail: vladimir@research.haifa.ac.il)

(Received August 5, 2002)

Abstract. A *Kautz s-ary closed sequence* is a circular sequence of l s-ary digits $0, 1 \ldots, s-1$ such that consecutive digits are distinct and all subsequences of length q are distinct, too [3]. Kautz sequences (of the maximal length $s(s-1)^{q-1}$) can also be represented by the series $\mathcal{H}_s = \{H_{s,q}\}_{q=1}^{\infty}$ $(s \geq 2)$ of Kautz digraphs [3]. Namely, $H_{s,1} = K_s$, where K_s is a complete s-vertex digraph without self-loops, and $H_{s,q+1} = \Gamma(H_{s,q}) = \Gamma^q K_s$, where Γ is the operator transforming an arbitrary (di-)graph G into its arc-graph $\Gamma(G)$ [8].

Under $s, q \geq 2$, the number of the Kautz sequences of the maximal length $s(s-1)^{q-1}$ is proven to equal $s^{s-2}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s+q-2}$. The demonstration is based on our recent results concerning the characteristic polynomial and permanent of the arc-graph [8], applied herein to the Kautz digraphs.

Wherever possible, the main subject is discussed in the wider context of related combinatorial problems, which first includes counting the *linear Kautz sequences*, whose number under the maximal length $s(s-1)^{q-1} + q - 1$ is equal to $s^{s-1}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s-1}$.

Obtained results can be used for calculating the number of monocyclic and linear compounds, formed from s sorts of atoms, obeying the specified combinatorial restrictions. The former is equivalent to finding the number of respective *necklaces with s kinds of beads*.

1. INTRODUCTION

The Kautz sequences [1–3], as they were introduced in the summary, are only one specific type of sequences that can occur in nature or be obeyed in a targeted human practise. The closest relatives of these are De Bruijn sequences [4–7], which additionally allow its adjacent ciphers to be equal. Our accepted paper [7] is just devoted to the enumeration of De Bruijn sequences and some their generalizations; ready results from [7] will be cited by us for comparison, later on. As well as [7], the present paper will be based on our previous results concerning the characteristic polynomial and permanent of the arc-graph [8], specially applied herein to Kautz graphs (see [3]). We shall also actively adopt the general ideology of [1–7], among which [5] specifically concerns Ch. 9 of the famous book by Hall. According to this, the enumeration of the sequences under consideration can be reduced to the count of Eulerian circuits in special ancillary digraphs bearing namesake's name of Kautz digraphs (see [3]).

In more detail, all necessary aspects of our work will be discussed in the main part of the text. Right now, the author would like to emphasize just two specific reasons that had influenced him considerably and provoked his work.

The first was the study of complex sequences being carried on by the research group under the supervision of Profs. Edward Trifonov and Alexander Bolshoy, in the Genome Diversity Center of the University of Haifa (see [9–14]). In particular, Dr. Valery Kirzhner defined a minimal generating sequence in DNA as the sequence of minimal length that produces all possible amino acids; thus, it should contain all triplets of nucleotides, taking into account the table of identity of some triplets. Under this, at the first stage of work, one can disregard the equivalency of some triplets. Such a minimal sequence is, in some sense, the most complex [9–14]; and the mathematical formalization of it leads to De Bruijn sequences and, under additional restrictions, to Kautz sequences ([7]).

The second was that the properties of closed and unclosed Kautz sequences (as well as De Bruijn ones) can be utilized in the synthetical chemistry of cyclic and linear molecules, respectively. Cases in point are engineering and design of new reagents for Analytical Chemistry or drugs that employ the principles of Combinatorial Chemistry. At the first stage of synthesis, when the general prognosis should be done, the researcher is much interested in devising "the most concentrated" all-inclusive molecule which allows one to simultaneously incorporate, in one reagent, all spatial compositions of reactive groups to be attested. Moreover, such a substance should enable every mentioned composition of groups (in our case, displayed by a different segment of a Kautz sequence) to contest for the best credits under equal starting conditions. Then, when the optimal molecular substructures are already determined, one can turn to the synthesis of rather simple molecules that exclude "badly behaved" parts of the first "supermolecule". Clearly, such a tack could economize syntheticist's time.

The last chemical example, even though it was described briefly, puts forward the idea of replacing an intact Kautz $[s(s-1)^{q-1}]$-sequence, of maximal length, with all possible sets of shorter sequences (collectively comprising the same set of $s(s-1)^{q-1}$ q-subwords). Here, the solution for distributing a complete cycle immediately comes from our recent finding for the permanent of the arc-graph [8].

In our opinion, the above problems and their solutions can better be discussed in the wider context of similar combinatorial questions. However, planning to consider some additional problems in the subsequent sections, we have no intention whatever to make a detailed survey in this paper. For this reason, all references will be given in minimal numbers. We would like only to stress that other trends also exist and are all interesting as well. Wherever possible, we shall also propose problems that the reader can try to solve. Our general goal is to enhance the interest of chemists in Mathematics and, conversely, attract mathematicians to the wider range of problems that come from Chemistry, Biology and other sciences.

Now we must supply mathematical requisites that will be used by us later, in the main section.

2. PRELIMINARIES

This section culls just all known facts from Combinatorics and (Spectral) Theory of Graphs that will be needed for proving our targeted results; all information concerning allied areas will be given in Miscellaneous.

2.1 KAUTZ SEQUENCES

A *Kautz s-ary closed sequence* is a circular sequence of l s-ary digits $0, 1 \ldots, s-1$ such that consecutive digits are distinct and all subsequences of length q are distinct, too [3]. Thus, Kautz sequences are non-DeBruijn sequences included in the respective De Bruijn s^q-sets [7], with an additional proviso that equal digits may never be adjacent therein. Kautz sequences (of the maximal length $s(s-1)^{q-1}$) can also be represented by the series $\mathcal{H}_s = \{H_{s,q}\}_{q=1}^{\infty}$ $(s \geq 2)$ of Kautz digraphs [3] that resemble De Bruijn graphs [4–7]. Namely, $H_{s,1} = K_s$, where K_s is a complete s-vertex digraph without self-loops, and $H_{s,q+1} = \Gamma(H_{s,q}) = \Gamma^q K_s$, where Γ is the operator transforming an arbitrary (di-)graph G into its arc-graph $\Gamma(G)$ [8]. Villar [3] proved that Kautz sequences exist for all lengths l except for 1 and $r(r-1)^{q-1} - 1$, where $q \geq 2$ and $r = s(s-1)^q$ is the number of arcs in a digraph $H_{s,q}$. In particular, for $s = 3$ and $q = 1, 2$, there exist the following Kautz sequences:

$q = 1$	012	021
$q = 2$	121323	
	123132	
	123213	.

Apparently, cutting a Kautz $[s(s-1)^{q-1}]$-cycle $(q \geq 2)$ in all $s(s-1)^{q-1}$ positions generates $s(s-1)^{q-1}$ distinct words since any such cycle is circularly asymmetric, by definition. However, every $[s(s-1)^{q-1}]$-word obtained in this fashion contains only $s(s-1)^{q-1} - q + 1$ basic subwords of length q, out of those belonging to the complete cycle. A minimal word of length $s(s-1)^{q-1} + q - 1$ that incorporates just the same set of $s(s-1)^{q-1}$ basic q-subwords as an intact complete cycle is called a *linear Kautz sequence*. Obviously, a linear Kautz sequence can be obtained by adding the first $q-1$ letters of any $[s(s-1)^{q-1}]$-word, obtained by cutting a complete Kautz cycle, to the end of this word. As a brief illustration, we shall consider the cases for $s = 3$ and $q = 1, 2$, as these follow from the above example for circular Kautz sequences:

$q = 1$	012	120	201	021	210	102
$q = 2$	1213231	2132312	1323121	3231213	2312132	3121323
	1231321	2313212	3132123	1321231	3212313	2123132
	1232131	2321312	3213123	2131232	1312321	3123213.

Another generalization of Kautz sequence is a *Kautz* $[s(s-1)^{q-1}]$*-set* of sequences which are not maximal Kautz sequences on their own, except for the case when a Kautz set consists of exactly one Kautz sequence, but collectively have the same aggregated length $s(s-1)^{q-1}$ and also produce the same set of all s-ary words of length q; see Theorem 10 and Corollary 10.1, in Section 3.

In order to proceed, we need to introduce some graph-theoretical notions (see [15–22; 8]). A *directed graph*, or *digraph*, D of order n consists of a finite nonempty set V of different objects that are called *vertices*, or *points*, together with a given set E containing m ordered pairs of different vertices of the set V. A pair (u, v), or uv, of vertices from V is called an *arc* of a digraph D that emanates from a vertex u and enters a vertex v; under $u = v$, an arc uu (vv) is called a *self-loop* lying in the point u (v). If an arc uv exists, in D, we say that a vertex u is adjacent to a vertex v; and a vertex u and an arc uv are *incident* to each other, as well as an arc uv and a vertex v are. The *out-degree* $d^+(v)$ of a vertex v is the number of arcs that go out of it, including self-loops; symmetrically, the *in-degree* $d^-(v)$ of v is the number of arcs (and self-loops) that come into it. In lieu of the term *degree*, we also use its synonym *valency*, which may seem preferable while describing chemical objects.

Following [3], we need to define the series $\mathcal{H}_s = \{H_{s,q}\}_{q=1}^{\infty}$ $(s \geq 2)$ of special digraphs that will be used by us in the further proof; here, the numbers s and q have the same interpretation as above. Initially, we set $H_{s,1}$ to be a complete s-vertex digraph without self-loops. The set $V_{s,q}$ of vertices of a digraph $H_{s,q-1}$ $(q \geq 2)$ consists of all $s(s-1)^{q-2}$ ordered sequences, or words, of $q-1$ letters over the alphabet A, wherein no pair of adjacent letters are equal, while the set E of arcs is in one-one correspondence with all $s(s-1)^{q-1}$ words of q letters, over A, with the same adjacency restrictions. Under this, the arc uv labeled by a word $a_1a_2\cdots a_{q-1}a_q$ emanates from a vertex $u = a_1a_2\cdots a_{q-1}$ and enters a vertex $v = a_2\cdots a_{q-1}a_q$. In other words, arcs $a_1a_2\cdots a_{q-1}a_q$ and $a_2a_3\cdots a_qa_{q+1}$ share a common incident vertex $a_2a_3\cdots a_{q-1}a_q$. It is easy to see that the arc set $E_{s,q}$ of a digraph $H_{s,q}$ is simultaneously the vertex set $V_{s,q+1}$ of the next digraph $H_{s,q+1}$, in $\mathcal{H}_s$ (see [1–3]). But what is rather more important, $H_{s,q+1}$ $(q \geq 1)$ can be obtained from $H_{s,q}$ by the process that can locally be called taking the arc-graph $\Gamma(H_{s,q})$ of a digraph $H_{s,q}$ (see [8]); under this, $H_{s,q+1} = \Gamma(H_{s,q})$. The members of the series $\mathcal{H}_s$ are called *Kautz graphs* (see [3]). Herein, we shall adapt the methods applied in [4–7], wherein calculating the number of complete s^q-cycles was reduced to calculating the number of Eulerian circuits in the respective De Bruijn graph $G_{s,q}$.

2.2 COUNTING EULERIAN CIRCUITS IN DIGRAPHS

A digraph D is called *Eulerian* if there exists a closed spanning walk W traversing every arc, in D, exactly once and consistently with its orientation; under this, the number of arcs entering any vertex of D equals the number of arcs emanating from it. The mentioned closed walk W, in D, is called an *Eulerian circuit*. The circular order of arcs in an Eulerian circuit is of value because one and the same Eulerian digraph D admits more than one Eulerian circuit whenever the order of circularly touring its arcs may be varied. The last circumstance plays a crucial role when Eulerian circuits formalize the cyclical motion of particles in the respective models of statistical physics, where every possible closed walk of a particle must necessarily be taken into account [8]. All the above can readily be adapted to undirected graphs if one considers every edge as a pair of opposite darts. In the last sense, any connected undirected graph G admits at least one Eulerian circuit passing along every edge strictly twice and just in opposite directions.

The adjacency matrix of an unweighted digraph D with n vertices is an $n \times n$ matrix $C = C(D) = \{c_{ij}\}_{i,j=1}^{n}$ of zeros and ones, wherein an entry $c_{ij} = 1$ iff (if and only if) there is an arc ij (or a self-loop ii, if $i = j$) that goes out of a vertex i and enters a vertex j of D (see [15–18; 21, 22; 7, 8]). Another matrix pertaining to D is its *Laplace, Kirchhoff, or admittance*, matrix $T = T(D) = \{t_{ij}\}_{i,j=1}^{n}$, whose entries are defined as follows (see [15–18]):

$$t_{ij} = \begin{cases} c_{ij}, & \text{if } i \neq j; \text{ and} \\ c_{ii} - d^{+}(i), & \text{if } i = j. \end{cases}$$

Thus, the sum of entries in each column of T equals 0. Here, we do not consider an equivalent version T^* of T, wherein similar manipulations involve the columns of the original matrix C, instead. The reader can consider T^* on his/her own, as an exercise, substituting the respective in-degrees $d^-(j)$ for the out-degrees $d^+(i)$, in the definition of T above.

Every Laplace matrix $T(D)$ (or $T^*(D)$) of an Eulerian digraph D has the property that all its cofactors T_{ij} (or T^*_{ij}) are equal; moreover, here, $T_{ij} = T^*_{ij}$ as well (see [15–18]). Just in case, we recall that a cofactor T_{ij} is the respective minor, of T, multiplied by $(-1)^{i+j}$, where the mentioned minor is in turn the determinant $\det M_{ij}$ of an $(n-1)\times(n-1)$ matrix M_{ij}, obtained by scoring out the i^{th} row and j^{th} column in T.

The common cofactor $c(D) = T_{ij} = \text{const}$ of the Laplace matrix of an Eulerian digraph D is equal to the number of oriented spanning trees that go out of (or come into) any vertex i of D (see [15–18]).

At this point, we shall cite the famous matrix-tree theorem for graphs (see [15-18]), which was first proven by De Bruijn and van Aardenne-Ehrenfest [19], viz.:

Theorem 1. *The number $\varepsilon(D)$ of Eulerian circuits in a labeled Eulerian digraph D is equal to*

$$\varepsilon(D) = c\prod_{i=1}^{n}(d_i - 1)!\,, \tag{1}$$

where c is the common value of cofactors T_{ij} in T; and $d_i = d^+(i) = d^-(i)$.

Theorem 1 plays a very important role herein due to the following statement that comes hand in hand with it (see [4–7]):

Proposition 2. *The number of maximal Kautz sequences of length $s(s-1)^{q-1}$ over the alphabet A ($|A| = s \geq 2; q \geq 2$) is equal to the number $\varepsilon(H_{s,q-1})$ of Eulerian circuits in the respective Kautz graph $H_{s,q-1}$.*

Proof. (Sketch.) By the definition of a Kautz digraph $H_{s,q-1}$, every arc of it corresponds to a distinct word of length q over the alphabet A; and all these arcs together exactly comprise all $s(s-1)^{q-1}$ possible s-ary words of q letters. Since each Eulerian circuit, in $H_{s,q-1}$, traverses each of its arcs exactly once, it is in one-one correspondence with one Kautz sequence cycle. Hence, we at once arrive at the proof.

Some facts from the Spectral Theory of Graphs [18] are needed for us right now, before beginning the next subsection. Let I denote the identity matrix, that is, a diagonal matrix, whose diagonal entries are all 1s while the other entries are all 0s. The *characteristic polynomial* $P(D;x)$ of a (di-)graph D is the characteristic polynomial of its adjacency matrix $C(D)$ (see [18]); that is,

$$P(D;x) = P(C(D);x) = \det[xI - C(D)].$$

Similarly, the *Laplacian polynomial* of D is defined (see [18]):

$$L(D;x) = P(T(D);x) = \det[xI - T(D)].$$

Herein, we need to employ the spectral method [18] of calculating the common cofactor $c = c(D)$. Since all cofactors of T are equal to c, one can deduce, in particular, that the principal $(n-1)\times(n-1)$ minors of T are all equal to c. From the Spectral Theory of Graphs (or Matrices) [18], it immediately follows that

$$c = c(D) = \frac{1}{n} L'(D;x)\,|_{x=0}\ , \tag{2}$$

where $L'(D;x) = \frac{d}{dx} L(D;x)$.

However, for all regular digraphs (with $d^+(i) = d^-(i) = d = \text{const}$, as we have for Kautz graphs) the Laplacian polynomial $L(D;x)$ can readily be calculated through the respective characteristic polynomial as follows:

$$L(D;x) = P(D;x+d). \tag{3}$$

Therefore, we arrive at an equivalent result, earlier derived for multigraphs by Hutschenreuther [20] (see p. 39 in [18]), viz.:

Proposition 3. *The common value c of the cofactors T_{ij} in T can be calculated as*

$$c = c(D) = \frac{1}{n} P'(D;x)\,|_{x=d} \tag{4}$$

We shall use this result in the next subsection.

2.3 SPECTRAL PROPERTIES OF THE ARC-GRAPH

Part of the information about the properties of the arc-graph will be borrowed by us from our previous paper [8]; other properties will be proven directly in this subsection.

Let $D = D(V, E)$ be a digraph with the set V of vertices and set E of arcs (self-loops, if any, are considered as self-adjacent arcs whose head and tail coincide); $|V| = n$, $|E| = m$. The arc-graph $\mathit{\Gamma}(D) = \mathit{\Gamma}(E, U)$ of a digraph D is a derivative digraph whose vertex set $V(\mathit{\Gamma})$ is the set E of arcs of D; each ordered pair ij and kl of arcs, of D, is a pair of adjacent vertices in $\mathit{\Gamma}$ iff the head j of ij coincides with the tail k of kl $(j = k)$, whether the remaining tail i and head l coincide or not.

For the sake of completeness, note that the arc-graph $\mathit{\Gamma}(H)$ of an undirected graph $H = H(V, E)$ can also be constructed if we initially replace each edge ij with a pair of opposite darts $(1 \leq i, j \leq |V| = n; |E| = m)$, which results in the so-called symmetric digraph $S = S(H) = S(V, E^*)$ $(|E^*| = 2|E| +$ number of self-loops, if any), and then revert to the above pattern.

Rosenfeld [8] obtained the following general result:

Theorem 4. *Let $P(G; x)$ and $P(\mathit{\Gamma}(G); x)$ be the characteristic polynomial of an arbitrary weighted (di-)graph G and that of its arc-graph $\mathit{\Gamma}(G)$, respectively. Then*

$$P(\mathit{\Gamma}(G); x) = x^{m-n} P(G; x), \tag{5}$$

where m and n are the numbers of vertices in $\mathit{\Gamma}(G)$ and G, respectively.

In other words, the spectra of $\mathit{\Gamma}(G)$ and G may differ only in the number of zero eigenvalues and this difference in the multiplicities is $|m - n|$.

The Greek character "$\mathit{\Gamma}$" in "$\mathit{\Gamma}(G)$" can be considered as an operator $\mathit{\Gamma}$ transforming a graph G into another one $\mathit{\Gamma}(G)$. This operator has some remarkable properties.

In particular, it can give for any Eulerian digraph G with not less than 2 proper arcs ij $(i \neq j)$ out-going from each of its vertices i, and an arbitrary number $\# \geq 0$ of self-loops, an infinite series of such digraphs: $\Gamma^0(G) := G, \Gamma^1(G) = \Gamma(G), \Gamma^2(G) = \Gamma\left(\Gamma^1(G)\right), \ldots, \Gamma^{q+1}(G) = \Gamma(\Gamma^q(G))$ $(q \geq 0)$, whose spectra differ only in the number of zero eigenvalues.

The reader familiar with [4–7] can immediately see that an instance of the last series $\{\Gamma^q(G)\}_{q=1}^{\infty}$ of digraphs is the series $\mathcal{H}_s$ of the Kautz graphs, whose original definition obeys the same Γ-constructive property (see above). In other words, this is tantamount to the following statement:

Proposition 5. *The series $\mathcal{H}_s = \{H_{s,q}\}_{q=1}^{\infty}$ $(s \geq 2)$ of Kautz digraphs is a recurrent sequence of digraphs, wherein $H_{s,1}$ is a s-vertex complete digraph, without self-loops, and $H_{s,q+1} = \Gamma(H_{s,q})$.*

Proof. To prove it, one should compare the criteria of the adjacency of arcs in $H_{s,q}$, reconsidered as vertices of $H_{s,q+1}$, given in [1–3] and in [8]. Since the two criteria coincide for constructing all the graphs $H_{s,q+1}$ in $\mathcal{H}_s$, the proof is immediate.

Now we can readily calculate the characteristic polynomial of a digraph $H_{s,q-1}$ $(s, q \geq 2)$; the solution will be stated as

Lemma 6. *The characteristic polynomial of $H_{s,q-1}$ is*

$$P(H_{s,q-1}; x) = x^{s(s-1)^{q-2}-s}(x - s + 1)(x + 1)^{s-1} \qquad (s, q \geq 2). \tag{6}$$

Proof. By virtue of Proposition 5, the repetitive application of Theorem 4 demonstrates that every digraph $H_{s,q-1}$ $(s, q \geq 2)$ possesses exactly s nonzero eigenvalues: one eigenvalue $\lambda = s - 1$ and $s - 1$ equal eigenvalues $\lambda = -1$ (namely, those of the complete s-vertex digraph $H_{s,1}$, without self-loops). Since the number of vertices in a digraph $H_{s,q-1}$ is equal to $s(s-1)^{q-2}$, it possesses exactly $s(s-1)^{q-2} - s$ zero eigenvalues. Considering all $s(s-1)^{q-2}$ eigenvalues together, we at once arrive at the proof.

Proposition 3 and Lemma 6 immediately afford, as their common corollary, the following

Lemma 7. *The common value* c *of cofactors* $T_{ij}(H_{s,q-1})$ *in a Laplace matrix* $T(H_{s,q-1})$ *of a digraph* $H_{s,q-1}$ *is equal to*

$$c = c(G_{s,q}) = s^{s-2}(s-1)^{s(s-1)^{q-2}-q-s+2} \qquad (s, q \geq 2). \tag{7}$$

Proof. First, calculate $P'(H_{s,q-1}; x)$, using the R.H.S. of (6) for it:

$$\begin{aligned} [x^{s(s-1)^{q-2}-s}(x-s+1)(x+1)^{s-1}]' &= \\ x^{s(s-1)^{q-2}-s-1}\{[s(s-1)^{q-2}-s](x-s+1)(x+1)^{s-1} &+ \\ (s-1)x(x-s+1)(x+1)^{s-2} + x(x+1)^{s-1}\}. \end{aligned}$$

Hence, under $x = s-1$, Proposition 3 gives

$$\begin{aligned} \frac{1}{s(s-1)^{q-2}} P'(H_{s,q-1}; x) &= \frac{1}{s(s-1)^{q-2}} (s-1)^{s(s-1)^{q-2}-s-1} [0(x+1)^{s-1} + \\ 0x(x+1)^{s-2} + s^{s-1}(s-1)] &= s^{s-2}(s-1)^{s(s-1)^{q-2}-q-s+2}, \end{aligned}$$

which is the proof.

Another important property of the operator Γ is that Γ "unties" every Eulerian circuit θ of a graph $\Gamma^q(G)$, transferring it into an oriented cycle $\Gamma(\theta)$ of Γ^{q+1} $(q \geq 0)$ with the same weight $w(\Gamma(\theta)) = w(\theta)$. Here, we recall that the weight $w(\sigma)$ of any cycle σ, in an arbitrary digraph D, is the product of the weights of arcs comprising σ. Moreover, Γ assures one-to-one correspondence between the set of all Eulerian circuits on G and the set of all oriented cycles of $\Gamma(G)$.

We shall also present a partial result for the tail coefficient of the permanental polynomial $P^+(\Gamma(G); x)$ of the arc-graph $\Gamma(G)$ of an Eulerian digraph G. In particular, G may be the above symmetric digraph $S(H)$ and, consequently, the arc-graph $\Gamma(H)$ of an undirected graph H can also be considered below in place of $\Gamma(G)$. The reader interested in calculating the tail coefficient for all sorts of weighted Eulerian (di-)graphs can see [8], where this problem was completely resolved.

First, it is worth recalling that the permanental polynomial $P^+(H;x)$ of a weighted digraph H is the permanental polynomial $P^+(C(H);x)$ of its adjacency matrix $C(H)$; herein, $P^+(C(H);x) = \mathrm{per}[xI + C(H)]$, where I is a diagonal identity matrix (see p. 34 in [18]). Thus, the tail coefficient of $P^+(C(H);x)$ is simply $\mathrm{per}C(H)$ of the adjacency matrix $C(H)$. Below, we shall derive a corollary of the general weighted version that was proven by Rosenfeld [8], viz.:

Proposition 8. *Let $C(\mathit{\Gamma}(G))$ be the adjacency matrix of the arc-graph $\mathit{\Gamma}(G)$ of an unweighted Eulerian digraph G. Then*

$$\mathrm{per}[C(\mathit{\Gamma}(G))] = \prod_{i=1}^{n} d_i!\,, \tag{8}$$

where d_i stands for the out-degree of a vertex i in G; and the product of factorials $d_i!$ is taken over all (indices of) vertices of G.

We want to specially introduce the definition of Eulerian subcircuit because it may otherwise seem ambiguous. Namely, an *Eulerian subcircuit* of a digraph D is the Eulerian circuit of its Eulerian subgraph $D_1 \subseteq D$ that takes into account exactly one circular order in which all arcs of D_1 can be traversed. In general, there may be more than one circular order for passing all arcs of D_1; therefore, the number of Eulerian subcircuits corresponding to D_1 may be more than 1.

Graph-theoretically, $\mathrm{per}[C(\mathit{\Gamma}(G))]$ is the number of ways in which all arcs of G can be covered by its arc-disjoint Eulerian subcircuits (see [8]). To facilitate referring to this fact in the subsequent text, we shall derive the following working corollary of the last proposition:

Corollary 8.1. *Let $H_{s,q-1}$ $(s, q \geq 2)$ be a Kautz digraph. Then the number of ways in which all $s(s-1)^{q-1}$ arcs of $H_{s,q-1}$ can be covered by Eulerian subcircuits is $[(s-1)!]^{s(s-1)^{q-2}}$.*

Proof. Setting the values $d_i = s - 1$ and $n = s(s-1)^{q-2}$ in (8) at once affords the proof.

Also, due to the above "untying" properties of the operator Γ, the permanent per$[C(\Gamma(G))]$ is the number of spanning cycle covers of Γ (that collectively cover all vertices of Γ). Therefore, we can end this subsection by formulating another corollary, viz.:

Corollary 8.2. *Let $H_{s,q}$ $(s, q \geq 2)$ be a Kautz digraph. Then the number of ways in which all $s(s-1)^{q-1}$ vertices of $H_{s,q}$ can be covered by oriented cycles is $[(s-1)!]^{s(s-1)^{q-2}}$.*

Proof. Recalling that all $s(s-1)^{q-1}$ vertices of $H_{s,q}$ are exactly all arcs of $H_{s,q-1}$ and applying Corollary 8.1 to the last digraph, we immediately arrive at the proof.

At this point, it is time to summarize the tack which will be followed by us, in the next section.

2.4 OUR TACK

We shall keep the general ideas expounded in [1–3], according to which the enumeration of Kautz $[s(s-1)^{q-1}]$-cycles, can be reduced to counting the number of Eulerian circuits in the respective Kautz graph $H_{s,q-1}$ $(s, q \geq 2)$. Under this, we shall employ our recent results concerning the spectral properties of iterated arc-graphs [8], which are exemplified herein by the Kautz graphs. It will enable us to obtain the overall solution for all $s, q \geq 2$. We also plan to discuss some related combinatorial problems, in Miscellaneous.

3. MAIN RESULTS

We at once begin this section with its master theorem:

Theorem 9. *For integers $s, q \geq 2$, there are exactly $s^{s-2}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s+q-2}$ Kautz sequences (cycles) of maximal length $s(s-1)^{q-1}$.*

Proof. By virtue of Lemma 7, c on the R.H.S. of (1) is equal to $s^{s-2}(s-1)^{s(s-1)^{q-2}-q-s+2}$ (see the R.H.S. of (7)); and, by definition of the Kautz digraphs $H_{s,q-1}$, the degree $d_i = s-1$ $(1 \leq i \leq s(s-1)^{q-2})$. With these specific values on the R.H.S. of (1), Theorem 1 gives

$$\begin{aligned}\varepsilon(H_{s,q-1}) &= s^{s-2}[(s-1)!]^{s(s-1)^{q-2}-s-q+2}[(s-2)!]^{s(s-1)^{q-2}} \\ &= s^{s-2}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s+q-2}.\end{aligned}$$

But, by virtue of Proposition 2, $\varepsilon(H_{s,q-1})$ is also the number of Kautz $[s(s-1)^{q-1}]$-cycles, whence the proof is immediate.

Here, we can also calculate the number of linear Kautz sequences, viz.:

Corollary *For integers $s, q \geq 2$, there exist exactly $s^{s-1}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s-1}$ linear Kautz sequences of maximal length $s(s-1)^{q-1}+q-1$.*

Proof. It immediately follows from Theorem 9 and the definition of a linear Kautz sequence. Namely,

$$s(s-1)^{q-1} \cdot s^{s-2}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s+q-2} = s^{s-1}[(s-1)!]^{s(s-1)^{q-2}}/(s-1)^{s-1},$$

as it was stated above.

Here, we can also calculate the number of Kautz $s(s-1)^{q-1}$-sets due to the following theorem:

Theorem 10. *For integers $s, q \geq 2$, the number of Kautz $s(s-1)^{q-1}$-sets is equal to $[(s-1)!]^{s(s-1)^{q-2}}$.*

Proof. This generalizes the Proof of Theorem 9, where the number of Kautz $s(s-1)^{q-1}$-cycles is calculated as the number $\varepsilon(H_{s,q-1})$ of Eulerian circuits of a Kautz graph $H_{s,q-1}$. Now, in lieu of that, we should consider the number of all possible covers of all arcs of $H_{s,q-1}$ by its Eulerian subcircuits. But the last number is given by Corollary 8.1 as $[s(s-1)!]^{s(s-1)^{q-2}}$. Hence, we immediately arrive at the proof.

Moreover, Corollary 9.1 and Theorem 10 enable one to uncover the following interesting fact, viz.:

Corollary 10.1. *For every integer $s \geq 2$, the ratio of the number of linear Kautz sequences of maximal length $s(s-1)^{q-1}+q-1$ to the number of Kautz $[s(s-1)^{q-1}]$-sets is independent of q and equals $s^{s-1}/(s-1)^{s-1}$.*

It is worth mentioning that the respective ratio in the case of De Bruijn linear sequences and De Bruijn s^q-sets even simply equals 1 [8] (see Subsection 4.1, below). Thus, there is some hope that there may also exist like rations for 'sequences' and 'sets' with different adjacency restrictions on characters, which similarly depend only on the cardinality s of the alphabet A and not on q. If it is really true such a fact can well be of use in practical calculations (see [9–14]). But at this point, we must stop our consideration of this topic and turn to discussing other combinatorial problems that, however, resemble by their appearance the above ones.

4. MISCELLANEOUS

This section is a small compilation that seems to be close to the main text, done at the author's choice. It is a mere discussion of known results and methods [4–7; 21–30] but contains, at the end, some open problems that can be proposed to the reader.

4.1. BE BRUIJN SEQUENCES

A cycle is a sequence $a_1a_2\cdots a_r$ taken in a circular order—that is, a_1 follows a_r, and $a_2\cdots a_ra_1,\ldots,a_ra_1\cdots a_{r-1}$ are all the same cycle as $a_1a_2\cdots a_r$. Given natural numbers $q \geq 1$ and $s \geq 2$, a cycle of s^q letters is called a *complete cycle* [4, 5], or *De Bruijn sequence*, if subsequences $a_i a_{i+1} \cdots a_{i+q-1}$ $(1 \leq i \leq s^q)$ consist of all possible s^q ordered sequences $b_1 b_2 \cdots b_q$ over the alphabet A $(|A| = s)$.

In 1946, De Bruijn [4] (see [5]) proved his famous theorem:

Theorem 11. *For $s = 2$ and each positive integer q there are exactly $2^{2^{q-1}-q}$ complete cycles of length 2^q.*

In particular, for $q = 1, 2, 3$, there exist the following complete cycles:

$$\begin{array}{ll} q = 1, & 01, \\ q = 2, & 0011, \\ q = 3, & 00010111, \\ & 00011101. \end{array}$$

Apparently, cutting a complete s^q-cycle $(q \geq 2)$ in all s^q positions generates s^q distinct words since any such cycle is circularly asymmetric, by definition. However, every s^q-word obtained in this fashion contains only $s^q - q + 1$ basic subwords of length q, out of those belonging to the complete cycle. A minimal word of length $s^q + q - 1$ that incorporates just the same set of s^q basic q-subwords as an intact complete cycle is called a *linear De Bruijn sequence*. Obviously, a linear De Bruijn sequence can be obtained by adding the first $q - 1$ letters of any s^q-word, obtained by cutting a complete cycle, to the end of this word.

The following result can be regarded as a corollary of De Bruijn's theorem:

Corollary 11.1. *For $s = 2$ and each positive integer q there are exactly $2^{2^{q-1}}$ linear De Bruijn sequences of length $2^q + q - 1$.*

As a brief illustration, we shall consider the cases $q = 1$ and 2, as these follow from the above example for circular De Bruijn sequences:

$$\begin{array}{ll} q = 1, & 01, \\ & 10, \\ q = 2, & 00110, \\ & 01100, \\ & 11001, \\ & 10011. \end{array}$$

Another generalization of complete cycle is a *De Bruijn s^q-set* of sequences which are not De Bruijn sequences on their own, except for the case when a De Bruijn set consists of exactly one De Bruijn sequence, but collectively have the same aggregated length s^q and also produce the same set of all s-ary words of length q; see Theorem 12 and Corollary 12.1, below ([7]).

Following [4–7], we need to define the series $\mathcal{G}_s = \{G_{s,q}\}_{q=1}^{\infty}$ $(s \geq 2)$ of special digraphs that were used by us in the previous proof [7]; here, the numbers s and q have the same interpretation as above. Initially, we set $G_{s,1}$ to be a one-vertex graph possessing s self-loops. The set $V_{s,q}$ of vertices of a digraph $G_{s,q}$ $(q \geq 2)$ consists of all s^{q-1} ordered sequences, or words, of $q-1$ letters over the alphabet A while the set E of arcs (and self-loops) is in one-one correspondence with all s^q words of q letters over A. Under this, the arc uv labeled by a word $a_1a_2\cdots a_{q-1}a_q$ emanates from a vertex $u = a_1a_2\cdots a_{q-1}$ and enters a vertex $v = a_2\cdots a_{q-1}a_q$. In other words, arcs $a_1a_2\cdots a_{q-1}a_q$ and $a_2a_3\cdots a_qa_{q+1}$ share a common incident vertex $a_2a_3\cdots a_{q-1}a_q$. It is easy to see that the arc set $E_{s,q}$ of a digraph $G_{s,q}$ is simultaneously the vertex set $V_{s,q+1}$ of the next digraph $G_{s,q+1}$, in $\mathcal{G}_s$ (see [4–7]). But what is rather more important, $G_{s,q+1}$ $(q \geq 1)$ can be obtained from $G_{s,q}$ by the process that can locally be called taking the arc-graph $\Gamma(G_{s,q})$ of a digraph $G_{s,q}$ (see [8]); under this, $G_{s,q+1} = \Gamma(G_{s,q})$. The members of the series $\mathcal{G}_s$ were called in [6] (see [3]) *De Bruijn graphs*. Herein, we shall adapt the methods applied in [4–7], wherein estimating the number of complete s^q-cycles was reduced to calculating the number of Eulerian circuits in the respective De Bruijn graph $G_{s,q}$.

The present author proved the following generalizations of Theorem 11 and Corollary 11.1 (see [7]).

Theorem 12. *For positive integers $s \geq 2$ and $q \geq 1$ there are exactly $(s!)^{s^{q-1}-q}$ complete cycles of length s^q.*

Theorem 12 gives, as its elementary corollaries, De Bruijn's Theorem (Theorem 11, herein) and Corollary 11.1. Moreover, we can formulate here "the generalized Corollary 11.1", viz.:

Corollary 12.1. *For positive integers* $s \geq 2$ *and* $q \geq 1$ *there exist exactly* $(s!)^{s^{q-1}}$ *linear De Bruijn sequences of length* $s^q + q - 1$.

We can also calculate the number of De Bruijn s^q-sets due to the following theorem:

Theorem 13. *For positive integers* $s \geq 2$ *and* $q \geq 1$*, the number of De Bruijn* s^q*-sets is equal to* $(s!)^{s^{q-1}}$.

Thus, one can come to the following common corollary of Corollary 12.1 and Theorem 13:

Corollary 13.1. *For positive integers* $s \geq 2$ *and* $q \geq 1$ *the number of linear De Bruijn sequences of length* $s^q + q - 1$ *equals the number of De Bruijn* s^q*-sets:* $(s!)^{s^{q-1}}$.

In our opinion, such a coincidence may lead to new like findings concerning De Bruijn sequences and/or their generalizations. But, here, we must stop our consideration of these and turn to discussing other combinatorial problems that, however, resemble by their appearance all the above ones.

4.2. OTHER SEQUENCES WITH ADJACENCY RESTRICTIONS

The enumeration of s-ary circular sequences of length q is tantamount to the enumeration of q-bead necklaces with s kinds of beads (and other combinatorial restrictions, if any). Namely, the latter interpretation was adopted by Lloyd [24] for enumerating s-ary q-sequences with any given restrictions put on the adjacency of different ciphers (which can be adjacent or self-adjacent and which not). However, it should be noted that his calculation always considers two circular sequences as equal if one can be obtained from the other by reading the original sequence in the opposite direction. The instances of De Bruijn and Kautz sequences, however, do not admit such reversing of the circular order. Nevertheless, the work of Lloyd [24] is of paramount importance for chemists whenever they want to know the number of cyclic substitutional isomers with a given number of sorts of substituents and specified adjacency restrictions put on them.

In Graph Theory, it is known that the number of colorations of a labeled l-cycle ($l \geq 2$) with s colors provided that no two adjacent vertices are colored the same color is equal to $(s-1)^l + (-1)^l(s-1)$; see Theorem IX.23 in [17]. The respective result for a labeled path spanning l vertices is $s(s-1)^{l-1}$; see Theorem IX.24 in [17]. Here, we recall that labeled graphs take into account no symmetry whatever, even if they possess it. Nevertheless, using the so-called inclusion/exclusion principle (see [21, 22, 29]) enables one to utilize the results concerning labeled graphs for enumerating the colorations of the respective unlabeled graphs with a given group of automorphisms (or symmetry group) [21, 22, 29], and even with a given monoid (semigroup) of endomorphisms [29]. From the general combinatorial point of view, the same procedure works equally well for s-ary sequences, too.

From among other sequences, we shall pick herein only few [25–28; 30]. In particular, [25, 26] investigate the *square-free words*; in these, subwords BBb, wherein B is an arbitrary block and b is the first letter of it, are forbidden. Overlapping words and special circular codes have been considered in [27] and [28], respectively. Finally, uncancelable sequences of the elements of a finite regular monoid R that exclude subsequences of type aba, wherein a and b are inverses in R, are of use in an algebraic treatment of genomic sequences [30], proposed by the present author.

Now we shall turn to posing some problems that follow from the whole text above.

4.3. OPEN PROBLEMS

The following problems will represent only a very small part of the problems that could be posed in such a case.

Problem 1. To enumerate s-ary Kautz sequences of length l; $s, q \geq 2$ and $2 \leq l \leq s(s-1)^{q-1}$.

Problem 2. To enumerate s-ary Kautz sequences without subwords of type aba; $s, q \geq 2$.

Problem 3. To enumerate s-ary circular sequences of length l $(1 \leq l < s^q)$ that are included in all De Bruijn s^q-sets with fixed positive integers s and q; $s \geq 2$ and $q \geq 1$.

Problem 4. To enumerate subsequences of length l $(1 \leq l < s^q + q - 1)$ of all linear s-ary De Bruijn sequences of length $s^q + q - 1$ with fixed positive integers s and q; $s \geq 2$ and $q \geq 1$.

Some other sequences, whose consideration is omitted herein, are planned to be considered in our next publications.

Acknowledgements: I am grateful to Dr. Valery Kirzhner, Profs. Edward Trifonov and Alexander Bolshoy for stimulating the idea of preparing this paper. Also, I sincerely thank my anonymous Reviewer for his/her expert work.

References

[1] W. H. Kautz, in: Design of Optimal Interconnection Networks for Multiprocessors, Architecture and Design of Digital Computers, Nato Advanced Summer Institute (1969), 249–272.

[2] M. A. Fiol, I. Alegre and J. L. A. Yerra, *Line Digraph Iterations and the* (d, k) *Problem for Directed Digraphs,* Proc. 10th Int. Symp. Comp. Arch. (1983), 174–177.

[3] J. L. Villar, *Kautz s-Ary Closed Sequences,* in Combinatorics '88, Proceedings of the International Conference on Incidence Geometries and Combinatorial Structures, Ravello, Italy, May 23–28, 1988. (A. Barlotti, G. Lunardon, F. Mazzocca, N. Melone, D. Olanda, A. Pasini and G. Tallini, eds.), Mediterranean Press, **2** (1991), 459–469.

[4] N. G. De Bruijn, *A Combinatorial Problem,* Nederl. Akad. Wetensch. Proc. 1946, v. 49, p. 758–764, Indagationes Math. **8** (1946), 461–467.

[5] M. Jr. Hall, *Combinatorial Theory*, John Wiley and Sons, Inc., New York, (1967), 91–99.

[6] A. Lempel, *m-Ary Closed Sequences,* J. Comb. Theory, **10** (1971), 253–258.

[7] V. R. Rosenfeld, *Enumerating De Bruijn Sequences,* Commun. Math. Comput Chem. (MATCH), **45** (2002), 71–83.

[8] V. R. Rosenfeld, *Some Spectral Properties of the Arc-Graph,* Commun. Math. Comput. Chem. (MATCH), **43** (2001), 41–48.

[9] E. N. Trifonov, *Making Sense of the Human Genome,* in Structure and Methods, v. 1, *Human Genome Initiative and DNA Recombination*, Adenine Press, N. Y., (1990), 69–76.

[10] E. N. Trifonov, *Informational Structure of Genetic Sequences and Nature of Gene Splicing,* in Advances in Biomolecular Simulations (E. N. Lavery, J.-L. Rivail and J. Smith, eds.), AIP Conference Proceedings 239, N. Y., (1991), 329–338.

[11] O. Popov, D. M. Segal and E. N. Trifonov, *Linguistic Complexity of Protein Sequences As Compared to Texts of Human Languages,* Biosystems, **38** (1996), 65–74.

[12] A. Bolshoy, K. Shapiro, E. N. Trifonov and I. Ioshiknes, *Enhancement of the Nucleosomal Pattern in Sequence of Lower Complexity,* Nucleic Acids Research **25** (16) (1997), 3248–3254.

[13] A. Gabrielian and A. Bolshoy, *Sequence Complexity and DNA Curvature,* Comput. Chem., **23** (1999), 263–274.

[14] O. G. Troyanskaya, O. Arbell, Y. Koren, G. M. Landau and A. Bolshoy, *Sequence Complexity Profiles of Prokaryotic Genomic Sequences: A Fast Algorithm for Calculating Linguistic Complexity,* Bioinformatics, (2002), submitted.

[15] F. Harary, *Graph Theory*, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1969.

[16] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York and London, 1973.

[17] W. T. Tutte, *Graph Theory*, Addison-Wesley, 1984.

[18] D. M. Cvetković , M. Doob and H. Sachs, *Spectra of Graphs: Theory and Application*, Academic Press, Berlin, 1980.

[19] N. G. De Bruijn and T. van Aardenne-Ehrenfest, *Circuits and Trees in Oriented Graphs,* Simon Stevin, **28** (1951), 203–217.

[20] H. Hutschenreuther, *Einfacher Beweis des Matrix-Gerüst-Satzes der Netzwerktheorie,* Wiss. Z. TH Ilmenau, **13** (1967), 403–404.

[21] A. Kerber, *Algebraic Combinatorics via Finite Group Actions*, Wissenshaftsverlag, Manheim, Wien, Zürich, 1991.

[22] A. Kerber, *Applied Finite Group Actions*, Springer-Verlag, Berlin, Heidelberg, New York, 1999.

[23] A. T. Balaban and F. Harary, *Chemical Graphs IV (Aromaticity VI): Dihedral Groups and Monocyclic Aromatic Compounds,* Rev. Roumaine Chim., **12** (1967), 1511–1515.

[24] E. K. Lloyd, *Necklace Enumeration with Adjacency Restrictions,* in Combinatorics. The Proceedings of the British Combinatorial Conference held in the University College of Wales, Aberystwyth, 2–6 July 1973 (T. P. McDonough and V. C. Mavron, eds.), Cambridge University Press, Cambridge, 1974, p. 97–102.

[25] E. D. Fife, *Irreducible Binary Sequences,* in Combinatorics on Words—Progress and Perspectives (L. J. Cumming, ed.), Academic Press, Toronto, New York, p. 91–100.

[26] R. O. Shelton, On the Structure and Extendibility of Square-Free Words, ibid., p. 101–118.

[27] J. Almeida, *Overlapping of Words in Rational Languages,* ibid., p. 119–131.

[28] J. Berstel and D. Perrin, Codes Circulaires, ibid., p. 133–165.

[29] V. R.Rosenfeld, *Yet Another Generalization of Pólya's Theorem: Enumerating Equivalence Classes of Objects with a Prescribed Monoid of Endomorphisms,* Commun. Math. Comput. Chem (MATCH), (43) (2001), 111–130.

[30] V. R. Rosenfeld, *An Algebraic Model of Closed Loops in Proteins,* Commun. Math. Comput. Chem. (MATCH), submitted.