

Perfect necklaces

Verónica Becher

Universidad de Buenos Aires

Workshop EPA!
Buenos Aires, 24 October , 2024

Perfect necklaces

A necklace is the equivalence class of a word under rotations.

Definition (Alvarez, Becher, Ferrari and Yuhjtman 2016)

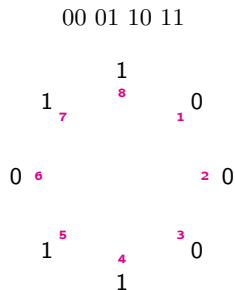
A necklace over a b -symbol alphabet is (n, k) -perfect if each word of length n occurs k times, at positions with different congruence modulo k , for any convention of the starting point.

The (n, k) -perfect necklaces have length kb^n .

De Bruijn circular sequences are exactly the $(n, 1)$ -perfect necklaces.

Example

All words of length 2 concatenated in lexicographical order, view it circularly.



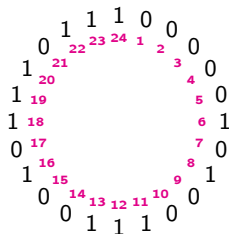
00 occurs twice (p:1,2);
01 occurs twice (p:3,6);
10 occurs twice (p:5,8);
11 occurs twice (p:4,7)

Each word of length 2 occurs 2 times at positions with **different** congruence modulo n .

Example

All words of length 3 concatenated in lexicographical order, view it circularly.

000 001 010 011 100 101 110 111



000 occurs three times (positions 1,2,3)

001 occurs three times (positions 4,9,14)

...

Each word of length 3 occurs n times at positions with **different** congruence modulo 3.

The ordered necklace is perfect

Definition

The concatenation of all words of length n over a b -symbol alphabet in lexicographic order is called the **ordered necklace** for length n .

Proposition (Alvarez, Becher, Ferrari and Yuhjtman 2016)

The ordered necklace for length n is (n, n) -perfect.

Astute graphs

Fix b -symbol alphabet.

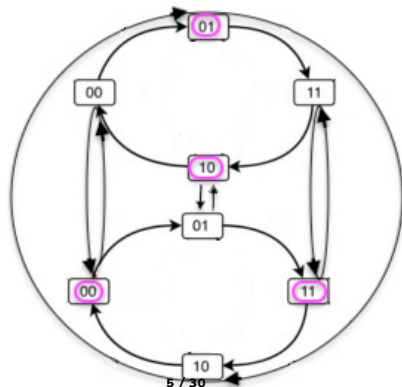
Consider the tensor product of the de Bruijn graph with a simple cycle.

The **astute graph** $G_b(n, k) = (V, E)$ is directed, with kb^n vertices.

$$V = \{0, \dots, b-1\}^n \times \mathbb{Z}/k\mathbb{Z}$$

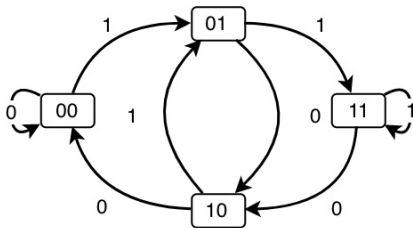
$$E = \{(u, m), (v, m+1) : u = a_1 \dots a_n, v = a_2 \dots a_n a_{n+1}\}$$

$$G_2(2, 2)$$



$G_b(n, 1)$ is the de Bruijn graph of words of length n over b -symbols.

$G_2(2, 1)$



Perfect necklaces characterization

Every Hamiltonian cycle in $G_b(n, k)$ yields an (n, k) -perfect necklace.

$G_b(n, k)$ is the line graph of $G_b(n - 1, k)$.

Thus, every Hamiltonian cycle in $G_b(n, k)$ is Eulerian in $G_b(n - 1, k)$,

Hence, every Eulerian cycle in $G_b(n - 1, k)$ yields one (n, k) -perfect necklace.

Each (n, k) -perfect necklace can come from several Eulerian cycles in $G_b(n - 1, k)$

Count

Theorem (Alvarez, Becher, Ferrari and Yuhjtman 2016)

The number of (n, k) -perfect necklaces over a b -symbol alphabet is

$$\frac{1}{k} \sum_{d_{b,k} | j | k} e(j) \varphi(k/j)$$

where

- ▶ if $k = p_1^{\alpha_1} \dots p_t^{\alpha_t}$, then $d_{b,k} = \prod p_i^{\alpha_i}$, where p_i divides both b and k ,
- ▶ $e(j) = (b!)^{j b^{n-1}} b^{-n}$ is the number of Eulerian cycles in $G_b(n-1, j)$
- ▶ φ is Euler's totient function.

Three families of perfect necklaces

Arithmetic, Nested, Succession

- Construction
- Count
- Discrete discrepancy

Arithmetic necklaces

Identify the words of length n over a b -symbol alphabet with the set of non-negative integers modulo b^n according to representation in base b .

Definition

Let $b \geq 2$ be an integer, let c be coprime with b . Let n be a positive integer. An arithmetic necklace is the concatenation of words of length n corresponding to the arithmetic progression with difference c :

$$\boxed{0} \quad \boxed{c \bmod b^n} \quad \boxed{2c \bmod b^n} \quad \dots \quad \boxed{(b^n - 1)c \bmod b^n}$$

With $c = 1$ we obtain the ordered necklace.

Theorem (Alvarez, Becher, Ferrari and Yuhjtman 2016)

For each n , the arithmetic necklaces are (n, n) -perfect.

Count

Given b and n , $\#$ numbers coprime to b and smaller than b^n .

Discrete discrepancy

Let A be b -symbol alphabet with equal probability.

The discrepancy $d : A^* \times \mathbb{N} \rightarrow \mathbb{N}$

$$d(w, \ell) = \max\{|v|_s - |v|_t| : s, t \in A^\ell, v \text{ substring of } w\}$$

where $|v|_s$ is the number of occurrences of s in v .

Example: $d(aaaa, 1) = 4$, $d(aaaa, 2) = 3$, $d(aaaa, 3) = 2$,
 $d(aaaa, 4) = 1$ and $d(abab, 1) = 1$, $d(abab, 2) = 2$, $d(abab, 3) = 1$.

Let $d(w) = \max_{1 \leq \ell < \lfloor \log |w| \rfloor} d(w, \ell)$.

Problem

What is minimal $d(w)$ among all words w (of a given size)?

Among all de Bruijn sequences?

Among all perfect necklaces?

Discrete discrepancy

Problem

What are the minimal and maximal discrete discrepancy for arithmetic necklaces?

The largest is presumably by the progression with difference 1.

For small discrete discrepancy:

Theorem (Levin 1999 Theorem 1, using Popov 1981)

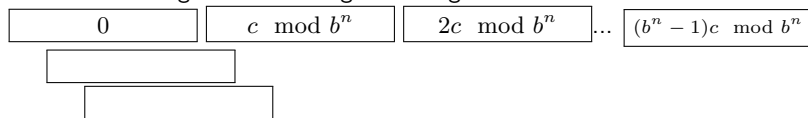
For every n there is an arithmetic necklace w such that $d(w) = O(n^3)$,

Conjecture (Капов 2019)

For every n there is an arithmetic necklace w such that $d(w) = O(n^2)$.

The discrete discrepancy of arithmetic necklaces

We need a sliding window of length n along this concatenation



These are nb^n windows.

Convert the nb^n windows to nb^n rationals in the unit interval (base- b expansion)

Consider the n progressions of b^n terms:

$$\begin{array}{ccccccc} 0, & \frac{c}{b^n} \bmod 1, & \frac{2c}{b^n} \bmod 1, & \dots, & \frac{(b^n - 1)c}{b^n} \bmod 1 \\ 0, & \frac{c}{b^{n-1}} \bmod 1, & \frac{2c}{b^{n-1}} \bmod 1, & \dots, & \frac{(b^n - 1)c}{b^{n-1}} \bmod 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0, & \frac{c}{b^2} \bmod 1, & \frac{2c}{b^2} \bmod 1, & \dots, & \frac{(b^n - 1)c}{b^2} \bmod 1 \\ 0, & \frac{c}{b} \bmod 1, & \frac{2c}{b} \bmod 1, & \dots, & \frac{(b^n - 1)c}{b} \bmod 1 \end{array}$$

Using classical discrepancy

By a classical result, for every $\alpha \in (0, 1)$ irrational,

$ND_N((n\alpha \bmod 1)_{n \geq 1}) \leq S(\alpha)$ stop at $t(N) + 1$, $q_{t(N)} \leq N \leq q_{t(N)+1}$

For $\alpha = [a_0; a_1, \dots, a_s]$ let $S(\alpha) = \sum_{i=1}^s a_i$.

$$D_N(x_1, x_2, \dots) = \sup_{(a,b) \subseteq (0,1)} \left| \frac{\#\{i : 1 \leq i \leq N, x_i \in (a,b)\}}{N} - (b-a) \right|$$

$$\frac{1}{2}d(w) \leq$$

$$D_N \left(0, \frac{c}{b^n} \bmod 1, \frac{c}{b^{n-1}} \bmod 1, \dots, \frac{c}{b} \bmod 1 \right) +$$

$$D_N \left(0, \frac{2c}{b^n} \bmod 1, \frac{2c}{b^{n-1}} \bmod 1, \dots, \frac{2c}{b} \bmod 1 \right) +$$

...

$$D_N \left(0, \frac{(b^n - 1)c}{b^n} \bmod 1, \frac{(b^n - 1)c}{b^{n-1}} \bmod 1, \dots, \frac{(b^n - 1)c}{b} \bmod 1 \right)$$

Levin 1999: For every $b \geq 2$ and n there is c coprime with b such that

$$\sum_{i=1}^n S(c/b^i) < Kn^3, \text{ where } K \text{ is constant.}$$

Definition (minimizer)

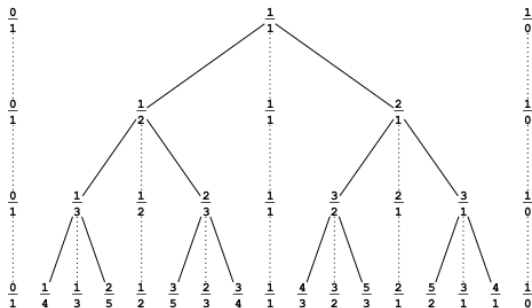
Let $b \geq 2$ be an integer and let n be a positive integer.

A minimizer for (b, n) is a positive integer c that minimizes $\sum_{i=1}^n S(c/b^i)$.

n	$b = 2$		$b = 3$		$b = 10$	
	c	$\sum_{i=1}^n S(c/b^i)$	r	$\sum_{i=1}^n S(c/b^i)$	r	$\sum_{i=1}^n S(c/b^i)$
1	1	2	1	3	3	6
2	1	6	2	9	27	17
3	3	11	5	18	173	36
4	3	19	31	29	2627	62
5	5	29	92	44	22627	91
6	19	39	140	63	262113	128
7	37	52	857	85	2262113	170
8	45	67	2570	109	16172177	227
9	151	83	9131	138	226542279	286
10	151	102	12262	172	—	—
11	807	125	31907	207	—	—
12	867	151	46787	245	—	—
13	3367	174	311411	286	—	—
14	3433	201	1288610	332	—	—
15	4825	231	3761986	379	—	—
16	13893	260	—	—	—	—
17	51351	289	—	—	—	—
18	79655	322	—	—	—	—
19	79655	357	—	—	—	—
20	444567	390	—	—	—	—

Stern-Brocot tree

The Stern-Brocot tree is a binary tree whose vertices are the positive rational numbers. The root is 1 (row $r = 0$). The left subtree, the Farey tree, contains the rationals less than 1.



The number x is at row r if and only if $S(x) = r + 1$.
 For b and n , find d coprime with b , between 1 and $b^n - 1$ minimizing

$$\sum_{i=1}^n \text{row}(c/b^i).$$

Around Zaremba's conjecture

For the case of $b = 2$, we need

For every n ,
find c odd between 1 and $2^n - 1$
that minimizes $S(c/2) + S(c/2^2) + \dots + S(c/2^n)$

Theorem (Neiderreter 1986, Zaremba's conjecture for the powers of 2)

For every n there is a c such that all the coefficients in the continued fraction expansion of $a/2^n$ are bounded by 3.

Zaremba's 1971 conjecture predicts that every integer appears as the denominator of a finite continued fraction whose coefficients are bounded by an absolute constant.

Nested perfect necklaces

Definition (Becher & Carton 2019)

An (n, k) -perfect necklace over a b -symbol alphabet is *nested* if $n = 1$ or it is the concatenation of b nested $(n - 1, k)$ -perfect necklaces.

This is a nested $(2, 2)$ -perfect necklace for $b = 2$

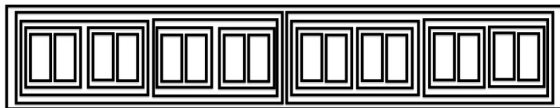
$$\underbrace{0011}_{(1,2)\text{-perfect}} \quad \underbrace{0110}_{(1,2)\text{-perfect}}$$

The ordered perfect necklace is not nested. For example, $b = 3$, $n = 2$,

$$\underbrace{00 \ 01 \ 02}_{\text{not } (1,2)\text{-perfect}} \quad \underbrace{10 \ 11 \ 12}_{\text{not } (1,2)\text{-perfect}} \quad \underbrace{20 \ 21 \ 22}_{\text{not } (1,2)\text{-perfect}}$$

Nested perfect necklaces

For example, in the binary alphabet and n is a power of 2,



- | | | |
|-----------|--|------------|
| 1 | nested (n, n) -perfect necklace | determines |
| 2 | nested $(n - 1, n)$ -perfect necklaces | determine |
| 2^2 | nested $(n - 2, n)$ -perfect necklaces | determine |
| \dots | | |
| 2^{n-1} | $(1, n)$ -perfect necklaces | |

Levin's necklace, n power of 2

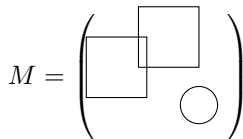
For n a power of 2, M. Levin (1999) defines a matrix M in $\mathbb{F}_2^{n \times n}$ using Pascal triangle matrix modulo 2,

$$M := (p_{i,j})_{i,j=0,1,\dots,n-1} \text{ where } p_{i,j} := \binom{i+j}{j} \pmod{2}.$$

M is upper triangular and it has the following property on submatrices.

Lemma (Levin 1999 from Bicknell and Hoggart 1978; Mereb 2023)

For Pascal triangle matrix modulo 2, each square submatrix at the left or at the top has determinant computed in \mathbb{Z} equal to 1 or -1 .



The diagram shows a large square representing a matrix M . Inside this square, there are two smaller squares that overlap each other. One square is positioned higher and further to the left, while the other is lower and further to the right. Below the large square, there is a small circle.

Then, if these determinants are computed in $\mathbb{Z}/b\mathbb{Z}$, for any integer $b \geq 2$, they are equal to 1 or -1 .

Levin's necklace, n power of 2

Definition (Levin 1999)

Let integer $b \geq 2$ and let n be a power of 2.

Identify the set of non-negative integers modulo b^n according to representation in base b with the vectors w_0, \dots, w_{b^n-1} in $(\mathbb{Z}/b\mathbb{Z})^n$.

Let $M \in \mathbb{F}_2^{n \times n}$ be the Pascal triangle matrix modulo 2.

Define the necklace (computation is done in $\mathbb{Z}/b\mathbb{Z}$)

$$Mw_0 \dots Mw_{b^n-1}.$$

For example, for $b = 2$,

$$\begin{array}{l} n = 2^0 \quad 01 \\ n = 2^1 \quad 0011 \ 1001 \\ n = 2^2 \quad 0000 \ 1111 \ 1010 \ 0101 \ 1100 \ 0011 \ 0110 \ 1001 \ 1000 \ 0111 \ 0010 \ 1101 \ 0100 \ 1011 \ 1110 \\ \dots \end{array}$$

Levin's necklace is nested perfect, n power of 2

Theorem (Becher and Carton 2019)

Let $b \geq 2$ be a integer and let n be a power of 2. The necklace given by the Pascal triangle matrix modulo 2 is nested (n, n) -perfect.

Construction of nested (n, n) -perfect necklaces, n power of 2

Definition (Pascal-like family $\mathcal{P} \subseteq \mathbb{F}_2^{n \times n}$)

Let n be a power of 2.

Let $(\eta_j)_{0 \leq j < n}$ such that $\eta_0 = 0, \eta_j \leq \eta_{j+1} \leq \eta_j + 1$ (non decreasing step)

Define $M^\eta = (p_{i,j}^\eta)_{0 \leq i, j < n}$ in $\mathbb{F}_2^{n \times n}$,

$$p_{i,j}^\eta = \binom{i + j - \eta_j}{j} \pmod{2}$$

For each M in $\mathcal{P} = \{M^\eta : \eta \text{ non decreasing step}\}$, for every integer $b \geq 2$,

$$Mw_0 \dots Mw_{b^n - 1}$$

(multiplication in $\mathbb{Z}/b\mathbb{Z}$) is a nested (n, n) -perfect necklace.

Count of binary nested (n, n) -perfect necklaces, n power of 2

Theorem (Becher and Carton 2019)

There are 2^{2n-1} *binary* nested (n, n) -perfect necklaces, n power of 2.

Proof.

For each M in \mathcal{P} and for each z in \mathbb{F}_2^n , $M(w_0 \oplus z) \dots M(w_{2^n-1} \oplus z)$ is a binary nested (n, n) -perfect necklace.

If $z' = Mz$, $M(w_0 \oplus z) \dots M(w_{2^n-1} \oplus z) = Mw_0 \oplus z' \dots Mw_{2^n-1} \oplus z'$,

matrices $\in \mathbb{F}_2^{n \times n}$ in $\mathcal{P} \times$ # vectors $z \in \mathbb{F}_2^n = 2^{2n-1} \leq$ Total.

By a graph theoretical argument we know that there can be no more. \square

Construction nested (n, n) -perfect necklaces, prime b , n power of b

Definition (Pascal-like family $\mathcal{P} \subseteq \mathbb{F}_b^{n \times n}$ Hofer and Larcher, 2022)

Let b be a prime.

Let n be a power of b .

Let $(u_j)_{0 \leq j < n}$ with $u_j \not\equiv 0 \pmod{b}$.

Let $(\eta_j)_{0 \leq j < n}$ such that $\eta_0 = 0, \eta_j \leq \eta_{j+1} \leq \eta_j + 1$.

Define $M^{u, \eta} = (p_{i,j}^{u, \eta})_{0 \leq i, j < n}$ in $\mathbb{F}_b^{n \times n}$,

$$p_{i,j}^{u, \eta} = \binom{i + j - \eta_j}{j} u_j \pmod{b}.$$

For each M in \mathcal{P} ,

$$Mw_0 \dots M(w_{b^n-1})$$

(multiplication in \mathbb{F}_b) is a nested (n, n) -perfect necklace.

Count, we know very little

Base/n	necklace in base $b = 2$	necklace in base $b \geq 3$
n a power of 2	2^{2n-1}	?
n a power of prime $b \geq 3$?	?

Discrepancy

Theorem (Levin 1999; Becher and Carton 2019, Hofer and Larcher 2022,2023)

	<i>discrete discrepancy necklace in base $b \geq 2$</i>
n a power of 2	$d(w) = O(n^2)$
n a power of prime b	

In case of the canonical Pascal triangle matrix in \mathbb{F}_b , prime $b \geq 2$,
 $d(w) = \Theta(n^2)$.

Nested perfect arrays

The positions of an array are partitioned into $p \times q$ residue classes according to their respective residue classes modulo p and modulo q : An array of size $n \times n$ occurs at position (k, ℓ) in a given array if it is equal to the subarray of size $n \times n$ at position (k, ℓ) .

Definition (Perfect array)

An array A is (s, t, p, q) -perfect, if each $s \times t$ array has the same number of occurrences in A in each residue class modulo (p, q) .

Definition (Square nested perfect array)

Assume a b -symbol alphabet. An (n, n, n, n) -perfect array is nested if, for $k = 1, \dots, n$, each subarray of size $nb^{nk/2} \times nb^{nk/2}$ is (k, k, n, n) -perfect. Becher and Carton (2023) constructed binary square nested arrays using Pascal triangle matrix.

Problem

- Construct other nested perfect arrays.
- Determine a walk in the array that has low discrepancy.

Succession necklaces

Definition

The succession necklaces for (n, k) are (n, k) -perfect necklaces that correspond to Eulerian cycles in $G_b(n-1, k)$ obtained by joining cycles given by a succession rule.

We extend the shift registers of Golomb 1967 to construct some: (n, k) -rotation cycles and (n, k) -increment cycles.

We do not know how to count them.

Observation

For every n , the ordered necklace of words of length n is arithmetic and succession.

Discrepancy of succession necklaces

Theorem (Álvarez, Becher, Mereb, Pajor and Soto 2023)

We construct an $(n, 1)$ -perfect necklace by joining $(n, 1)$ -increment cycles

*for $b = 2$, symbol discrepancy is n , and this is optimal
for $b \geq 3$, symbol discrepancy is $n + 1$.*

Nikolai Korobov (1955)

What is the minimum $D_N((b^n x \bmod 1)_{n \geq 0})$ for some real number x and integer b ?

Some references

- C. Aistleitner, B. Borda, M. Hauke. On the distribution of partial quotients of reduced fractions with fixed denominator. *Trans. Amer. Math. Soc.* 377: 1371-1408, 2024.
- N. Alvarez, V. Becher, M. Mereb, I. Pajor and C. Soto. de Bruijn sequences with minimal discrepancy, submitted, 2024, <https://arxiv.org/abs/2407.17367>
- N. Alvarez, V. Becher, P. Ferrari and S. Yuhjtman. Perfect necklaces, *Advances of Applied Mathematics* 80:48–61, 2016.
- V. Becher and O. Carton. Normal numbers and nested perfect necklaces, *Journal of Complexity* 54:101403, 2019.
- V. Becher and O. Carton. Nested perfect arrays, *IEEE Transactions on Information Theory* 70(10): 7463 - 7471, 2024.
- P. Flajolet, P. Kirschenhofer, R. Tichy R. Discrepancy of Sequences in Discrete Spaces. In: Halász G., Sós V.T. (eds) *Irregularities of Partitions*. Algorithms and Combinatorics 8. Springer, Berlin, Heidelberg, 61-70, 1989.
- P. Flajolet, P. Kirschenhofer, R. Tichy. Deviations from Uniformity in Random Strings. *Probability Theory and Related Fields* 80, 139-150, 1988.
- P. Gauna. Números normales muy rápidos. *Tesis de Licenciatura en Ciencias de la Computación*. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires, 2019.
- S. Golomb. *Shift register sequences*. Aegean Park 1967.
- G. Larcher and R. Hofer. The exact order of discrepancy for Levin's normal number in base 2, *Journal de Théorie des Nombres de Bordeaux* 35(3): 999-1023, 2023. , 2022.
- G. Larcher and R. Hofer. Discrepancy bounds for normal numbers generated by necklaces in arbitrary bases, *Journal of Complexity* 78:1017672023, 2023
- M. B. Levin. On the discrepancy estimate of normal numbers. *Acta Arithmetica* 88(2):99–111, 1999.
- M. Mereb. Determinants of matrices related to the Pascal triangle. *Periodica Mathematica Hungarica* 89: 168-174, 2024.