# New constructions for orientable sequences

## 5th Pythagorean Conference, Kalamata, 2–6 June 2025

Chris Mitchell and Peter Wild

5th June 2025

# 1980: Finite Geometries & Designs, UK

# 1. Introduction: What are orientable sequences?

- A $k$-ary de Bruijn sequence of order $n$ is an 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 of order $n$ (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] ($n = 3$).

# The de Bruijn digraph

- The de Bruijn digraph is a key tool for analysing and constructing both de Bruijn and orientable sequences.
- This graph, otherwise known as the de Bruijn-Good graph, $B_k(n)$ 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 $i$ $(0 \leq i \leq n-2)$.
- It is simple to see that $B_k(n)$ is Eulerian, i.e. it is connected and every vertex has in-degree equal to its out-degree.
- 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 $B_k(n)$ — which must exist given $B_k(n)$ is Eulerian.
- There are, of course, efficient algorithms for finding such circuits.

# The Lempel Homomorphism

- The Lempel $D$-function, originally defined only for $k = 2$, maps $B_2(n)$ to $B_2(n-1)$.
- $D$ 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 $B_2(n)$ to $B_2(n-1)$.
- Can extend definition to $k$-ary case, where $D$ maps the $k$-ary $n$-tuple $(a_0, a_1, \ldots, a_{n-1})$ to $(a_1 - a_0, a_2 - a_1, \ldots, a_{n-1} - a_{n-2})$, where computations take place modulo $k$.
- The inverse of $D$ has been widely used, e.g. to recursively construct de Bruijn sequences, observing that $D^{-1}$ maps a circuit in $B_k(n-1)$ to a set of $k$ circuits in $B_k(n)$.

# Upper bounds on the period of orientable sequences

- 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 $k > 2$, so 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.
- A bound for the $k > 2$ case which is a little sharper than the trivial bound was recently established (Alhakim, M, Szmidt & Wild, 2024).

# 2. New upper bounds on the period

- In recent work we have established new upper bounds on the period of a $k$-ary orientable sequence (for $k > 2$), sharper than the 2024 bound.
- These bounds all derive from simple observations regarding the subgraph of the de Bruijn graph defined by the edges of an orientable sequence.
- If $S$ is a $k$-ary orientable sequence of order $n$ — an $\mathcal{OS}_k(n)$ — then we define $B_S$ to be the subgraph of $B_k(n-1)$ with edges corresponding to the $n$-tuples appearing in either $S$ or $S^R$ (where $S^R$ is the reverse of $S$).
- The $n$-tuples appearing in either $S$ or $S^R$ are, of course, all distinct since $S$ is orientable.
- Since $S$ and $S^R$ define edge-disjoint (but not vertex-disjoint) Eulerian circuits in $B_S$, it follows that $B_S$ must be Eulerian.
- This simple observation leads to the improved bounds, given we can identify cases where certain edges cannot occur in $B_S$.

# Degree-parity constraints

- An $n$-tuple $(a_0, a_1, \ldots, a_{n-1})$ is said to be *symmetric* if and only if $(a_0, a_1, \ldots, a_{n-1})$ is a palindrome.
- Both $S$ and $S^R$ correspond to an Eulerian circuit in $B_S$, and these circuits are edge disjoint and cover all the edges of $B_S$.
- If **a** is symmetric then both circuits pass through this vertex equally many times.
- It follows that **a** has even in-degree and even out-degree in $B_S$.
- If $k$ is odd then every vertex in $B_k(n)$ has odd in-degree (and out-degree).
- Hence if $k$ is odd and $s$ is an $\mathcal{OS}_k(n)$ then, for every vertex corresponding to a symmetric $(n-1)$-tuple, at least one incoming edge and at least one outgoing in $B_k(n-1)$ cannot occur in $B_S$.
- This limits the edges that can be contained in $B_S$, and hence upper-bounds its period.

# Semi-symmetry constraints

- An $n$-tuple $(a_0, a_1, \ldots, a_{n-1})$ is said to be *left-semi-symmetric* if and only if $(a_0, a_1, \ldots, a_{n-2})$ is a palindrome.
- E.g. for $n = 5$ and $k = 3$, (02201) is left-semi-symmetric, since 0220 is a palindrome.
- In the de Bruijn digraph $B_k(n)$, one of the edges incoming to such a vertex will be a palindrome, and hence cannot occur in an orientable sequence.
- So, if $S$ is orientable, the in-degree of a vertex corresponding to a left-semi-symmetric tuple in $B_S$ will be less than the maximum, and hence so will the out-degree.
- This limits the edges that can be contained in $B_S$, and hence upper-bounds its period.
- An analogous argument applies to right-semi-symmetric tuples.

# Interactions

- The eagle-eyed amongst you will have immediately spotted that we cannot simply add together the numbers of excluded edges from these arguments as we may be double counting.
- As a result, we need to carefully (and rather painfully) examine a number of special cases.
- In the next slide, the bound resulting from these observations are tabulated for small $k$ and $n$, with the 'old' bound given in brackets for comparison.

# Bounds — new and (old) — on the period of an $\mathcal{OS}_k(n)$

| $n$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ | $k=7$ | $k=8$ |
|---|---:|---:|---:|---:|---:|---:|
| 2 | 3 | 4 | 10 | 12 | 21 | 24 |
| | (3) | (4) | (10) | (12) | (21) | (24) |
| 3 | 9 | 20 | 50 | 84 | 147 | 216 |
| | (9) | (22) | (50) | (87) | (147) | (220) |
| 4 | 30 | 112 | 280 | 612 | 1134 | 1984 |
| | (33) | (118) | (290) | (627) | (1155) | (2012) |
| 5 | 99 | 452 | 1450 | 3684 | 8085 | 15896 |
| | (105) | (478) | (1490) | (3777) | (8211) | (16124) |
| 6 | 315 | 1958 | 7550 | 23019 | 58065 | 130332 |
| | (336) | (2014) | (7680) | (23217) | (58464) | (130812) |
| 7 | 972 | 7844 | 38100 | 138144 | 408072 | 1042712 |
| | (1032) | (8062) | (38640) | (139317) | (410256) | (1046524) |
| 8 | 3096 | 32390 | 193800 | 837879 | 2876496 | 8382492 |
| | (3189) | (32638) | (194630) | (839157) | (2879835) | (8386556) |
| 9 | 9423 | 129572 | 971350 | 5027304 | 20149437 | 67059992 |
| | (9645) | (130558) | (974390) | (5034957) | (20166027) | (67092476) |

# 3. Methods of construction

- 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 M & 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.

# A different approach: Antisymmetric subgraphs of the de Bruijn digraph

- A subgraph $T$ of the de Bruijn digraph $B_k(n)$ is said to be *antisymmetric* if the following property holds.
- Suppose $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})$ and $\mathbf{y} = (y_0, y_1, \ldots, y_{n-1})$ are $k$-ary $n$-tuples, i.e. vertices in $B_k(n)$.
- Then if $(\mathbf{x}, \mathbf{y})$ is an edge in $T$, then $(\mathbf{y}^R, \mathbf{x}^R)$ is *not* an edge in $T$.

# From subgraph to sequence

- If $S$ is an $\mathcal{OS}_k(n)$ of period $m$, then $B_S$ is an antisymmetric Eulerian subgraph of $B_k(n-1)$ containing $m$ edges.
- Antisymmetry follows from the definition of orientable.
- **More importantly**, if $T$ is an antisymmetric Eulerian subgraph of $B_k(n-1)$ with $m$ edges, then there exists an $\mathcal{OS}_k(n)$ $S$ of period $m$ with edge set $T$.
- Why? Since $T$ is Eulerian there exists an Eulerian circuit. This Eulerian circuit corresponds to an $n$-window sequence, which is orientable since $T$ is antisymmetric.

# A simple construction for an antisymmetric subgraph

- Construct the edge set such that an edge connects $(a_0, a_1, \dots, a_{n-1})$ to $(a_1, a_2, \dots, a_n)$ if and only if

$$a_n - a_0 \in \{1, 2, \dots, \lfloor (k-1)/2 \rfloor\}.$$

- Every vertex has in-degree and out-degree $\lfloor (k-1)/2 \rfloor$. If $k \geq 5$ then $T$ is connected, i.e. $T$ is Eulerian.
- $T$ is antisymmetric since every edge $(a_0, a_1, \dots, a_n)$ satisfies $a_n - a_0 \in \{1, 2, \dots, \lfloor (k-1)/2 \rfloor\}$, and hence $-a_0 - (-a_n) = a_0 - a_n \in \{\lfloor (k+2)/2 \rfloor, \lfloor (k+4)/2 \rfloor, \dots, k-1\}$
- Thus $T$ yields an $\mathcal{OS}_k(n+1)$ of period $k^n \lfloor (k-1)/2 \rfloor$ (for $k \geq 5$).
- If $n = 2$, or $n = 3$ and $k$ odd, the period meets the upper bound.

# A small example

Consider the case $k = 5$ and $n = 3$. The 50 3-tuples in $T$ are listed in the table, and a period of an $\mathcal{OS}_5(3)$ containing these 50 3-tuples is:

[00123 40112 23344 00213 24304 21431 03142 03204 10224 41133].

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 001 | 002 | 102 | 103 | 203 | 204 | 304 | 300 | 400 | 401 |
| 011 | 012 | 112 | 113 | 213 | 214 | 314 | 310 | 410 | 411 |
| 021 | 022 | 122 | 123 | 223 | 224 | 324 | 320 | 420 | 421 |
| 031 | 032 | 132 | 133 | 233 | 234 | 334 | 330 | 430 | 431 |
| 041 | 042 | 142 | 143 | 243 | 244 | 344 | 340 | 440 | 441 |

# Antinegasymmetry

- A subgraph $T$ of the de Bruijn digraph $B_k(n-1)$ is said to be *antinegasymmetric* if the following property holds.
- Suppose $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})$ and $\mathbf{y} = (y_0, y_1, \ldots, y_{n-1})$ are $k$-ary $n$-tuples, i.e. vertices in $B_k(n)$.
- Then if $(\mathbf{x}, \mathbf{y})$ is an edge in $T$, then $(-\mathbf{y}^R, -\mathbf{x}^R)$ is *not* an edge in $T$.

# From antinegasymmetry to antisymmetry

- If $T$ is an antinegasymmetric subgraph of the de Bruijn digraph $B_k(n-1)$ with edge set $E$, then $D^{-1}(E)$, of cardinality $k|E|$, is the set of edges for an antisymmetric subgraph of $B_k(n)$, which, abusing our notation slightly, we refer to as $D^{-1}(T)$.
- If every vertex of $T$ has in-degree equal to its out-degree, then the same applies to $D^{-1}(T)$.
- If $T$ is connected and its edge set contains the all-one tuple, then $D^{-1}(T)$ is connected, i.e. in this case if $T$ is Eulerian then so is $D^{-1}(T)$.

# Constructing antinegasymmetric subgraphs

- If $0 \leq u \leq k-1$, set $f(u) = u$ if $u \neq 0$ and $f(u) = k/2$ if $u = 0$.
- Suppose $\mathbf{u} = (u_0, u_1, \ldots, u_{n-1})$ is a $k$-ary $n$-tuple.
- The *pseudoweight* of $\mathbf{u}$ is defined to be the sum

$$w^*(\mathbf{u}) = \sum_{i=0}^{n-1} f(u_i)$$

  where the sum is computed in $\mathbb{Q}$.
- If $E$ is the set of all $k$-ary $n$-tuples with pseudoweight less than $nk/2$, then $E$ is the set of edges for an antinegasymmetric Eulerian subgraph of the de Bruijn digraph $B_k(n-1)$.
- Moreover, $E$ contains the all-one $n$-tuple.
- Hence $D^{-1}(E)$ is a negasymmetric Eulerian subgraph of $B^k(n)$.
- This approach yields orientable sequences with largest possible period for $n = 3$ (all $k$) and $n = 4$ ($k$ odd).

# An example

Suppose $k = 3$ and $n = 3$. The ten 3-ary 3-tuples having pseudoweight less than 4.5 are listed below — this is $E$.

| | | |
|---|---|---|
| 111 | | |
| 011 | 101 | 110 |
| 001 | 010 | 100 |
| 112 | 121 | 211 |

$D^{-1}(E)$ consists of the 30 4-tuples given below.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0120 | 1201 | 2012 | | | | | | |
| 0012 | 1120 | 2201 | 0112 | 1220 | 2001 | 0122 | 1200 | 2011 |
| 0001 | 1112 | 2220 | 0011 | 1122 | 2200 | 0111 | 1222 | 2000 |
| 0121 | 1202 | 2010 | 0101 | 1212 | 2020 | 0201 | 1012 | 2120 |

An $\mathcal{OS}_5(3)$ of period 30 containing these 4-tuples is:

[01201 21202 01012 22011 20011 12200]

.

# 4. Open questions

- Prior to the work described, the only cases where the largest period was known was for $n = 2$ (and a couple of other cases established by exhaustive search).
- The new bounds and new construction methods mean we have now resolved the maximum period question for $n = 3$ (all $k$) and $n = 4$ (odd $k$).
- However, apart these small values of $n$, there is a gap between the period of the longest known $\mathcal{OS}_k(n)$ and the best 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 $k = 2$

| Order ($n$) | Maximum known 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`

# Largest known periods for $k > 2$

Table: Largest known periods for an $\mathcal{OS}_k(n)$ (and bounds)

| $n$ | $k = 3$ | $k = 4$ | $k = 5$ | $k = 6$ | $k = 7$ | $k = 8$ |
|---|---:|---:|---:|---:|---:|---:|
| 2 | **3** | **4** | **10** | **12** | **21** | **24** |
| | (3) | (4) | (10) | (12) | (21) | (24) |
| 3 | **9** | **20** | **50** | **84** | **147** | **216** |
| | (9) | (20) | (50) | (84) | (147) | (216) |
| 4 | **30** | 88 | **280** | 534 | **1134** | 1800 |
| | (30) | (112) | (280) | (612) | (1134) | (1984) |
| 5 | 93 | 372 | 1390 | 3360 | 7763 | 15120 |
| | (99) | (452) | (1450) | (3684) | (8085) | (15896) |
| 6 | 288 | 1608 | 7160 | 21150 | 56056 | 124320 |
| | (315) | (1958) | (7550) | (23019) | (58065) | (130332) |
| 7 | 882 | 7308 | 36890 | 135450 | 403389 | 1034264 |
| | (972) | (7844) | (38100) | (138144) | (408072) | (1042712) |
| 8 | 2691 | 30300 | 187980 | 821940 | 2844408 | 8315496 |
| | (3096) | (32390) | (193800) | (837879) | (2876496) | (8382492) |

- Upper bound values are given in brackets.
- Figures in bold represent maximal lengths.
- As of 25/5/25 we believe we can increase the 288 for $n = 6$, $k = 3$ to 303.

# 5. 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/`