Constructing orientable sequences

(Recent) results and open questions

Chris J Mitchell

Royal Holloway, University of London

me@chrismitchell.net
www.chrismitchell.net

February 2025

1. Introduction: What are orientable sequences?

- I'm sure you're familiar with de Bruijn sequences, i.e. infinite periodic sequences of elements from $\{0, 1, ..., k-1\}$ in which every possible k-ary n-tuple occurs exactly once in a period.
- The period must be k^n, and there are many known methods of construction.
- Earliest known reference to constructing (and enumerating) such sequences is due to Sainte-Marie (1894), but better known work is by de Bruijn (1946) and Good (1947).
- Examples for $k = 2$ are: [0011] ($n = 2$), and [00010111] ($n = 3$).
- There are many applications, for example in stream ciphers, position location, and genome sequencing.
- De Bruijn sequences are examples of n-window sequences, periodic sequences in which any n-tuple occurs *at most once* in a period.

Orientable sequences

- An orientable sequence (an $\mathcal{OS}_k(n)$) is a k-ary n-window sequence with the added property that an n-tuple occurs at most once in a period of a sequence *or its reverse*.
- First introduced in 1992, they have potential application in certain position location applications.
- For the binary case, a simple example for $n = 5$ has period 6 — a single period is [001011].
- The sequence and its reverse contain twelve distinct 5-tuples: 00101, 00110, 01001, 01011, 01100, 01101, and the complements of these 5-tuples.
- Examples for $k = 3$ are: [012] ($n = 2$) and [001201122] ($n = 3$).

Information Security Department

2. Upper bounds on the period

- Since any n-tuple can only occur once in a period in either direction, and symmetric n-tuples cannot occur, a trivial bound on the period of an $\mathcal{OS}_k(n)$ is
$$\frac{k^n - k^{\lfloor (n+1)/2 \rfloor}}{2}.$$
- However, apart from when $n = 2$ and k is odd, this bound is not sharp.
- The binary case is different from $k > 2$ — in particular, constant $(n-1)$-tuples and $(n-2)$-tuples cannot occur in a binary sequence, whereas they can for $k > 2$.
- This means that an $\mathcal{OS}_2(n)$ cannot exist for $n < 5$.
- Dai, Martin, Robshaw & Wild (1993) gave a bound for the binary case which is significantly sharper than the trivial bound.

The Dai-Martin-Robshaw-Wild upper bound ($k = 2$)

Suppose S is an $\mathcal{OS}_2(n)$ ($n \geq 5$). Then the period of S is at most:

$$\begin{array}{rll} 2^{n-1} - 41/9 \times 2^{n/2-1} + n/3 + 16/9 & \text{if} & n \equiv 0 \pmod 4 \\ 2^{n-1} - 31/9 \times 2^{(n-1)/2} + n/3 + 19/9 & \text{if} & n \equiv 1 \pmod 4 \\ 2^{n-1} - 41/9 \times 2^{n/2-1} + n/6 + 20/9 & \text{if} & n \equiv 2 \pmod 4 \\ 2^{n-1} - 31/9 \times 2^{(n-1)/2} + n/6 + 43/18 & \text{if} & n \equiv 3 \pmod 4 \end{array}$$

Dai et al. upper bound values ($k = 2$)

Order (n)	Maximum period	Maximum period (simple bound)
5	6	14
6	17	28
7	40	60
8	96	120
9	206	248
10	443	496

The naive bound is given for comparison purposes.

A general bound

- We can establish a bound for the $k > 2$ case which is a little sharper than the trivial bound (Alhakim, Mitchell, Szmidt & Wild, 2024).
- Suppose that $S = (s_i)$ is an $\mathcal{OS}_k(n)$ ($k \geq 2$, $n \geq 2$). Then the period of S is at most:

$$(k^n - k^{\lceil n/2 \rceil} - k^{\lceil (n-1)/2 \rceil} + k)/2 \quad \text{if } k \text{ is odd,}$$
$$(k^n - k^{\lceil n/2 \rceil} - k)/2 \quad \text{if } k \text{ is even.}$$

Further, if k is odd and $n \geq 6$ then the period of S is at most

$$(k^n - 2k^{n/2} - k(n-2)/2 + 2k)/2 \quad \text{if } n \text{ is even,}$$
$$(k^n - k^{(n+1)/2} - 2k^{(n-1)/2} + k + k^2)/2 \quad \text{if } n \text{ is odd.}$$

Bound values ($k > 2$)

Order (n)	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
2	3	4	10	12	21
3	9	22	50	87	147
4	33	118	290	627	1155
5	105	478	1490	3777	8211
6	336	2014	7680	23217	58464
7	1032	8062	38640	139317	410256
8	3189	32638	194630	839157	2879835

3. Binary sequence constructions

- The first published construction method for orientable sequences is due to Dai et al. (1993).
- It involves joining orientable cycles of length n, where the cycles come in pairs made up of a cycle and its reverse.
- Dai et al. showed using a graph-theoretic argument that is existential rather than constructive that one of every pair of these cycles can be joined to give an orientable sequence.
- The method produces sequences which have asymptotically optimal period.
- As far as I am aware, nothing further was published on these sequences for almost 40 years; however, since 2022, a number of new results have been established.

The de Bruijn digraph

- A construction method for binary orientable sequences (Mitchell and Wild, 2022) relies on a graph homomorphism first described by Lempel in 1970.
- The de Bruijn-Good graph $G_{n,k}$ is a directed graph with vertex set $\{0, 1, \ldots, k-1\}^n$.
- An edge connects $(a_0, a_1, \ldots, a_{n-1})$ to $(b_0, b_1, \ldots, b_{n-1})$ iff $a_{i+1} = b_i$ for every i $(0 \leq i \leq n-2)$.
- If we identify an edge from $(a_0, a_1, \ldots, a_{n-1})$ to $(b_0, b_1, \ldots, b_{n-1})$ with the $(n+1)$-tuple $(a_0, a_1, \ldots, a_{n-1}, b_{n-1})$, then a de Bruijn sequence of order $n+1$ corresponds to an Eulerian circuit in $G_{n,k}$.
- There are, of course, efficient algorithms for finding such circuits.

The Lempel Homomorphism

- The Lempel D-function, originally defined only for $k = 2$, maps $G_{n,2}$ to $G_{n-1,2}$.
- D maps any binary n-tuple $(a_0, a_1, \ldots, a_{n-1})$ to $(a_1 - a_0, a_2 - a_1, \ldots, a_{n-1} - a_{n-2})$.
- D is a graph homomorphism from $G_{n,2}$ onto $G_{n-1,2}$.
- We can extend the notation to allow D to be applied to periodic binary sequences, so D maps the set of periodic binary sequences to itself.
- If S is a sequence of period m, then $D(S)$ will clearly have period dividing m.

Information Security Department

Inverse Lempel and de Bruijn sequences

- Can also define D^{-1} where if S is a periodic binary sequence then $D^{-1}(S)$ is set T of all binary sequences such that $D(T) = S$.
- If S is an n-window sequence then it is straightforward to see that any $(n+1)$-tuple will appear at most once in a period of one of the sequences in $D^{-1}(S)$.
- In the special case where S is a de Bruijn sequence of order n, then $D^{-1}(S)$ contains a complementary pair of sequences, both of period 2^n, in which every $(n+1)$-tuple appears exactly once in a period of one of the sequences.
- As Lempel showed, one of the two sequences will contain the $(n+1)$-tuple $(0101\ldots)$, and the other will contain the $(n+1)$-tuple $(1010\ldots)$, and hence they both contain the n-tuple $(0101\ldots)$.
- They can thus be joined to form a de Bruijn sequence of order $n+1$.

Information Security Department

Inverse Lempel and orientable sequences I

- The Lempel homomorphism can also be applied to generate orientable sequences (Mitchell & Wild, 2022).
- If S is an $\mathcal{OS}_2(n)$ of period m and weight w, then $D^{-1}(S)$ contains either an $\mathcal{OS}_2(n+1)$ of period $2m$ (if w is odd) or a pair of sequences of period m which are 'collectively' orientable (if w is even).
- However, if w is odd, the weight of the $\mathcal{OS}_2(n+1)$ will have weight m, and so even if w is odd and m is odd, the homomorphism can only be applied recursively twice before yielding sequences in pairs rather than the single long sequence desired.

Inverse Lempel and orientable sequences II

- The solution is as follows. Suppose S is an orientable sequence of order n containing exactly one occurrence of 1^{n-4}. If S has even weight then leave it alone; otherwise change 1^{n-4} to 1^{n-3} to make it have odd weight (and the result is still orientable).
- Given a suitable starter sequence S that is an $\mathcal{OS}_k(n-1)$, can guarantee that $D^{-1}(S)$ will be an $\mathcal{OS}_k(n)$ containing exactly one occurrence of 1^{n-4}, and can repeat indefinitely.
- This gives a simple recursive method of generating orientable sequences with large periods.

4. General sequence constructions

- As described by Alhakim et al. (2024), can use the inverse Lempel homomorphism to go from an $\mathcal{OS}_k(n)$ of period m to an $\mathcal{OS}_k(n+1)$ of period km.
- However, it is non-trivial to ensure that D^{-1} yields a single sequence of period km rather than a set of $(n+1)$-tuple-disjoint sequences with periods summing to km.
- Moreover, some variants of the (inverse) Lempel homomorphism only yield 'negative' orientable sequences, in which the collection of all n-tuples and reverse negative n-tuples in a period are all distinct.
- Various approaches have been devised to fix this in recent work by Gabrić & Sawada (2024) and Mitchell & Wild (2024). Gabrić & Sawada showed how to join the multiple cycles produced, and Peter Wild and I constructed 'starter sequences' with special properties enabling repeated use of the Lempel homomorphism.
- Sequences produced by Gabrić & Sawada have asymptotically maximal period.

Information Security Department

A new construction

- The following simple method of construction (Mitchell & Wild, 2025 (unpublished)) involves a subgraph $A_{n,k}$ of the de Bruijn graph $G_{n,k}$.
- As for $G_{n,k}$, the vertices are the k-ary n-tuples.
- An edge connects $(a_0, a_1, \ldots, a_{n-1})$ to $(b_0, b_1, \ldots, b_{n-1})$ iff
  - $a_{i+1} = b_i$ for every i $(0 \leq i \leq n-2)$ (as in the de Bruijn graph); and
  - $b_{n-1} - a_0 \in \{1, 2, \ldots, \lfloor (k-1)/2 \rfloor .\}$
- Every vertex has in-degree and out-degree $\lfloor (k-1)/2 \rfloor$. If $k \geq 5$ then $A_{n,k}$ is connected.
- Analogously to de Bruijn sequences, an Eulerian circuit in $A_{n,k}$ will yield an $\mathcal{OS}_k(n+1)$ of period $k^n \lfloor (k-1)/2 \rfloor$ (for $k \geq 5$), which is greater then $(k-1)/k$ times the upper bound for k odd, and greater then $(k-2)/k$ times the upper bound for k even.
- In fact if $n = 2$ or $n = 3$ and k odd, the period meets the upper bound.

5. Open questions

- Apart from a few small values of n and k, there is a gap between the period of the longest known $\mathcal{OS}_k(n)$ and the best upper bound.
- Also, for a few small values of n and k, exhaustive search has shown that the maximum period is strictly less than the upper bound.
- This suggests further research is needed on two main problems:
  - tightening the upper bounds;
  - constructing sequences with periods closer to the upper bounds;

  so that (ideally) there is no gap.
- Eliminating the gap altogether seems difficult.

Largest known periods for the binary case ($k = 2$)

Order (n)	Maximum period	Maximum period (Dai et al. bound)
5	**6**	6
6	**16**	17
7	**36**	40
8	92	96
9	174	206
10	416	443

- Figures in bold represent maximal lengths as verified by search.
- For further details see the excellent website maintained by Joe Sawada: `http://debruijnsequence.org/db/orientable`

Information Security Department

Largest known periods for $k > 2$

n	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
2	**3** (3)	**4** (4)	**10** (10)	**12** (12)	**21** (21)
3	**9** (9)	**20** (22)	**50** (50)	72 (87)	**147** (147)
4	**30** (33)	84 (118)	275 (290)	522 (627)	1127 (1155)
5	90 (105)	368 (478)	1385 (1490)	3360 (3777)	7756 (8211)
6	285 (336)	1608 (2014)	7155 (7680)	21150 (23217)	56049 (58464)
7	879 (1032)	7308 (8062)	36890 (38640)	135450 (139317)	403389 (410256)
8	2688 (3189)	30300 (32638)	187980 (194630)	821940 (839157)	2844408 (2879835)

- Upper bound values are given in brackets.
- Figures in bold represent maximal lengths.
- As of 6/2/25 I believe I can increase the 72 for $n = 3$, $k = 6$ to 78.

6. Literature

- (Mitchell & Wild, 2022): IEEE Trans on Inf Thy **68** (2022) 4782–4789.
- (Gabrić & Sawada, 2024): arXiv 2401.14341 and 2407.07029.
- (Mitchell & Wild, 2024): arXiv 2409.00672 and 2411.17273.

Other resources

- Joe Sawada's page: `http://debruijnsequence.org/db/orientable`
- The Combinatorial Object Server: `http://combos.org/`