Perfect necklaces

Verónica Becher

Universidad de Buenos Aires

Séminaire de l'équipe de Théorie Ergodique et Systèmes Dynamiques
L'Institut Galilée, Université Sorbonne Paris Nord
October 11, 2023

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.

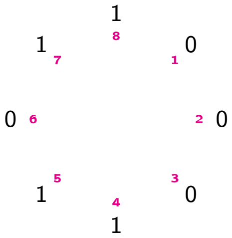
The (n, k)-perfect necklaces have length kb^n.

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

Example

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

$$00\ 01\ 10\ 11$$



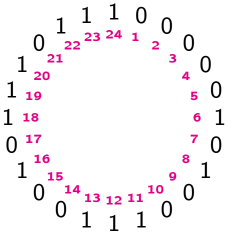
00 occurs twice (p:1,2);
01 occurs twice (p:3,6);
10 occurs twice (p:5,8);
11 occurs twice (p:4,7)

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.

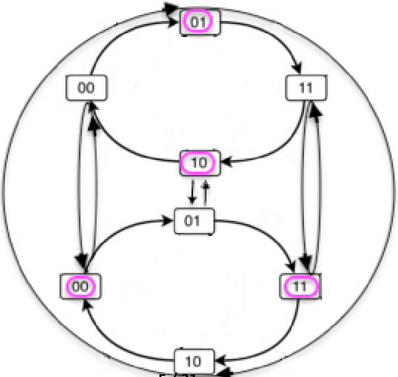
Proposition (Alvarez, Becher, Ferrari and Yuhjtman 2016)

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

Astute graphs

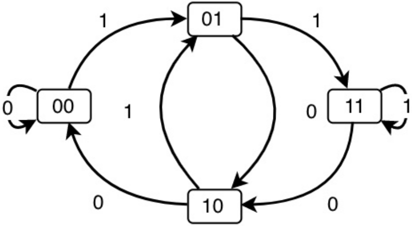
Fix b-symbol alphabet.
Consider the tensor product of the de Bruijn graph with a simple cycle.
The astute graph $G_b(n,k) = (V,E)$ is directed, with kb^n vertices.
$V = \{0,..,b-1\}^n \times \mathbb{Z}/k\mathbb{Z}$
$E = \{(u,m),(v,m+1) : u = a_1...a_n, v = a_2...a_n a_{n+1}\}$

$G_2(2,2)$



$G_b(n,1)$ 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 $k = p_1^{\alpha_1} ... p_t^{\alpha_t}$, then $d_{b,k} = \prod p_i^{\alpha_i}$, where p_i divides both b and k,*
- *$e(j) = (b!)^{jb^{n-1}} b^{-n}$ 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 d 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 d:

0	$d \mod b^n$	$2d \mod b^n$	...	$(b^n - 1)d \mod b^n$

With $d = 1$ 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

Fix b-symbol alphabet with the uniform measure.

$$\Delta_{\ell,N}(a_1 a_2 \ldots) = \max_{u \in \{0,..,b-1\}^\ell} \left| \frac{|a_1 a_2 \ldots a_{N+\ell-1}|_u}{N} - \frac{1}{b^\ell} \right|$$

where $\ell < \lfloor \log(N) \rfloor$. This is to obtain u.d. of almost all (with respect to the product measure) infinite sequences (Flajolet, Kirschenhofer and Tichy 1988)

Problem

What is minimal $\Delta_{\ell,N}(x)$ among all words x?

This is the discrete counterpart of Korobov's question (1955) on the minimum $D_N((b^n x \mod 1)_{n \geq 0})$ for some real number x and integer b.

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 discrepancy:

Theorem (Levin 1999 Theorem 1, using Popov 1981)

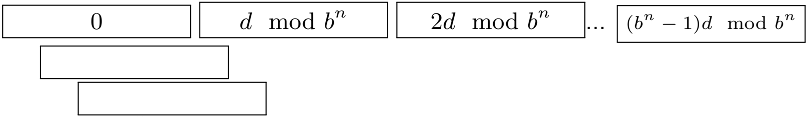
For every n there is an arithmetic necklace such that $N\Delta_N = O(n^3)$.

Conjecture (Becher and Carton 2019)

For every n there is an arithmetic necklace such that $N\Delta_N = O(n^2 \log n)$.

Using classical discrepancy

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)
We obtain n progressions of b^n terms:

$$\begin{array}{llll}
0, \quad \frac{d}{b^n} \mod 1, & \frac{2d}{b^n} \mod 1, & \ldots, & \frac{(b^n-1)d}{b^n} \mod 1 \\
0, \quad \frac{d}{b^{n-1}} \mod 1, & \frac{2d}{b^{n-1}} \mod 1, & \ldots, & \frac{(b^n-1)d}{b^{n-1}} \mod 1 \\
\quad \vdots & \vdots & \vdots & \vdots \\
0, \quad \frac{d}{b^2} \mod 1, & \frac{2d}{b^2} \mod 1, & \ldots, & \frac{(b^n-1)d}{b^2} \mod 1 \\
0, \quad \frac{d}{b} \mod 1, & \frac{2d}{b} \mod 1, & \ldots, & \frac{(b^n-1)d}{b} \mod 1
\end{array}$$

Classical discrepancy on arithmetic progressions

For $\alpha = [a_0; a_1, \ldots, a_s]$ let $S(\alpha) = \sum_{i=1}^{s} a_i$.
By a classical result,
$ND_N((k\alpha \mod 1)_{k\geq 1}) \leq S(\alpha)$ stop at $t(N)+1$, $q_{t(N)} \leq N \leq q_{t(N)+1}$

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

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

Since $\Delta_N \leq D_N$, for N between 1 and nb^n,

$$N\Delta_N(x) \leq \sum_{i=1}^{n} S(d/b^i) = O(n^3).$$

Our conjecture

Definition (minimizer)

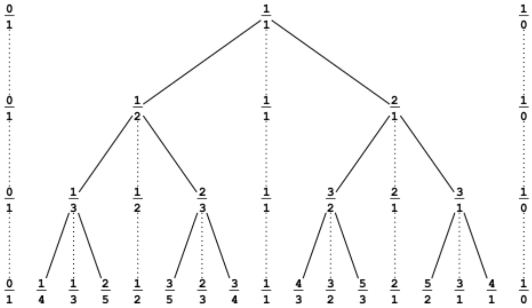
Let $b \geq 2$ be an integer and let n be a positive integer.

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

	$b = 2$		$b = 3$		$b = 10$	
n	d	$\sum_{i=1}^{n} S(d/b^i)$	r	$\sum_{i=1}^{n} S(d/b^i)$	r	$\sum_{i=1}^{d} S(d/b^i)$
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	[illegible]				

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

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

Around Zaremba's conjecture

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

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

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

For very n there is a 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 $b=2$

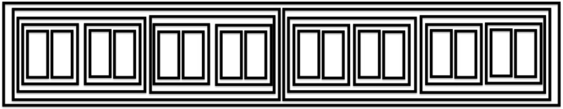
$$\underbrace{0011}_{\text{(1,2)-perfect}}\ \underbrace{0110}_{\text{(1,2)-perfect}}$$

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

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

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
...		
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,

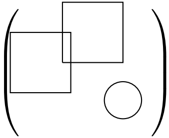
$$M := (p_{i,j})_{i,j=0,1,\ldots n-1} \text{ where } p_{i,j} := \binom{i+j}{j} \mod 2.$$

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.

$M =$



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 $w_0, \ldots w_{b^n-1}$ 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 $\mathbb{Z}/b\mathbb{Z}$)

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

For example, for $b = 2$,

$n = 2^0$ 0 1
$n = 2^1$ 00 11 10 01
$n = 2^2$ 0000 1111 1010 0101 1100 0011 0110 1001 1000 0111 0010 1101 0100 1011 111
...

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 $(\eta_j)_{0 \le j < n}$ such that $\eta_0 = 0, \eta_j \le \eta_{j+1} \le \eta_j + 1$ (non decreasing step)
Define $M^\eta = (p^\eta_{i,j})_{0 \le i,j < n}$ in $\mathbb{F}_2^{n \times n}$,

$$p^\eta_{i,j} = \binom{i + j - \eta_j}{j} \mod 2$$

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

$$M w_0 \ldots M w_{b^n - 1}$$

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

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 $(u_j)_{0\le j<n}$ with $u_j \not\equiv 0 \mod b$.
Let $(\eta_j)_{0\le j<n}$ such that $\eta_0 = 0, \eta_j \le \eta_{j+1} \le \eta_j + 1$.
Define $M^{u,\eta} = (p^{u,\eta}_{i,j})_{0\le i,j<n}$ in $\mathbb{F}_b^{n\times n}$,

$$p^{u,\eta}_{i,j} = \binom{i + j - \eta_j}{j} u_j \mod b.$$

For each M in $\mathcal{P}$,

$$M w_0 \dots M(w_{b^n-1})$$

(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)

Base/n	*necklace in base* $b \geq 2$
n *a power of* 2 n *a power of prime* b	$\Delta_N = O(n(\log N)/N)$

In case of the canonical Pascal triangle matrix in $\mathbb{F}_b$, prime $b \geq 2$, $\Delta_N = \Theta(n(\log N)/N)$.

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 $b = 2$, $N\Delta_{1,N} = n/2$, and this is optimal;
for $b \geq 3$, $N\Delta_{1,N} = (n+1)/2$.

Summary on necklaces and discrepancy

Arithmetic (n, n)-perfect necklaces
Exist arithmetic (n, n)-perfect $N\Delta_N = O(n^3)$, Levin 1999
Conjecture: Exist arithmetic (n, n)-perfect $N\Delta_N = O(n^2 \log n)$

Nested (n, n)-perfect necklaces,
n a power of 2, $b \geq 2$, $N\Delta_N = O(n \log N)$, Levin 1999; Becher and Carton 2019
n a power of b, b prime, $N\Delta_N = O(n \log N)$
In Levin's cases, this is exact Hofer and Larcher 2022,2023

Succession (n, k)-perfect necklaces Alvarez,Becher,Mereb,Pajor,Soto 2023
There exist succession (n, n)-perfect with $N\Delta_{1,N} = \Theta(n)$.

There are many open questions.

Nikolaĭ Korobov (1955)

What is the minimum discrepancy associated to a normal number?

Some references

- C. Aistleitner, B. Borda, M. Hauke. On the distribution of partial quotients of reduced fractions with fixed denominator. Arxiv 14095, 2022.
- N. Alvarez, V. Becher, M. Mereb, I. Pajor and C. Soto. An algorithm to onstruct a de Bruijn sequence with minimal discrepancy, *manuscript*, 2022.
- 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.
- 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. *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, 2022.
- G. Larcher and R. Hofer. Discrepancy bounds for normal numbers generated by necklaces in arbitrary bases, Journal of Complexity, 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*, 2023.

Classical definitions

Let b be a base. Let $\Sigma_b = \{0, \dots b-1\}$.

Rotation $r : \Sigma_b^n \to \Sigma_b^n$ moves the last character of the string to the front.
For example: $r(0010) = 0001; r(0001) = 1000; r^2(0010) = 1000$
The equivalence classes of Σ_b^n under rotation are the necklaces of size n.

Since r is invertible with $r^{-1} = r^{n-1}$, $\langle r \rangle$ is a group that acts on Σ_b^n.

Incremented rotation $\iota : \Sigma_b^n \to \Sigma_b^n$ increments the last character of the string (modulo b) and moves that incremented character to the front.
For example, if $b = 3$, we have: $i(0021) = 2002$, $i(2002) = 0200$

Since ι is invertible, $\langle \iota \rangle$ is a group that acts on Σ_b^n.

Succession perfect necklaces

Definition

For positive n and k we define $r_k : \Sigma_b^n \times \mathbb{Z}/k\mathbb{Z} \to \Sigma_b^n \times \mathbb{Z}/k\mathbb{Z}$,

$$r_k(s,t) = (r(s), t+1)$$

The (n,k)-rotation necklaces are the orbits of $\langle r_k \rangle$ on the set $\Sigma_b^n \times \mathbb{Z}/k\mathbb{Z}$

Applying Burnside's Lemma,

Proposition (ABMPS 2022)

The (n,k)-rotation necklaces correspond to some simple cycles in graph $G_b(n-1,k)$ and determine a partition of $G_b(n-1,k)$. The total number is

$$\frac{\gcd(n,k)}{n} \sum_{\gcd(n,k)|d|n} \varphi(n/d) b^d$$

Succession perfect necklaces

Definition

For a positive integer k we define $\iota_k : \Sigma_b^n \times \mathbb{Z}/k\mathbb{Z} \to \Sigma_b^n \times \mathbb{Z}/k\mathbb{Z}$,

$$\iota_k(s,t) = (\iota(s), t+1)$$

The (n,k)-increment necklaces are the orbits of $\langle \iota_k \rangle$ on the set $\Sigma_b^n \times \mathbb{Z}/k\mathbb{Z}$
Applying Burnside's Lemma,

Proposition (ABMPS 2022)

The (n,k)-increment necklaces correspond to some simple cycles in graph $G_b(n-1,k)$ and determine a partition of $G_b(n-1,k)$.
The total number is

$$\frac{\gcd(\gcd(k,bs)n/s, kb)}{nb} \sum_{\gcd(n,\operatorname{lcm}(k,bs))|d|n} \varphi(n/d) \cdot b^d$$

where s is the smallest divisor of n such that n/s is coprime with b.

Nested marvelous (semi-perfect) necklaces

Definition

A necklace over a a b-symbol alphabet is nested (n, k)-marvelous if all words of length n occur exactly k times, and in case $n > 1$ it is the concatenation of b nested $(n-1, k)$-marvelous necklaces.

This is nested $(3, 3)$-marvelous, not perfect,

$$000111\ 011001\ 000111\ 101010$$

Theorem (Frizzo 2020; Larcher and Hofer 2022)

For every number x whose base-b expansion is the concatenation of nested (n, n)-marvelous necklaces, for n a power of b or a power of 2, $D_N(\{b^t x\}_{t\geq 0})$ is $O\left((\log N)^2/N\right)$.