

Permutons and entropy

TDK thesis

Authors: Péter László Fazekas, Balázs Szepesi, Sára Szepessy

Project supervisor: Balázs Maga

ELTE-TTK Department of Analysis

2024

Abstract

For any permuton μ we can define the pattern density distribution over $\text{Sym}(n)$ ($n > 0$) through sampling. In our project, we studied the asymptotics of the Shannon entropy sequence of these distributions which we denote $H(\mu^n)$. We calculated the asymptotic value of $H(\mu^n)$ for some special cases. We gave formulas for certain permuton constructions. For pattern avoiding permutons it can be shown that $H(\mu^n)/n \leq c$ for some c constant. We designed a permuton which contains any pattern with positive density, yet satisfies the inequality above, displaying the lack of equivalence of these properties. One can verify that the identity

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = \int_0^1 \log |f'(x)| dx$$

holds for any permuton μ corresponding to a measure preserving function f which consists of finitely many continuously differentiable pieces. We show that this identity does not generalize for infinitely many linear pieces, but show some special cases where it does.

1. Introduction

A modern approach to the study of large permutations is analyzing their limit objects called permutons. Our involvement started with our Research Experience for Undergraduates (REU) project, during which our research group investigated questions related to the Shannon entropy sequence one can naturally associate with a permuton. After discussing the relevant notions, we summarize our results, which are related to a yet unpublished paper of our supervisor, however, do not depend on the content of it. At this point, it should be noted that our REU research group included two foreign students from the programme Budapest Semesters in Mathematics, Bowen Li and Aayan Pathan, to whom we are grateful for their helpful presence in the starting stages of this project.

The main result of this thesis concerns limit properties of the Shannon-entropy of permutons. A permutation π is a bijective function of $\{1, 2, \dots, n\}$ to itself. The length of π is n and we use the notation $\pi(i)$ as the image of i . The set of all permutations of length n is denoted by $\text{Sym}(n)$. By $[n]_{<}^m$ we mean the set of m -tuples in $[n]$ whose elements are in strictly increasing order. Permutons were introduced in [1], from where we essentially copy the necessary definitions.

Definition 1.1. (pattern density) For positive integers $k, n \in \mathbb{N}$, let $\tau \in \text{Sym}(k)$ and $\pi \in \text{Sym}(n)$. The number of occurrences Λ of the permutation τ in π is the number of k -tuples $(x_1, x_2, \dots, x_k) \in [n]_{<}^k$ such that $\pi(x_i) < \pi(x_j)$ if and only if $\tau(i) < \tau(j)$. The density of the permutation τ as a pattern of π is given by

$$t(\tau, \pi) = \begin{cases} \binom{n}{k}^{-1} \Lambda(\tau, \pi) & \text{if } k \leq n \\ 0 & \text{if } k > n \end{cases}$$

Example 1.2. $\pi = 13254$, $\tau = 12$. If we take 3, 5 in π , it gives the pattern 12, if we take 5, 4, it gives 21. Taking uniform random choice of k distinct entries, the probability of seeing 12 is the density of τ in π .

Example 1.3. If $\pi = 132$, $\tau = 12$, then $t(\tau, \pi) = 2/3$, since $\underline{132} = 12$, $\underline{13}\underline{2} = 12$ and $\underline{13}\underline{2} = 21$.

Definition 1.4. (convergence of a permutations sequence) A permutation sequence $(\pi_n)_{n \in \mathbb{N}}$ is convergent if, for every fixed permutation τ , the sequence of real numbers $(t(\tau, \pi_n))_{n \in \mathbb{N}}$ converges.

If $\exists K \in \mathbb{N}$ such that $\forall \pi_i \in (\pi_n)_{n \in \mathbb{N}} |\pi_i| < K$, then $\exists m \in \mathbb{N}$ such that $\forall n_1, n_2 > m \pi_{n_1} = \pi_{n_2}$ so the limit object is a permutation which is not the interesting case. When the length of the permutations tend to infinity we use the following limit object:

Definition 1.5. (permuton) A permuton μ is a Borel probability measure on $[0, 1]^2$ with uniform marginals, that is, $\mu(Z \times [0, 1]) = \mu([0, 1] \times Z)$ is equal to the Lebesgue measure of Z for any Borel set $Z \subset [0, 1]$.

Suppose we sample points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ independently at random from the measure μ . Let $(x_1^*, x_2^*, \dots, x_n^*)$ and $(y_1^*, y_2^*, \dots, y_n^*)$ denote the values of $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$ respectively, rearranged in increasing order. The result of this sampling is a permutation π , where $\pi(i) = j$ if and only if $x_i^* = x_j$ and $y_j^* = y_i$ for some $i \in [n]$. We denote the distribution of this random permutation by μ^n .

These measures encapsulate all information about the permuton, making their study worthwhile. One possible direction is to examine the Shannon entropy of these measures, which serves as a metric for their 'spread,' and is defined as follows:

Definition 1.6. (entropy) The Shannon entropy or just entropy of a discrete random variable X is

$$H[X] = - \sum_x \mathbb{P}(X = x) \log \mathbb{P}(X = x) = -\mathbb{E}[\log \mathbb{P}(X)]$$

when the sum exists. The joint entropy of two random variables, $H[X, Y]$, is the entropy of their joint distribution.

Applying entropy on the random variable series $(\mu^n)_{n \in \mathbb{N}}$ received from the sampling method as defined above we gain the sequence $(H(\mu^n))_{n \in \mathbb{N}}$ (for simplicity, we will denote it by $H(\mu^n)$ from now on) that we call pattern entropy sequence. In this thesis, we investigate questions concerning the asymptotics of $H(\mu^n)$ of pattern entropy sequences of permutons.

We can interpret the weight assigned to τ by μ^n as the density of τ within μ . Using this interpretation, we define the concept of a sequence of permutations converging to a permuton.

Definition 1.7. Let $(\pi_n)_{n \in \mathbb{N}}$ be a sequence of permutations such that $|\pi_n| \rightarrow \infty$. Let μ be a permuton. We say that $(\pi_n)_{n \in \mathbb{N}}$ converges to μ or write $\pi_n \rightarrow \mu$, if:

$$\forall \tau \in \text{Sym}(n) : \lim_{n \rightarrow \infty} t(\tau, \pi_n) = t(\tau, \mu)$$

Example 1.8. Let $(\pi_n)_{n \in \mathbb{N}}$ be the permutation series where $\forall n \in \mathbb{N} |\pi_n| = n$ and $\forall k \in [n] \pi_n(k) = k$. In this case the limit of this series is the identity permuton, where we take uniform distribution on the diagonal line $\{(x, x) : x \in [0, 1]\}$. A non-trivial permutation series converging to the same permuton is the following: $(\pi_n)_{n \in \mathbb{N}}$, $\forall n \in \mathbb{N} |\pi_n| = n$ and $\forall k \in [n-1] \pi_n(k) = k+1$, $\pi_n(n) = 1$.

Definition 1.9. (pattern-avoidance) We say that a permutation π is τ -avoiding if $t(\tau, \pi) = 0$. Similarly, a permuton μ is τ -avoiding if $t(\tau, \mu) = 0$. A permutation or permuton is pattern avoiding if there exists a pattern τ for which it is τ -avoiding.

Pattern-avoidance is a key aspect in the study of permutations, with numerous results addressing the structures that emerge [2]. An example for these results is the Stanley-Wilf conjecture proven by Marcus and Tardos [3]:

Theorem 1.10. (*Marcus-Tardos theorem*) *For any τ permutation there exists C_τ such that the number of τ -avoiding permutations of length n is at most C_τ^n*

The aspects of pattern-avoidance is examined among permutons as well. The following claim is a direct consequence of the Marcus-Tardos theorem regarding permutons:

Proposition 1.11. $H(\mu^n) = O(n)$ for a pattern-avoiding permuton μ .

In the beginning of Section 3 we give some basic examples about the asymptotics of $H(\mu^n)$ on particular permutons. Later we introduce the weighted product of two permutons [3.1] and give a formula concerning its entropy and present some applications. We also give some asymptotic boundaries for certain cases. In the second part of this section [3.2] experiment with a definition for quasi-pattern-avoidance which has been inspired by the relationship between pattern-avoiding permutons and the asymptotics of $H(\mu^n)$ of their entropy. Using this new definition we make some claims regarding some special cases.

It is shown in 1.11 that pattern-avoiding permutons have an asymptotic upper bound on their entropy. In Section 4.1 we show a permuton for which the upper bound holds, yet it contains any permutation with positive density, showcasing that there is no equivalence between the two.

Calculating a permuton's asymptotics of $H(\mu^n)$ is generally not an easy task, but with certain restrictions we arrive at an unexpectedly elegant formula, identical to Rokhlin's formula about the Kolmogorov–Sinai entropy of the underlying measure-preserving function [5]. One can verify that the identity

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = \int_0^1 \log |f'(x)| dx$$

holds for any permuton μ corresponding to a measure preserving function f which consists of finitely many continuously differentiable pieces. We investigated further this expression and in Section 4.2 we give a family of permutons which contain infinitely many continuously differentiable pieces and the formula does not work, concluding that the formula cannot be generalized further. In Section 5, we present an open problem suggesting that $H(\mu^n)$ is non-decreasing for all permutons.

2. Preliminaries

Throughout the thesis we use the term o to denote asymptotic growth referred to "small-O order", O to denote "big-O order" and Θ to denote asymptotic growth which is in "big-O order" and not in "small-O order".

The Lebesgue measure on the unit square is a permuton which has some interesting properties. Any permutation series π_i , where for each i we pick independently and uniformly randomly from $\text{Sym}(i)$, will converge to the Lebesgue measure. For any permutation τ with length k , the density of the permutation is $t(\tau, \mu) = \frac{1}{n!}$ since the sampling gives any permutation with the

same probability. By using these properties of the Lebesgue measure it can be shown that its entropy is asymptotically $n \log(n)$, which is the greatest achievable among permutons. We use this property later in our thesis.

To establish a clear foundation, we introduce a key definition and a well-known proposition related to entropy, both of which are essential for the subsequent analysis. For clarity and completeness, we take these from [4].

Definition 2.1. The conditional entropy of X , given Y , $H[X|Y]$ is:

$$H[X|Y] = H[X, Y] - H[Y]$$

Here are some well-known properties of the Shannon entropy:

Proposition 2.2. Let X, Y be discrete random variables. The following properties hold:

- $H[X] + H[Y] \geq H[X, Y]$, with equality iff X and Y are independent.
- $H[X, Y] \geq H[X]$
- $H[X|Y] \leq H[X]$, with equality iff X is independent of Y

A well-known property for Bernoulli distribution, which we use later:

Lemma 2.3. The entropy of Bernoulli process on n trials is asymptotically equivalent to

$$\frac{1}{2} \log_2[2\pi enp(1-p)].$$

Measure-preserving functions are of large importance as they naturally give rise to permutons.

Definition 2.4. We say that a function $f: [0, 1] \rightarrow [0, 1]$ is *measure preserving*, if $\forall A \subset [0, 1]$

$$\lambda_1(f^{-1}(A)) = \lambda_1(A)$$

where λ_1 denotes the one dimensional Lebesgue measure.

For a measure preserving function, one can define a corresponding permuton the following way:

Definition 2.5. If f is a measure-preserving function and $A, B \subset [0, 1]$ are measurable sets, then we define

$$\mu_f(A \times B) = \lambda_1(f^{-1}(B) \cap A).$$

This measure can, of course, be extended to all Lebesgue-measurable sets. We refer to this extended measure as the permuton corresponding to f .

It is straightforward to verify that the measure defined above indeed is a permuton.

In their paper, Hoppen et al. [1] provide a characterization of convergence for permuton sequences, relating it to weak convergence of probability measures and convergence via the so-called rectangular distance. This latter concept proves to be a useful tool when handling such limits.

Definition 2.6 (rectangular distance). The rectangular distance of two permutons μ and ν is $d_{\square}(\mu, \nu) := \sup_{R=[a,b] \times [c,d]} |\mu(R) - \nu(R)|$.

We recall some equivalences from [1]:

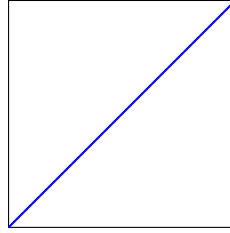
Theorem 2.7. *The followings are equivalent :*

1. $\mu_n \rightarrow \mu$, i.e., $\forall \tau, t(\tau, \mu_n) \rightarrow t(\tau, \mu)$
2. $\mu_n \xrightarrow{w} \mu$ (weak convergence)
3. $\forall x_0, y_0 : \mu_n([0, x_0] \times [0, y_0]) \rightarrow \mu([0, x_0] \times [0, y_0])$
4. $d_{\square}(\mu_n, \mu) \rightarrow 0$.

In the following, we will use this to prove Theorem 4.1.

3. Calculating asymptotic Shannon entropy of a permuton

The most trivial example one can think of is when μ is the uniform distribution supported along the diagonal line segment as following:

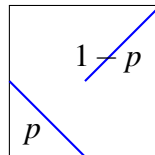


In this case the entropy is

$$H[\mu^{(n)}] = 0$$

for all n .

Proposition 3.1. *Define the permutons μ_f and μ_g corresponding to the measure-preserving functions $f(x) = 1 - x$ for $x \in [0, 1]$ and $g(x) = x$ for $x \in [0, 1]$. We scale μ_f to fit within the square $[0, p] \times [0, p]$ with weight p and μ_g to fit within the square $[p, 1] \times [p, 1]$ with weight $1 - p$. The permuton μ is then defined as the sum of these two measures.*

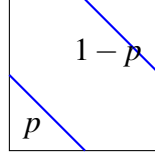


Then we have

$$\lim_{n \rightarrow \infty} \frac{H[\mu^{(n)}]}{\log n} = \frac{1}{2}.$$

Proof. Given sampling n points, we will have $\binom{n}{k} p^k (1-p)^{n-k}$ probability reaching the permutation $k(k-1)\dots 1(k+1)\dots n$, which is same as getting k heads in Bernoulli process. Then, by Lemma 2.3, we have the desired entropy. $\square$

Proposition 3.2. Let $f(x) = g(x) = 1 - x$ for $x \in [0, 1]$ and define permuton μ as above.



Then we have

$$\lim_{n \rightarrow \infty} \frac{H[\mu^{(n)}]}{\log n} = \frac{1}{2}.$$

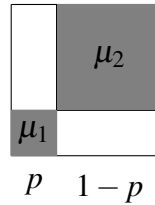
Proof. Essentially same proof as the previous proposition. $\square$

3.1. Combining two permutons

Let μ_1 and μ_2 be two permutons. We define their weighted product $\mu_1 \otimes_p \mu_2$ as follows:

$$\mu_1 \otimes_p \mu_2 = p \cdot (\mu_1)_{[0,p] \times [0,p]} + (1-p) \cdot (\mu_2)_{[p,1] \times [p,1]},$$

where $p \cdot (\mu_1)_{[0,p] \times [0,p]}$ denotes the measure μ_1 scaled down to fit within the square $[0, p] \times [0, p]$ with weight p , and $(1-p) \cdot (\mu_2)_{[p,1] \times [p,1]}$ denotes the measure μ_2 scaled down to fit within the square $[p, 1] \times [p, 1]$ with weight $1-p$.



Let $Y \in \{0, 1, \dots, n\}$ be the random variable, which is the number of points arrived in the bottom square given n samples of $\mu_1 \otimes_p \mu_2$.

Theorem 3.3 (weighted product formula).

$$H[(\mu_1 \otimes_p \mu_2)^{(n)} | Y] = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \left(H[\mu_1^{(k)}] + H[\mu_2^{(n-k)}] \right)$$

Proof. By definition we have

$$H[(\mu_1 \otimes_p \mu_2)^{(n)} | Y] = \sum_{k=0}^n P(Y = k) \left(- \sum_{\substack{\tau \in \text{Sym}(k) \\ \tau' \in \text{Sym}(n-k)}} P(\tau \circ \tau' | Y = k) \log P(\tau \circ \tau' | Y = k) \right)$$

which can be rewritten in terms of the product entropy as

$$\sum_{k=0}^n P(Y = k) H[\mu_1^{(k)}, \mu_2^{(n-k)}].$$

Notice that we used the fact that the probability $P(\tau \circ \tau' \mid Y = k)$ can be factorized as follows:

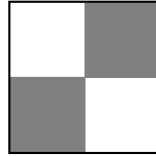
$$P(\tau \circ \tau' \mid Y = k) = \frac{P(Y = k) \cdot \mu_1^{(k)}(\tau) \cdot \mu_2^{(n-k)}(\tau')}{P(Y = k)} = \mu_1^{(k)}(\tau) \cdot \mu_2^{(n-k)}(\tau').$$

This factorization allows us to express the entropy $H[(\mu_1 \otimes_p \mu_2)^{(n)} \mid Y = k]$ in terms of the entropies $H[\mu_1^{(k)}]$ and $H[\mu_2^{(n-k)}]$, resulting in the expression

$$H[\mu_1^{(k)}, \mu_2^{(n-k)}] = H[\mu_1^{(k)}] + H[\mu_2^{(n-k)}],$$

since $\mu_1^{(k)}$ and $\mu_2^{(n-k)}$ are independent. This completes the proof. $\square$

Proposition 3.4. *Let μ be the permuton obtained by placing a uniform distribution with total mass 1 over $[0, \frac{1}{2}] \times [0, \frac{1}{2}] \cup [\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$.*



Then we have $H(\mu) = \Theta(n \log n)$.

Proof. Let λ represent the Lebesgue measure on the unit square. By Theorem 3.3, we can express the entropy of μ as follows:

$$H(\mu \mid Y) = \sum_{k=0}^n \frac{\binom{n}{k}}{2^n} \left(H[\lambda^{(k)}] + H[\lambda^{(n-k)}] \right),$$

Expanding each entropy term, we obtain:

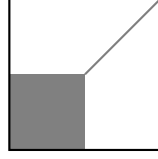
$$H(\mu \mid Y) = \sum_{k=1}^{n-1} \frac{\binom{n}{k}}{2^n} (k \log k + (n-k) \log(n-k)).$$

To analyze the asymptotics of $H(\mu^n)$ of this sum, we observe that by symmetry, it suffices to estimate the expected value $\mathbb{E}[X \log X]$ for $X \sim \text{Binomial}(n, 1/2)$. This expectation asymptotically behaves as $\Theta(n \log n)$, a result from the known asymptotics for binomial entropy.

To interpret this result, observe that $H(\mu \mid Y)$ only provides a lower bound for the actual entropy $H(\mu)$. Since an upper bound of $\Theta(n \log n)$ holds for the entropy of any permuton, this bound applies to $H(\mu)$ as well. Thus, combining the lower bound from $H(\mu \mid Y)$ and the universal upper bound for permuton entropy, we conclude that $H(\mu) = \Theta(n \log n)$.

Therefore, we conclude that $H(\mu) = \Theta(n \log n)$ as desired. $\square$

Proposition 3.5. Let μ be the following permuton: the Lebesgue measure distributed uniformly over the square $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$, and a uniform distribution with probability $\frac{1}{2}$ supported along the diagonal line segment $\{(x, x) : x \in [\frac{1}{2}, 1]\}$.



Then we have $H(\mu) = \Theta(n \log n)$.

Proof. Essentially the same proof as the previous proposition. Here, by expanding the entropy term we get:

$$H(\mu | Y) = \sum_{k=1}^{n-1} \frac{\binom{n}{k}}{2^n} (k \log k + 0).$$

Therefore, the lower and upper bounds are the same as above.

□

Based on the basic properties of Shannon entropy, we have

$$H[\mu^{(n)} | Y] \leq H[\mu^{(n)}] \leq H[\mu^{(n)}, Y] = H[\mu^{(n)} | Y] + H[Y], \quad (3.1)$$

and thus we get the following:

Corollary 3.6. Let μ be a permuton.

1. If $H(\mu^{(n)}) = c \cdot n + o(n)$ for some $c \in \mathbb{R}$, then the entropy of any weighted product of μ and μ is also $c \cdot n + o(n)$.
2. If $H(\mu^{(n)}) = o(n)$, then the entropy of any weighted product of μ and μ is also $o(n)$.
3. If $H(\mu^{(n)}) = \omega(n)$, then the entropy of any weighted product of μ and μ is also $\omega(n)$.

Proof. 1. Since $H(\mu) = \theta(n)$, assume that $H(\mu^{(n)}) = c \cdot n + o(n)$. By Theorem 3.3, we have $H[(\mu \otimes_p \mu)^{(n)} | Y] = \sum_{k=0}^{n-1} \binom{n}{k} p^k (1-p)^k (c \cdot k + o(n) + c \cdot (n-k) + o(n)) = c \cdot n + o(n)$. Then, by expression 3.1, we have $c \cdot n + o(n) \leq H[(\mu \otimes_p \mu)^{(n)}] \leq c \cdot n + o(n) + \theta(\log n) = c \cdot n + o(n)$. Here, $H(Y) = \theta(\log n)$ is the entropy of the binomial distribution.

2. Since $H(\mu^{(n)}) = o(n)$, we have $H(\mu^{(n)})$ is concave in terms of n . Then, by Jensen's inequality and Theorem 3.3, we have $H[\mu^{(n)} | Y] = \sum_{k=0}^{n-1} \binom{n}{k} p^k (1-p)^k (H(\mu^{(k)}) + H(\mu^{(n-k)})) \leq \sum_{k=0}^{n-1} \binom{n}{k} p^k (1-p)^k (2H(\mu^{(n/2)})) = 2H(\mu^{(n/2)}) = 2o(n/2) = o(n)$. Then, by expression 3.1, we have $o(n) \leq H[(\mu \otimes_p \mu)^{(n)}] \leq o(n) + \theta(\log n) = o(n)$.

3. Essentially same proof as above (convexity instead of concavity for this time).

□

3.2. Quasi pattern avoiding permutons

Earlier (see Proposition 1.11), we discussed that, as a consequence of the Marcus-Tardos theorem, for a pattern-avoiding permuton μ , the growth rate of the Shannon entropy sequence is $O(n)$. However, as we will show later, the reverse does not hold: in 4.1, we present a construction where the entropy growth rate is $O(n)$, yet the permuton is not pattern-avoiding.

In this section, we establish a weaker form of pattern avoidance that is indeed implied by the entropy growth rate being $O(n)$.

Definition 3.7. Let μ be a permuton. We say that μ is *quasi pattern avoiding* if for each $\varepsilon > 0$ there exists a pattern τ such that

$$t(\tau, \mu) \leq \frac{\varepsilon}{|\tau|!}.$$

A permuton μ being pattern-avoiding intuitively means that there exists a pattern τ whose density is arbitrarily smaller compared to what it would be under a uniform distribution.

Theorem 3.8. *If*

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} < \infty$$

then μ is quasi pattern avoiding.

Proof. We prove the theorem by proving the contrapositive: when a permuton μ is not quasi pattern avoiding, then

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} = \infty.$$

Let μ be a permuton such that $\exists \varepsilon > 0 \forall \tau$

$$t(\tau, \mu) > \frac{\varepsilon}{|\tau|!}.$$

It is straightforward to show that the function $x \log(1/x)$ attains its maximum on the interval $(0, 1)$ at $x = 1/e$ (this holds for any base of the logarithm). On the interval $(0, 1/e]$, the function $x \log(1/x)$ is increasing. Using this fact, we obtain the following lower bound:

$$\begin{aligned} H(\mu^{(n)}) &= \sum_{|\tau|=n} t(\tau, \mu) \log \frac{1}{t(\tau, \mu)} > \sum_{\substack{|\tau|=n \\ t(\tau, \mu) \leq 1/e}} t(\tau, \mu) \log \frac{1}{t(\tau, \mu)} \geq \\ &\geq \sum_{\substack{|\tau|=n \\ t(\tau, \mu) \leq 1/e}} \frac{\varepsilon}{n!} \log \frac{n!}{\varepsilon} \geq (n! - 2) \frac{\varepsilon}{n!} \log \frac{n!}{\varepsilon} = \\ &= \varepsilon \log n! - \varepsilon \log \varepsilon - 2 \frac{\varepsilon \log n! / \varepsilon}{n!} \end{aligned}$$

since $\log n! \sim n \log n$ we have that

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} = \infty,$$

which gives the theorem. □

It is not that hard to find an example where the reverse does not hold: Let μ be the permuton shown in Proposition 3.4. and

$$\tau = n(n-1) \dots 1.$$

Then

$$t(\tau, \mu) = (1/2)^{n-1} (1/n!)$$

so even though the growth rate of $H(\mu^n)$ is $\Theta(n \log n)$ we have that μ is quasi pattern avoiding.

Using similar ideas, one can generalize Theorem 3.8 further as follows:

Theorem 3.9.

1. If

$$\liminf \frac{H(\mu^{(n)})}{n} < \infty,$$

then for all $\varepsilon > 0$ and $f(n) = o(n \log n)$ there exists $n > 0$ s.t. at least $1 - \frac{1}{f(n)}$ portion of $\text{Sym}(n)$ has density $< \frac{\varepsilon}{n!}$.

2. If

$$\limsup_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} < \infty,$$

then for all $\varepsilon > 0$ and $f(n) = o(n \log n)$ there exists N such that $n > N$ implies that at least $1 - \frac{1}{f(n)}$ portion of $\text{Sym}(n)$ has density $< \frac{\varepsilon}{n!}$.

Proof. Again, we prove both statements, by proving the contrapositives.

1. Suppose that there exists $\varepsilon > 0$ and $f(n) = o(n \log n)$ s.t. for all $n > 0$ more than

$$\frac{1}{f(n)}$$

portion of $\text{Sym}(n)$ has density $\geq \frac{\varepsilon}{n!}$, this means that

$$\begin{aligned} H(\mu^{(n)}) &= \sum_{|\tau|=n} t(\tau, \mu) \log \frac{1}{t(\tau, \mu)} > \sum_{\substack{|\tau|=n \\ \varepsilon/n! \leq t(\tau, \mu) \leq 1/e}} t(\tau, \mu) \log \frac{1}{t(\tau, \mu)} \geq \\ &\geq \sum_{\substack{|\tau|=n \\ \varepsilon/n! \leq t(\tau, \mu) \leq 1/e}} \frac{\varepsilon}{n!} \log \frac{n!}{\varepsilon} \geq \left(\frac{n!}{f(n)} - 2 \right) \frac{\varepsilon}{n!} \log \frac{n!}{\varepsilon} = \\ &= \frac{\varepsilon \log n!}{f(n)} - \frac{\varepsilon \log \varepsilon}{f(n)} - 2 \frac{\varepsilon \log n! / \varepsilon}{f(n) n!} \approx \varepsilon \frac{n \log n}{f(n)} \end{aligned}$$

which tends to infinity, since $f(n) = o(n \log n)$. This means that

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} = \infty$$

so we are done.

2. Suppose that there exists $\varepsilon > 0$ and $f(n) = o(n \log n)$ such that for every N there exists $n > N$ s.t. more than

$$\frac{1}{f(n)}$$

portion of $\text{Sym}(n)$ has density $\geq \frac{\varepsilon}{n!}$. This gives an increasing sequence n_i for which the same holds, so

$$\begin{aligned} H(\mu^{(n_i)}) &= \sum_{|\tau|=n_i} t(\tau, \mu) \log \frac{1}{t(\tau, \mu)} > \sum_{\substack{|\tau|=n_i \\ \varepsilon/n_i! \leq t(\tau, \mu) \leq 1/e}} t(\tau, \mu) \log \frac{1}{t(\tau, \mu)} \geq \\ &\geq \sum_{\substack{|\tau|=n_i \\ \varepsilon/n_i! \leq t(\tau, \mu) \leq 1/e}} \frac{\varepsilon}{n_i!} \log \frac{n_i!}{\varepsilon} \geq \left(\frac{n_i!}{f(n_i)} - 2 \right) \frac{\varepsilon}{n_i!} \log \frac{n_i!}{\varepsilon} = \\ &= \frac{\varepsilon \log n_i!}{f(n_i)} - \frac{\varepsilon \log \varepsilon}{f(n_i)} - 2 \frac{\varepsilon \log n_i! / \varepsilon}{f(n_i) n_i!} \approx \varepsilon \frac{n_i \log n_i}{f(n_i)} \end{aligned}$$

which tends to infinity as $i \rightarrow \infty$, since $f(n) = o(n \log n)$. This means that

$$\limsup_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} = \infty,$$

so we are done. □

Even though this theorem seems stronger than the previous one, we still suspect that the reverse is not true. To substantiate this hypothesis we show that there exists a permutation for which a large portion of $\text{Sym}(n)$ has reasonably small densities but the corresponding entropy sequence has $\Theta(n \log n)$ growth rate.

Claim 3.10. *Let μ be the permutation shown in 3.4. At least*

$$1 - \frac{3}{n}$$

portion of $\text{Sym}(n)$ has the exact density

$$\frac{1}{n! 2^{n-1}}.$$

in μ .

Proof. To prove this, we show a subset of $S \subset \text{Sym}(n)$ such that

$$\frac{|S|}{n!} \geq 1 - \frac{3}{n}$$

and for any $\tau \in S$,

$$t(\tau, \mu) = \frac{1}{n! 2^n}.$$

We choose S to be the set of permutations in $\text{Sym}(n)$ that can only be sampled by choosing all the points either from the top or the bottom square.

The patterns not contained in S are precisely those that corresponds to samples with points in both the bottom and top squares. These are exactly the patterns $\pi \in \text{Sym}(n)$ for which there is an integer $1 \leq s < n$ which separates the permutation, meaning that for all $0 < i \leq s < j \leq n$

$$\pi(i) < \pi(j)$$

holds.

For a fix s we call the indexes smaller or equal than s the left side and those bigger the right side of the separator s . The patterns not in S which have separator s , are exactly the ones for which the numbers $1, \dots, k$ are contained in the left side and the other $n - k$ in the right side of s .

Let $r(k)$ be the number of patterns with minimal separator k . Then $r(k)$ is the number of patterns π where $\{1, \dots, k\}$ are on the left side of the minimal separator and $\pi(k) \neq k$, so

$$r(k) = (k-1)(k-1)!(n-k)!$$

for $k \neq 1$ and

$$r(1) = (n-1)!$$

The portion of patterns not in S can now be upper bounded:

$$\begin{aligned} \frac{\sum_{k=1}^{n-1} r(k)}{n!} &= \frac{1}{n} + \sum_{k=2}^{n-1} (k-1) \frac{(k-1)!(n-k)!}{n!} = \frac{1}{n} + \frac{1}{n} \sum_{k=2}^{n-1} (k-1) \frac{1}{\binom{n-1}{k-1}} = \\ &= \frac{1}{n} \left(1 + \sum_{k=1}^{n-2} k \frac{1}{\binom{n-1}{k}} \right) \leq \frac{1}{n} \left(1 + \frac{2}{n-1} + \sum_{k=2}^{n-3} k \frac{1}{\binom{n-1}{k}} \right) \leq \\ &\leq \frac{1}{n} \left(1 + \frac{2}{n-1} + \frac{2(n-4)(n-3)}{(n-2)(n-1)} \right) < \frac{3}{n} \end{aligned}$$

Where we used the trivial bound

$$\binom{n}{k} \geq \binom{n}{2}$$

for $2 < k < n-2$. The upper bound gives the claim. □

4. Exotic piecewise linear functions

The following section explores special cases related to extending the formula

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = \int_0^1 \log |f'(x)| dx \tag{4.1}$$

to certain classes of permutons and provides an example where the formula does not apply. Throughout this section, we focus on permutons corresponding to measure-preserving functions (see Definiton 2.5).

4.1. A permuton with accumulating structure

Earlier (see Proposition 1.11), we discussed that, as a consequence of the Marcus-Tardos theorem, for a pattern-avoiding permuton μ , the growth rate of the Shannon entropy sequence is $O(n)$. In this section, we demonstrate that the reverse is not true by providing a construction in which, despite the growth rate being $O(n)$, the permuton is not pattern-avoiding.

The construction is as follows: first, divide the unit square into two equal rectangles with a vertical line at $x = 1/2$. Place a diagonal in the left rectangle. Next, split the right rectangle into two equal parts with a vertical line at $x = 1/4$, and place a diagonal in the leftmost of the two new rectangles.

We continue this process by always placing a diagonal in the leftmost empty rectangle and then dividing the next rectangle into two equal parts. The result of repeating this procedure indefinitely is shown in Figure 4.1.

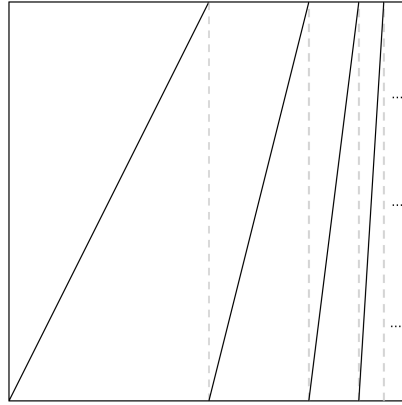


Figure 4.1: The infinite curtain permuton

The corresponding function is measure-preserving, so by (2.5) we can associate a permuton μ with this construction. It is evident that μ is not pattern-avoiding, as every pattern $\tau \in \text{Sym}(n)$ appears with positive probability. We only need to verify that the growth rate of its Shannon entropy sequence is $O(n)$.

Theorem 4.1. *Let μ be the permuton from the above construction. Then*

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = 2 \log 2.$$

Proof. Sampling n points from μ is equivalent with sampling the y coordinates independently, and then choosing a point to be on the i th diagonal with probability $1/2^i$. Since the Shannon-entropy of n independent copies of the distribution

$$p_i = \frac{1}{2^i} \quad i = 1, 2, \dots$$

is exactly $(2 \log 2)n$ we immediately get

$$H(\mu^n) \leq (2 \log 2)n.$$

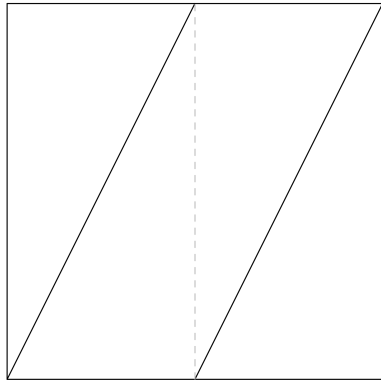
This tells us that

$$\limsup_{n \rightarrow \infty} \frac{H(\mu^n)}{n} \leq (2 \log 2). \quad (4.2)$$

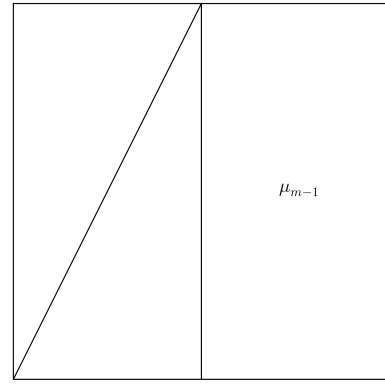
To give a lower bound, first we recursively define a permuton sequence that converges to μ .

Let μ_1 be the permuton obtained by dividing the unit square into two rectangles using the line $x = 1/2$ and placing a diagonal in each rectangle (see Figure 4.2a). Now, we define the m th permuton recursively. To obtain μ_m , we again divide the unit square into two equal rectangles, but instead of placing a second diagonal, we insert μ_{m-1} into the second rectangle (see Figure 4.2b).

Thus, μ_m represents the permuton formed by retaining the first m diagonals of μ and replacing the remaining infinitely many diagonals with a single diagonal. Using Theorem 2.7 it is easy to verify that the sequence μ_m indeed converges to μ .



(a) The permuton μ_1



(b) The permuton μ_m

Figure 4.2: Defining a permuton sequence, which converges to μ

After sampling n points from a permuton let

$$(x_1, y_1), \dots, (x_n, y_n)$$

be the points increasingly ordered by their y -coordinates. Lets define the random variable

$$Y_m \in \{0, 1\}^n$$

the following way:

$$(Y_m)_i = 1 \text{ iff } 1 - 1/2^m < x_i.$$

Lemma 4.2.

$$H(\mu^n | Y_m) \geq H(\mu_m^n | Y_m)$$

Proof. Let us call points with x -coordinate greater than $1 - 1/2^m$ the *end points*. These are the points that yield entries of 1 in Y_m .

When sampling from μ , the end points can form various patterns, whereas, when sampling from μ_m , the only possible pattern for the end points is the identity permutation. This implies that for

a fixed value of Y_m , multiple events from sampling μ merge into a single event when sampling μ_m . This proves that

$$H(\mu^n | Y_m = \mathbf{y}) \geq H(\mu_m^n | Y_m = \mathbf{y}),$$

which is sufficient, as conditional entropy is a convex combination of the entropies of the distributions given each particular condition. □

By the previous lemma, and basic properties of Shannon-entropy, we get the following chain of inequalities:

$$H(\mu^n) \geq H(\mu^n | Y_m) \geq H(\mu_m^n | Y_m) = H(\mu_m^n, Y_m) - H(Y_m) \geq H(\mu_m^n) - H(Y_m).$$

Since Y_m is a vector of n independent Bernoulli $[1 - 1/2^m]$ distributions, therefore

$$H(Y_m) = nH(\text{Bernoulli}[1 - 1/2^m]) \stackrel{\text{def}}{=} \varepsilon_m n.$$

Here $\varepsilon_m \rightarrow 0$ as m tends to infinity. By 4.1 we know that

$$\lim_{n \rightarrow \infty} \frac{H(\mu_m^n)}{n} = H(\mathbf{p}_m)$$

where

$$\mathbf{p}_m = (1/2, 1/4, \dots, 1/2^m, 1/2^m).$$

This gives

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^n)}{n} \geq \lim_{n \rightarrow \infty} \frac{H(\mu_m^n)}{n} - \varepsilon_m = H(\mathbf{p}_m) - \varepsilon_m$$

and so

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^n)}{n} \geq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{H(\mu_m^n)}{n} - \varepsilon_m = \lim_{m \rightarrow \infty} H(\mathbf{p}_m) - \varepsilon_m = H(\mathbf{p}) = 2 \log 2$$

where

$$\mathbf{p} = (1/2, 1/4, 1/8, \dots).$$

This and (4.2) gives the theorem. □

One can generalize the argument above for similar constructions.

Definition 4.3. Let $p \in (0, 1)$. Let μ_m be the permuton defined by the following recursion (similarly as done before):

We call the limit of the above permuton sequence an *infinite p -curtain*.

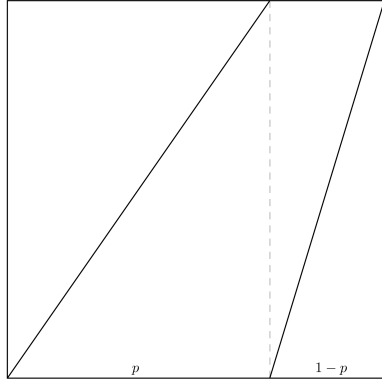
By applying a similar proof with adjusted parameters, we obtain an analogous result to Theorem 4.1 for the general p -curtain.

Theorem 4.4. Let $p \in (0, 1)$ and

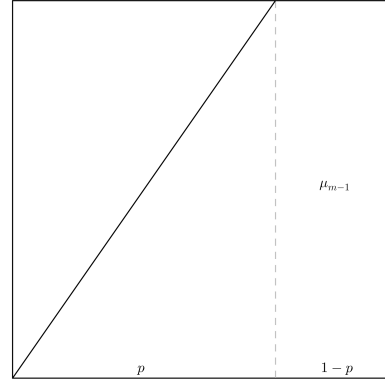
$$\mathbf{p} = (p, p(1-p), p(1-p)^2, \dots).$$

Let μ be the infinite p -curtain. Then

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = H(\mathbf{p}) = \frac{H(\text{Bernoulli}[p])}{p}.$$



(a) The permuton μ_1



(b) The permuton μ_m

This generalizes the formula 4.1 to infinite p -curtains. Of course this is just a very special case of a more general kind of permuton, namely that consisting of infinitely many linear pieces.

Let $f: [0, 1] \rightarrow [0, 1]$ be a measure preserving function consisting only of linear pieces. This means that we can partition $[0, 1)$ into intervals

$$I_l = [a_l, b_l) \quad l \in \mathbb{N}$$

such that $a_l < b_l$, and for $x \in I_l$

$$f(x) = p_l x + c_l \pmod{1}$$

for some constants $p_l, c_l \in [0, 1]$. Let μ_f be the corresponding permuton. Does the formula (4.1) apply to μ_f ? As we will demonstrate later, this is not true in general.

For now let us consider the case when

$$\begin{aligned} c_l &= a_l \frac{1}{a_l - b_l} \\ p_l &= \frac{1}{b_l - a_l}, \end{aligned} \tag{4.3}$$

and in the partition, every interval (except the interval ending in 1 if there exists such) has a right neighbour, so for any l there exists k s.t. $b_l = a_k$.

This gives a similar permuton to an infinite p -curtain, but with pieces of thicknesses $b_l - a_l$.

Definition 4.5. Let $f: [0, 1] \rightarrow [0, 1]$ be a measure preserving function such that (4.3) holds. Then we call μ_f a *general infinite curtain* with distribution $\mathbf{p}$, where $\mathbf{p} = (p_0, p_1, \dots)$ with $p_i = b_i - a_i$.

We suspect that using similar arguments one could prove that

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = H(\mathbf{p})$$

even holds for any general infinite curtain. However, some technical challenges arise since, in a general infinite curtain, multiple (even infinitely many) accumulation points could exist, which would complicate the proof of Lemma 4.2. For a general infinite curtain where 1 is the only accumulation point, the same argument can be applied.

4.2. A diagonally structured family of permutons with interesting properties

Theorem 4.6. *The identity*

$$\lim_{n \rightarrow \infty} \frac{H(\mu^n)}{n} = \int_0^1 \log |f'(x)| dx$$

does not hold for permutons created by measure-preserving functions, consisting of infinitely many pieces.

Proof. We give a family of permutons corresponding to measure preserving functions which consist of infinitely many linear pieces. For piecewise linear functions if the derivative is ± 1 the identity gives 0, so it will suffice to give permutons for which these limits are non-zero and are piecewise linear with slopes ± 1 .

Take a random permutation sequence $\pi_1, \pi_2, \dots$ such that $|\pi_k| = k$. This sequence converges to the Lebesgue measure permuton (λ) with probability one. Hence, for all τ :

$$\lim_{i \rightarrow \infty} t(\tau, \pi_i) = \frac{1}{|\tau|!}$$

with probability one. This means, that $\forall \varepsilon > 0$ and $n \in \mathbb{N}$, $\exists \pi$ permutation such that $\forall \tau, |\tau| \leq n$:

$$\left| t(\tau, \pi) - \frac{1}{|\tau|!} \right| < \varepsilon$$

holds. Divide the unit square into $|\pi| * |\pi|$ identical square pieces. Let μ_π be a permuton, which has the identity measure (or reverse identity measure) in the $(i, \pi[i])$ square piece, and has 0 measure everywhere else. With a small enough ε we can achieve for all $k \leq n$ that

$$H(\mu_\pi^{(k)}) \approx \log k!.$$

Define a $\pi_1, \pi_2, \dots$ permutation sequence, such that for an $n_i \in \mathbb{N}$ and $\varepsilon_i > 0$ defined later and for all $k \leq n_i$:

$$H(\mu_{\pi_i}^{(k)}) > \log k! - \varepsilon_i.$$

Consider the unit square, divided into smaller squares μ_i defined by

$$\mu_i = \left[\sum_{j=1}^{i-1} \frac{1}{2^j}, \sum_{j=1}^i \frac{1}{2^j} \right] \times \left[\sum_{j=1}^{i-1} \frac{1}{2^j}, \sum_{j=1}^i \frac{1}{2^j} \right],$$

where each μ_i occupies a distinct region of the unit square.

Let μ be a permuton constructed by embedding the probability measures μ_{π_i} into these sub-squares μ_i in the following way: in each sub-square μ_i , place the measure μ_{π_i} , scaled appropriately to fit within the bounds of μ_i . Explicitly, μ assigns zero measure outside of these squares,

and within each μ_i , the measure is a scaled copy of μ_{π_i} . Choose μ_{π_i} to have identities in the subsquares in μ_{π_1} and reverse identities everywhere else, this way the formula will give 0 as desired.

Let Y_i be a random variable, which takes the value of how many points we sampled from the i th block, when sampling. We know, that:

$$H(\mu^{(n_i)}) \geq H(\mu^{(n_i)}|Y_i) \geq \sum_{k=0}^{n_i} \binom{n_i}{k} \left(\frac{1}{2^i}\right)^k \left(1 - \frac{1}{2^i}\right)^{n_i-k} H(\mu_i^{(k)})$$

Since $H(\mu_i^{(k)}) \geq \log k! - \varepsilon_i$

$$H(\mu^{(n_i)}) \geq \left(\sum_{k=0}^{n_i} \binom{n_i}{k} \left(\frac{1}{2^i}\right)^k \left(\frac{1}{2^i}\right)^{n_i-k} \log k! \right) - \varepsilon_i.$$

Let $X_i \sim \text{Binomial}\left(n_i, \frac{1}{2^i}\right)$. The sum on the right hand side can be written up as the expected value of $\log X_i!$

Since $\log X_i!$ is convex we can use the Jensen inequality:

$$H(\mu^{(n_i)}) \geq \mathbb{E}[\log(X!)] - \varepsilon_i \geq \log(\Gamma(\mathbb{E}[X] + 1)) - \varepsilon_i = \log\left(\Gamma\left(\frac{n_i}{2^i} + 1\right)\right) - \varepsilon_i$$

Where Γ is the gamma function. We know that $\Gamma(n+1) = n!$ and $\mathbb{E}[\text{Binomial}(n, p)] = np$.

Choosing the right n_i and ε_i values we can see that the value of

$$\frac{H(\mu^{(n_i)})}{n_i}$$

can be arbitrarily large, so we can achieve that

$$\limsup_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} > 0.$$

□

5. Open problems

Conjecture 5.1. $H(\mu^n)$ is non-decreasing for all permutons, with equality if and only if the permuton is the identity permuton, which corresponds to the identity function on $[0, 1]$, or the reverse identity permuton, which corresponds to the function $f(x) = 1 - x$ on $[0, 1]$.

We attempted to prove Conjecture 5.1 but were unsuccessful. While exploring different proof techniques, we identified an alternative statement that, if proven, would imply the original conjecture, with the exception of the statement regarding equality. We need to extend the definition of pattern entropy sequence to permutations: for any permutation π , $H(\pi^n)$ denotes the Shannon entropy of the probability measure given by densities of n -length patterns in π .

Conjecture 5.2. $H(\pi^n)$ is non-decreasing for $n \leq \frac{|\pi|}{2}$, with equality if and only if the permutation is the identity permutation (given by $12 \dots n$) or the reverse identity permutation (given by $n(n-1) \dots 1$).

In our computational work, we programmed calculations for these values using shorter permutations, and the results confirmed the accuracy of the formula.

If $\pi_n \rightarrow \mu$ is a sequence of permutations (such a sequence exists for every permutation), then, as the individual pattern densities converge, we have $H(\pi_n^k) \rightarrow H(\mu^k)$ for every k . Therefore, since for $n > 2k$, $H(\pi_n^k) \leq H(\pi_n^{k+1})$, if Conjecture 5.2 holds, this inequality should also hold in the limit, thereby confirming Conjecture 5.1, save for the case of equality, as the limiting inequality might be saturated in the limit, even if the ones in the sequence are sharp. This could be handled by a quantitative form of Conjecture 5.2, which includes a lower bound on the growth of $H(\pi^n)$.

Acknowledgements

We would like to thank our supervisor, Balázs Maga, for his invaluable guidance throughout this project, providing quick and insightful feedback, encouragement, knowledge, and direction. We are also grateful for his support during the REU research program, his encouragement to develop and submit this work as a TDK thesis, and his assistance at every step of the way. Additionally, we extend our thanks to our REU group mates, Bowen Li and Aayan Pathan, for their collaboration, insights, and support throughout this research journey.

Author details

Péter László Fazekas, ELTE Faculty of Sciences Mathematics MSc, peterzasz@gmail.com
Balázs Szepesi, ELTE Faculty of Sciences Mathematics BSc, balazsszepesi@icloud.com
Sára Szepessy, ELTE Faculty of Sciences Mathematics BSc, szepessy.sara@gmail.com

References

- [1] C. Hoppen, Y. Kohayakawa, C. G. Moreira, B. Rath, R. M. Sampaio. “Limits of permutation sequences.” arXiv, 2011. [Online]. Available: <https://arxiv.org/abs/1103.5844>.
- [2] F. Garbe, J. Hladký, G. Kun, K. Pekárková. “On pattern-avoiding permutations” arXiv, 2022. [Online]. Available: <https://arxiv.org/abs/2208.12712>.
- [3] A. Marcus, G. Tardos. “Excluded permutation matrices and the Stanley–Wilf conjecture.” *Journal of Mathematical Sciences*, vol. 124, no. 1, 2003, pp. 123–152. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0097316504000512>.
- [4] “Shannon Entropy and Kullback-Leibler Divergence”, 1990. [Online]. Available: <https://www.stat.cmu.edu/~cshalizi/754/2006/notes/lecture-28.pdf>.

- [5] Y. Coudène, "Ergodic Theory and Dynamical Systems", Universitext, Springer-Verlag London, 2016. Available: <https://www.ime.usp.br/~sylvain/RokhlinFormulaEntropy.pdf>

6. Alapfogalmak, jelölések

6.1. Permutációk, permutonok

Munkánk során permutációsorozatok határértékét, az úgynevezett permutonokat vizsgáltuk. Egy n hosszú permutáció nem más, mint az $\{1, 2, \dots, n\}$ számok egy sorbarendezeése. Ezen permutációk halmazát $Sym(n)$ jelöli. Például $231 \in Sym(3)$. Egy π permutáció hosszát $|\pi|$ -vel jelöljük.

Definíció 6.1. Ha adott egy $\pi \in Sym(n)$, és egy $\tau \in Sym(k)$ permutáció, ahol $k \leq n$, akkor definiálhatjuk a τ permutáció *sűrűségét* a π permutációban, mint annak a valószínűségét, hogy k különböző elemet sorsolva a π permutációból a τ permutációt látjuk. Ilyenkor a τ permutációt mintának is hívjuk, és a τ minta sűrűségét a π permutációban $t(\tau, \pi)$ -vel jelöljük.

Például a $\pi = 132$ permutációból az első és második elemet sorsolva az 13 számokat kapjuk, ami értelemszerűen az 12 permutációnak felel meg. Az első és utolsó elemet sorsolva az 12 a második és utolsó elemet sorsolva pedig a 21 permutációt kapjuk. Ezek szerint $t(12, 132) = 2/3$ és $t(21, 132) = 1/3$.

Definíció 6.2. Azt mondjuk, hogy egy $\pi_1, \pi_2, \dots$ permutációsorozat konvergens, ha $i < j$ esetén $|\pi_i| < |\pi_j|$, és tetszőleges τ minta esetén a $t(\tau, \pi_n)$ valós számsorozat konvergens.

Definíció 6.3. Legyen adott egy μ valószínűségi mérték a $[0, 1]^2$ egységnégyzeten, melynek a marginális eloszlásai egyenletesek, azaz tetszőleges $Z \subset [0, 1]$ Borel-mérhető halmaz esetén $\mu(Z \times [0, 1]) = \mu([0, 1] \times Z)$ megegyezik a Z halmaz Lebesgue-mértékével. Ekkor a μ mértéket *permutonnak* hívjuk.

Ha egy μ permutonból uniform véletlen kisorsoljuk az $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ pontokat, akkor ezen pontoknak megfeleltethetünk egy $\pi \in Sym(n)$ mintát a következő módon: Jelölje $(x_1^*, x_2^*, \dots, x_n^*)$ és $(y_1^*, y_2^*, \dots, y_n^*)$ az $x_1, x_2, \dots, x_n$ és $y_1, y_2, \dots, y_n$ értékek növekvő sorrendbe rendezett sorát. Ekkor a sorsolás eredménye az a π permutáció lesz amelyre $\pi(i) = j$ pontosan akkor teljesül, ha van olyan l amelyre $x_i^* = x_l$ és $y_j^* = y_l$.

Például, ha egy μ permutációból az $(1/3, 1/3)$ és $(2/3, 2/3)$ pontokat sorsoljuk, akkor az 12 mintát, míg ha az $(1/3, 2/3)$ és $(2/3, 1/3)$ pontokat sorsoljuk, akkor a 21 mintát kapjuk.

Definíció 6.4. Ha adott egy μ permuton, és egy k hosszú τ minta, akkor definiálhatjuk a τ minta *sűrűségét* a μ permutonban mint annak a valószínűségét, hogy az előbb leírt módon k pontot sorsolva μ -ból a τ mintát kapjuk. Ezt a sűrűséget $t(\tau, \mu)$. A sűrűségek segítségével kapunk egy eloszlást az n hosszú permutációkon.

Ennek segítségével már definiálhatjuk, hogy mikor igaz, hogy egy permutációsorozat egy μ permutonhoz tart.

Definíció 6.5. Azt mondjuk, hogy egy $\pi_1, \pi_2, \dots$ permutációsorozat tart a μ permutonhoz, ha tetszőleges τ permutációra $\lim_{n \rightarrow \infty} t(\tau, \pi_n) = t(\tau, \mu)$.

Hoppen és társai megmutatták, hogy tetszőleges konvergens $\pi_1, \pi_2, \dots$ permutációsorozathoz létezik olyan μ permuton, hogy a permutációsorozat tart hozzá. Fordítva tetszőleges permutonhoz létezik olyan permutációsorozat, amely tart hozzá [1].

6.2. Entrópia és mintakerülő permutációk

Egy μ permutonban a mintasűrűségek tetszőleges n esetén meghatároznak egy eloszlást az n hosszú permutációk halmazán, ahol egy $\pi \in \text{Sym}(n)$ permutáció valószínűsége $t(\pi, \mu)$. Ezt az eloszlást μ^n jelöli. Kutatásunk középpontjába ezen eloszlások entrópiájának vizsgálata került.

Definíció 6.6. Legyen adott egy X diszkrét valószínűségi változó, melynek értékkészlete R . Ezen valószínűségi változó Shannon-entrópiája vagy egyszerűen entrópiája a

$$H(X) \stackrel{\text{def}}{=} -E(\log P(X)) = -\sum_{x \in R} P(X=x) \log P(X=x)$$

valós szám. Hasonlóan, ha ξ egy diszkrét valószínűségi eloszlás az Ω halmazon, akkor ennek entrópiája a

$$H(\xi) \stackrel{\text{def}}{=} -\sum_{\omega \in \Omega} \xi(\omega) \log \xi(\omega)$$

valós szám.

Egy eloszlás entrópiája megmutatja, hogy az eloszlás mennyire „kaotikus”. Például legyen μ a Lebesgue-mérték megszorítása az egységnégyzetre. Ekkor könnyen meggondolható, hogy μ egy permuton, és a hozzá tartozó μ^n eloszlás az egyenletes eloszlás. Ezek szerint

$$H(\mu^n) = \log n! \approx n \log n.$$

Nagyságrendileg ez a legnagyobb permuton-entrópia, amit csak kaphatunk. Dolgozatunk során a $H(\mu^n)$ entrópiasorozatot vizsgáltuk különböző permutonokra.

Definíció 6.7. Azt mondjuk, hogy a $\pi \in \text{Sym}(n)$ permutáció *elkerüli* a $\tau \in \text{Sym}(k)$ mintát (ahol $k \leq n$), ha $t(\tau, \pi) = 0$. Hasonlóan a μ permuton *elkerüli* a $\tau \in \text{Sym}(k)$ mintát, ha $t(\tau, \mu) = 0$.

Egy permutáció/permuton *mintakerülő*, ha létezik olyan minta amelyet elkerül.

A mintakerülés egy fontos fogalom a permutációk vizsgálásánál. Marcus és Tardos belátta a következőt [3]:

Tétel 6.1. Tetszőleges τ permutáció esetén létezik egy C_τ konstans, hogy a τ -kerülő n hosszú permutációk száma legfeljebb C_τ^n .

Ezen tétel következményeként megkapjuk a következőt:

Következmény 6.8. Ha egy μ permuton mintakerülő, akkor $H(\mu^n) = O(n)$.

Ez motivált minket abban, hogy megvizsgáljuk a következő kérdést:

Kérdés 6.9. Van további kapcsolat a $H(\mu^n)$ nagyságrendje és a mintakerülés között?

Ha adott egy $f: [0, 1] \rightarrow [0, 1]$ mértéktartó függvény, akkor ahhoz definiálhatunk egy permutont a következő módon:

Definíció 6.10. Legyen $f: [0, 1] \rightarrow [0, 1]$ egy mértékértartó függvény. Legyenek $A, B \subset [0, 1]$ mérhető halmazok. Ekkor legyen

$$\mu_f(A \times B) = \lambda_1(f^{-1}(B) \cap A)$$

ahol λ_1 az 1 dimenziós Lebesgue-mértéket jelöli. μ_f természetesen tetszőleges mérhető halmazra kiterjeszthető, és ez a kiterjesztett mérték egy permuton lesz. Az így kapott permutont az f függvényhez tartozó permutonnak hívjuk.

A $H(\mu^n)$ entrópia sorozat asszimptotikájának pontos meghatározása nem egyszerű feladat. Témavezetőnk Maga Balázs belátta, hogy ez abban az esetben, ha a μ permuton egy olyan mértékértartó $f: [0, 1] \rightarrow [0, 1]$ függvénynek felel meg, amely véges sok folytonosan differenciálható darabból áll, akkor

$$\lim_{n \rightarrow \infty} H(\mu^n)/n = \int_0^1 \log |f'(x)| dx.$$

Ez a következő kérdéshez vezetett minket:

Kérdés 6.11. Működik-e a formula abban az esetben, ha f nem véges, hanem végtelen sok folytonosan differenciálható darabból áll?

7. Módszerek, eredmények

Munkánk során különböző permutonkat konstruáltunk, és ezek entrópiasorozatainak asszimptotikus értékét számoltuk ki. A triviális esetek vizsgálata után úgy készítettünk újabb bonyolultabb permutonokat, hogy korábbiakat ragasztottunk össze. A kapott permutonok entrópiáját a Shannon-entrópia alapvető tulajdonságait és elemi kombinatorikát használva számoltuk ki.

Az 6.11. kérdés kapcsán a következő eredményeket értük el:

- A mértékértartó függvények egy olyan családjára esetén ahol a függvények végtelen sok lineáris darabból állnak megmutattuk, hogy működik a formula.
- Konstruáltunk egy ellenpéldát, ahol nem működik a formula.

Az 6.9. kérdés kapcsán a következő eredményeket értük el:

Létrehoztunk egy végtelen sok lineáris darabból álló f függvényhez tartozó μ_f permutont amely nem mintakerülő, viszont $H(\mu^n) = O(n)$. Ezzel bemutattuk, hogy ekvivalencia nem áll fenn a mintakerülés és az entrópiasorozat $O(n)$ nagyságrendje között. Ezután megvizsgáltuk, hogy a $H(\mu^n) = O(n)$ nagyságrend okoz-e mégis valami mintakerüléshez hasonlót. Bevezettük az ε mintakerülés fogalmát:

Definíció 7.1. Azt mondjuk, hogy egy μ permuton ε -mintakerülő, ha létezik olyan n és olyan n hosszú τ permutáció melyre $t(\tau, \mu) \leq \varepsilon/n!$. Ilyenkor azt mondjuk, hogy a μ permuton a τ mintát ε -elkerüli.

Ezzel kapcsolatban beláttuk a következőt:

Tétel 7.1. Ha

$$\liminf_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} < \infty,$$

akkor tetszőleges $\varepsilon > 0$ és $f(n) = o(\log n)$ esetén létezik $n > 0$, hogy μ az n hosszú permutációk $1 - \frac{1}{f(n)}$ részét ε -kerüli.

Ha

$$\limsup_{n \rightarrow \infty} \frac{H(\mu^{(n)})}{n} < \infty,$$

akkor tetszőleges $\varepsilon > 0$ és $f(n) = o(\log n)$ esetén létezik N , hogy $n > N$ esetén μ az n hosszú permutációk $1 - \frac{1}{f(n)}$ részét ε -elkerüli.