Georgia Southern University

Digital Commons@Georgia Southern

Electronic Theses and Dissertations

Graduate Studies, Jack N. Averitt College of

Summer 2016

Combinatorial Optimization of Subsequence Patterns in Words

Matthew R. Just

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/etd

Part of the Discrete Mathematics and Combinatorics Commons

Recommended Citation

Just, Matthew R., "Combinatorial Optimization of Subsequence Patterns in Words" (2016). *Electronic Theses and Dissertations*. 1464.
https://digitalcommons.georgiasouthern.edu/etd/1464

This thesis (open access) is brought to you for free and open access by the Graduate Studies, Jack N. Averitt College of at Digital Commons@Georgia Southern. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

COMBINATORIAL OPTIMIZATION OF SUBSEQUENCE PATTERNS IN WORDS

by

MATTHEW JUST

(UNDER THE DIRECTION OF HUA WANG AND DANIEL GRAY)

ABSTRACT. Packing patterns in words concerns finding a word with the maximum number of a prescribed pattern. The majority of the work done thus far is on packing patterns into permutations. In 2002, Albert, Atkinson, Handley, Holton and Stromquist showed that there always exists a layered permutation containing the maximum number of a layered pattern among all permutations of length n. Consequently, the packing density for all but two (up to equivalence) permutation patterns up to length 4 can be obtained. In this thesis we consider the analogous question for colored patterns and permutations. By introducing the concept of colored blocks we characterize the optimal permutations with the maximum number of a given colored pattern when it contains at most three colored blocks. As examples, we apply this characterization to find the optimal permutations of various colored patterns and subsequently obtain their corresponding packing densities.

Index Words: permutation, colored permutation, pattern, pattern packing

Combinatorial Optimization of Subsequence Patterns in Words

by

Matthew Just

B.S., Georgia Southern University, 2012

A Thesis Submitted to the Graduate Faculty of Georgia Southern University
in Partial Fulfillment of the Requirements for the Degree

Masters of Science

Statesboro, Georgia

© 2016

Matthew Just

All Rights Reserved

Combinatorial Optimization of Subsequence Patterns in Words

by

Matthew Just

Major Professor:	Hua Wang
Committee:	Daniel Gray (co-advisor)
	Colton Magnant
	Andrew Sills

Electronic Version Approved:

July 2016

DEDICATION

To the memory of Dr. Yingkang Hu, who expected perfection and inspired greatness.

ACKNOWLEDGEMENTS

A patient advisor, a helpful committee, great friends, a caring family, and a loving wife.

Contents

List of Figures

1. Introduction

We begin our discussion with an overview of words. The combinatorics on words has been studied extensively since the late 19th century. The inception of the study is traditionally credited to Axel Thue (1863-1922) who worked extensively on square-free infinite words. An extensive overview on the history can be found in M. Lothaire's[1] *Combinatorics on Words* [8].

In the early 20th century, Major P. A. McMahon pioneered the study of combinatorics in his beautiful work *Combinatory Analysis* [9]. Included in this work were the first steps in the combinatorics of subsequence patterns in words. Most notable is a bijection between 123 avoiding permutations and a special class of partitions. In 1968, Donald Knuth showed in *The Art of Computer Programming* [7] that this bijection extends to any length-3 permutation. More recently the dual question has been considered, that of finding permutations that contain as many copies of another permutation as possible. This is the study of pattern packing, and will be the central focus of this thesis.

1.1. Words.

Consider a nonempty set S and a binary operation on S denoted by $\times$. If $\times$ is associative, then $(S, \times)$ is called a semigroup. Furthermore, if S contains an element ϵ such that $a \times \epsilon = \epsilon \times a = a$, then $(S, \times)$ is called a monoid. The element ϵ is called the identity of S under $\times$.

Let A denote a nonempty set of elements, which we call letters. For now, we will assume the elements in A are distinct, yet have no relation to each other.

[1] M. Lothaire is a pseudonym for a group of Mathematicians, most of whom were students of Marcel-Paul Schützenberger.

Define a binary operation $\times$ that concatenates two elements, and let S be the set generated by A and $\times$. For instance, let $A = \{a, b\}$. Then

$$S = \{a, b, aa, bb, ab, aaa, bbb, aab, aba, baa, \ldots\}.$$

We will refer to A as an alphabet and S as the set of words generated by A. For any word $x \in S$, the length of x, denoted by $|x|$, is the number of letters that make up x. We can define the unique word ϵ such that $|\epsilon| = 0$, referred to as the empty word. If we include $\epsilon \in S$ then our set of words is a monoid. We include several properties of monoids of words below.

Remark 1.1. Let S be a set of words closed under concatenation $\times$. Then

(a) Concatenation $\times$ is associative and non-commutative.

(b) The length of a word $|x|$ is a monoid homomorphism.

(c) The set A is the smallest subset of S that generates all of S by $\times$.

Suppose that $x, y, z \in S$. The associativity of $\times$ is seen immediately. As an example suppose again we have $A = \{a, b\}$ and let $x = aa$, $y = ab$, and $z = bb$. Then

$$
\begin{aligned}
(x \times y) \times z &= (aa \times ab) \times bb \\
&= aaab \times bb \\
&= aaabbb \\
&= aa \times (ab \times bb) \\
&= x \times (y \times z).
\end{aligned}
$$

Also $x \times y = aaab \neq abaa = y \times x$ and thus $\times$ is not commutative. If $x, y \in S$, then

$$|x \times y| = |a_{i_1} \ldots a_{i_k} a_{j_1} \ldots a_{j_k}| = 2k = k + k = |x| + |y|$$

which illustrates that the length is a monoid homomorphism. Finally suppose that A' is any set that generates all of S. Clearly $a, b \in A'$ since $a, b \in S$. Thus $A \subseteq A'$ and A is the smallest set that generates S.

Suppose we want to generate the set of all words in a concise way. Consider the words in S of length k. Taking the kth-power of the sum of the elements in $A = \{a_i\}_{1 \leq i \leq n}$

$$(a_1 + a_2 + \ldots + a_n)^k = \sum_{1 \leq j \leq n^k} \prod_{1 \leq i \leq n} a_{j_i}.$$

For instance, if $A = \{a, b\}$ and $k = 3$:

$$(a+b)^3 = aaa + aab + aba + abb + baa + bab + bba + bbb.$$

We see that the number of words of length k generated from an alphabet of cardinality n is n^k. A useful observation is that this expression can be viewed as a symmetric function of the variables a_i.

1.2. **Patterns in Words.** We want to build up some tools to discuss patterns in words. By a pattern we mean a word that is in some way contained in another word. For instance the word *aabbac* contains the smaller word *aab*. An important note is that the letters in the pattern *aab* are consecutive in the word *aabbac*. We could also consider the pattern *abc* which is made up of letters read left to right, not necessarily consecutive. We formalize this distinction in the following definition.

Definition 1.2. Let $x = x_1 x_2 \dots x_k$ be a word of length k generated from an alphabet $A = \{a_1, a_2, \dots, a_n\}$. A length-$m$ subword of x is a word $x_i x_{i+1} \dots x_{i+m-1}$ with $1 \leq i \leq k - m + 1$. A length-m subsequence of x is a word $x_{i_1} x_{i_2} \dots x_{i_m}$ with $i_j < i_{j+1}$ for each $1 \leq j \leq m - 1$.

This distinction splits the study of patterns in words into two categories: subword patterns and subsequence patterns. Our focus in this thesis will be on subsequence patterns.

Proposition 1.3. *Let x be a word with $|x| = k$. Then there are exactly $\binom{k}{m}$ length-m subsequences of x.*

Proof. We can assign a variable u to represent inclusion of an element in the subsequence and a variable v to represent exclusion. Thus we can construct the inclusion/exclusion sequences by expanding $(u+v)^k$. Each term in this expansion containing m us represents a unique length-m subsequence. By the Binomial Theorem

$$(u+v)^k = \sum_{m=0}^{k} \binom{k}{m} u^m v^{k-m}$$

and we see there are $\binom{k}{m}$ subsequences of length-m. □

For a word y of length m, we say y occurs as a subsequence pattern in the word x if there is a length-m subsequence of x that is in some sense equivalent to y. We will make this more precise by the following definition:

Definition 1.4. Let A be an alphabet equipped with a relation $\mathcal{R}$, denoted $(A, \mathcal{R})$. For a word $x = x_1 x_2 \ldots x_k$ generated by A, the ordered adjacency matrix of x is the adjacency matrix for the relation $\mathcal{R}$ restricted to

$$\{x_1 \ldots x_k\} \times \{x_1 \ldots x_k\}$$

where rows and columns $1, \ldots, k$ correspond to the elements $x_1, \ldots, x_k$.

Two words x and y generated by A are isomorphic if the ordered adjacency matrix for x and y are identical. Denote this by $x \sim y$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad\qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$
$$abb \qquad\qquad\qquad\qquad bcc$$

FIGURE 1. The two adjacency matrices for the words *abb* and *bcc* showing they are isomorphic under the relation $\mathcal{I}$.

A simple example is the set $(A, \mathcal{I})$ where $\mathcal{I}$ is the identity relation (every element is only related to itself). Suppose the alphabet is given by $A = \{a, b, c\}$. Then the length-3 words *abb* and *bcc* are isomorphic since their ordered adjacency matrices are equivalent (see Figure 1).

1.3. **Counting Occurrences of Patterns.** Suppose now we want to count the number of occurrences of a specific pattern y in a word x. Denote by $f(y, x)$ the number of occurrences of the pattern y in the word x. For example consider the set of words generated by the set $(\{a, b, c\}, \mathcal{I})$. Then we have

$$f(ab, abc) = 3, \tag{1.1}$$

$$f(aa, cbcbcb) = 6, \text{ and} \tag{1.2}$$

$$f(abc, aaabbb) = 0. \tag{1.3}$$

The number of *ab* patterns in (1.1) is equal to the number of subsequences in *abc*. Thus this word has the maximum number of occurrences of this pattern. There are no occurrences of *abc* in *aaabbb* in (1.3), thus this word avoids the pattern abc. It is convenient to define a statistic to measure the number of occurrences of a pattern in a word.

Definition 1.5. Let y and x be two words generated from a set $(A, \mathcal{R})$. The density of y in x is given by

$$\delta(y,x) = \frac{f(y,x)}{\binom{|x|}{|y|}}.$$

Per our examples we have

$$\begin{aligned} \delta(ab, abc) &= 1, \\ \delta(aa, cbcbcb) &= \frac{2}{5}, \text{ and} \\ \delta(abc, aaabbb) &= 0. \end{aligned}$$

Clearly we always have $0 \leq \delta(y,x) \leq 1$. This suggests a probabilistic interpretation of the density: $\delta(y,x)$ is the probability that a random subsequence of x is isomorphic to y.

Consider the simplest nontrivial pattern generated by $(\{a_j\}_{1\leq j\leq n}, \mathcal{I})$, namely aa. If we look at the generating function for the set of all words generated by this alphabet:

$$\frac{1}{1-u(a_1+\ldots+a_n)} = 1 + u(a_1 + \ldots a_n) + u^2(a_1a_1 + a_1a_2 + \ldots + a_na_n) + \ldots$$

where u is a variable tracking the length of each word (a_1a_2 has a factor of u^2 since this word is made up of two letters). We will use the fact that this F is a symmetric function if we treat $a_1 \ldots a_n$ as variables and allow them to group ($a_1a_2a_1 = a_1^2a_2$). (A symmetric function satisfies the property $F(a_1 \ldots a_n) = F(\sigma(a_1) \ldots \sigma(a_n))$ for any permutation $\sigma \in \mathcal{S}_n$, the symmetric group of order n).

Consider the words of length k. These are the words with a coefficient of u^k:

$$(a_1 + a_2 + \ldots + a_n)^k$$

A necessary and sufficient condition for the occurrence of an aa pattern in a word x is the containment of two copies of any a_j. More specifically, if c_j is the number of a_j letters in a word:

$$f(x, aa) = \sum_{j=1}^{n} \binom{c_j}{2}$$

By the multinomial theorem,

$$(a_1 + a_2 + \ldots + a_n)^k = \sum_{c_1+\ldots+c_n=k} \binom{k}{c_1, \ldots, c_n} a_1^{c_1} \ldots a_n^{c_n}$$

Partition	Number of Words, x	$f(aa, x)$
1111	24	0
112	144	1
22	48	2
13	36	3
4	4	6

FIGURE 2. Words of length 4 on 4 letters, stratified by occurrences of aa patterns by partitions.

The sequences $c_1c_2\dots c_n$ are compositions of k. We want to group the terms by partitions of k (that is, not distinguish between compositions that are equivalent under some permutation of the parts). For any such partition $\lambda = \lambda_1 \dots \lambda_k$, write the elements in terms of their multiplicities: $\lambda = k^{b_k}(k-1)^{b_{k-1}} \dots 1^{b_1}0^{b_0}$. Thus corresponding to λ there are in total

$$\binom{k}{\lambda_1 \dots \lambda_k}\binom{n}{b_0 \dots b_k}$$

words of length k corresponding to the partition λ on an alphabet of n letters. Each of these words has in total

$$\sum_{i=1}^{k} \binom{\lambda}{2}$$

total aa patterns.

For example, let $k = 4$ and $n = 4$. There are 5 partitions of 4, namely $1111, 112, 13, 22, 4$. The distribution of words and aa patterns are given in Figure 2.

1.4. **Pattern Packing.** It is often useful to ask a more specific question. For instance: How many words of a certain length avoid a pattern? This is the question of pattern avoidance. In the previous section we considered the pattern aa and words of length k. From our analysis it would be straightforward to determine how many words of length k avoid the pattern aa. For instance if $n = 4$ and $k = 4$ there are 24 words that avoid aa. This could have been proven directly. The number of words that avoid aa are just the words that have all distinct letters. If the number of letters equals the length of the word, this is just the permutations of length 4 ($4! = 24$).

A question dual in nature is this: among all words of length k, what is the maximum number of occurrences of a pattern in a single word? This is the question of pattern packing, and will be the central focus of the rest of this thesis. For the example in the last section, the maximum number of occurrences of aa was equal to the total number of subsequences of a word of length k. Better yet, we determined that there were exactly n words that achieved this maximum. In general it is ambitious to pursue this much detail. For our purposes we will be content simply finding the maximum, understanding that there may be multiple words that achieve it.

Definition 1.6. Let S be the set of words generated by an alphabet $(A, \mathcal{R})$. Let y be a pattern in S and let S_k be the subset of S containing all the words of length k. A word $\hat{x}$ (in general not unique) such that $f(y, \hat{x}) \geq f(y, x)$ for any other $x \in S_k$ is called an optimal word of length k. Denote this maximum number of occurrences of y in an $\hat{x}$ by $f_k(y)$.

It is an important observation that for a fixed y, $f_k(y)$ is a nondecreasing function with k. More importantly is the next proposition, first proved by Fred Galvin (see [10], Theorem 2.1).

Proposition 1.7. *Let S be the set of words generated by an alphabet $(A, \mathcal{R})$. Let y be a pattern in S and let S_k be the subset of S containing all the words of length $k \geq |y|$. Define*

$$\delta_k(y) = \frac{f_k(y)}{\binom{k}{|y|}}$$

Then $\delta_{k-1}(y) \geq \delta_k(y)$ for all $k \geq 2$.

Proof. Consider $f(y, \hat{x}) = f_k(y)$ for some optimal word $\hat{x}$. Let $S_{k-1}(\hat{x})$ be the set of all length-$(k-1)$ subsequences of $\hat{x}$. For any y-pattern in $\hat{x}$, there are exactly $k - |y|$ subsequences in $S_{k-1}(\hat{x})$ that contain this pattern. This is because each length-$(k-1)$ subsequence excludes one element from $\hat{x}$. Of the k subsequences in $S_{k-1}(\hat{x})$, $|y|$ of them will exclude one of the elements that compose this subsequence isomorphic to y. We can then write

$$\begin{aligned} f_k(y) &= \frac{1}{k - |y|} \sum_{z \in S_{k-1}(\hat{x})} f(y, z) \\ \frac{f_k(y)}{\binom{k}{|y|}} &= \frac{\frac{1}{k-|y|} \sum_{z \in S_{k-1}(\hat{x})} f(y, z)}{\binom{k}{|y|}} \\ &= \frac{\frac{1}{k} \sum_{z \in S_{k-1}(\hat{x})} f(y, z)}{\binom{k-1}{|y|}} \end{aligned}$$

Now because $f(y,z) \leq f_{k-1}(y)$

$$\begin{aligned}
\frac{f_k(y)}{\binom{k}{|y|}} &\leq \frac{\frac{1}{k}\sum_{z \in S_{k-1}(\hat{x})} f_{k-1}(y)}{\binom{k-1}{|y|}} \\
&\leq \frac{f_{k-1}(y)}{\binom{k-1}{|y|}}
\end{aligned}$$

□

We see now that $\delta_k(y)$ is a non-increasing sequence for any fixed pattern y. Because $\delta_k(y) \geq 0$, it then makes sense to take the limit of this sequence.

Definition 1.8. Let y be a word generated from an alphabet $(A, \mathcal{R})$. Define the packing density of y as follows

$$\delta(y) = \lim_{k \to \infty} \delta_k(y)$$

Returning to the example in the previous section, let us compute the packing density for the pattern aa. Considering an optimal word $\hat{x}$ of length k, we found that $f_k(aa) = \binom{k}{|y|}$ and thus $\delta_k(aa) = 1$. Taking the limit is trivial, and we find that $\delta(aa) = 1$.

It was taken for granted in this example that the packing density holds a dependence on the number of letters in the generating alphabet. For instance, consider the packing density of the pattern ab over the alphabet $(\{a,b,c\}, \mathcal{I})$. It is straight forward to verify that there is a length-k optimal word with the form

$a^{r_1}b^{r_2}c^{r_3}$. Here we have used exponent notation to signify repeated consecutive letters. It follows that $r_1 + r_2 + r_3 = k$. Basic calculus shows $r_1 = r_2 = r_3 = \frac{k}{3}$ and we have the following calculation for the packing density:

$$\begin{aligned}
\delta_k(ab) &= \frac{f(ab, a^{\frac{k}{3}}b^{\frac{k}{3}}c^{\frac{k}{3}})}{\binom{k}{2}} \\
&= \frac{\binom{3}{2}\binom{\frac{k}{3}}{1}^2}{\binom{k}{2}} \\
&= \frac{2k}{9(k-1)} \\
\delta(ab) &= \frac{2}{9}
\end{aligned}$$

If instead we had an alphabet with an infinite number of distinct letters, we see that the packing density would be equal to one. Unless otherwise specified we will assume the number of letters in our alphabet n equals the length of the word in question k.

2. Permutations

The majority of the previous work on subsequence patterns in words is in the area of permutations. Here we will think of permutations not as bijective maps but rather as a sequence of distinct, ordered letters. By convention we will choose the natural numbers as our alphabet, equipped with the usual total order $<$. For permutations, however, we will require the letters of each word to be totally ordered as well. That is, we will not allow repeated elements in the same word. We can clarify this in the following way.

Let $[k]$ be the totally ordered alphabet consisting of the numbers $1 < 2 < \ldots < k$. The set of permutations of length k, denoted $[k!]$, is the collection of all linear orderings of the letters in $[k]$. Note the notation $[k!]$ is chosen to signify the size of the set. This is easy to see. Starting with the first position in the permutation there are k letters to choose from, then $k-1$ for the second position, and so on. Thus there are $k(k-1)\ldots 1 = k!$ permutations of length k.

Recall that two words are isomorphic if their ordered adjacency matrices are equivalent. Thus two permutations are isomorphic if their elements are ordered in the same way. See Figure 3 for an example.

$$\underset{132}{\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}} \qquad\qquad \underset{586}{\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}$$

Figure 3. The two adjacency matrices for the permutations 132 and 586 showing they are isomorphic under the relation $<$.

Also a permutation τ, called a pattern, of length k occurs in another permutation π of length $n > k$ if there is a length-k subsequence of π that is isomorphic to τ. The number of occurrences of τ in π will again be denoted by $f(\tau, \pi)$. If $f(\tau, \pi) = 0$, π avoids τ. If $\hat{\pi} \in [n!]$ is such that $f(\tau, \hat{\pi}) \geq f(\tau, \pi)$ for any $\pi \in [n!]$, then $\hat{\pi}$ is an optimal permutation and we write $f_n(\tau) = f(\tau, \hat{\pi})$. The packing density of τ is also defined as before

$$\delta(\tau) = \lim_{n\to\infty} \frac{f_n(\tau)}{\binom{n}{k}}$$

For example we have the following:

$$f(12, 13524) = 7 \tag{2.1}$$

$$f(132, 132654) = 10 \tag{2.2}$$

$$f(321, 12435) = 0 \tag{2.3}$$

Here 2.3 shows that 12435 avoids the pattern 321. For 2.2 it can be shown that 132654 is an 132-optimal permutation of length 6. It is not hard to see that the packing density of an increasing/decreasing pattern is trivially equal to 1 (achieved by any increasing/decreasing permutation). In general it is very difficult to find the packing density of nontrivial permutations.

Definition 2.1. Let $\pi = \pi(1)\pi(2)\ldots\pi(n)$ be an arbitrary permutation of length n. Define a pattern $\tau = \tau(1)\tau(2)\ldots\tau(k)$ to be a permutation of length

$k \leq n$. Let a length-k subsequence $\pi(i_1)\pi(i_2)\ldots\pi(i_k)$ of π be isomorphic to τ if $\pi(i_j) < (>)\pi(i_k)$ whenever $\tau(j) < (>)\tau(k)$, for $j < k$. The number of τ-isomorphic subsequences in a permutation π is given by the pattern counting function

$$f(\tau, \pi)$$

If $f(\tau, \pi) > 0$ then π is said to contain τ. If $f(\tau, \mathit{\Pi}) = 0$ then π is said to avoid τ.

The pattern counting function can also be used to count patterns contained in sets of permutations. Let $[n!] = \{\pi_1, \pi_2, \ldots, \pi_{n!}\}$ be the set of all permutations of length n. The number of occurrences of any length k pattern in $[n!]$ is easily found.

Theorem 2.2. *Let $[n!]$ be the set of all permutations of length n and let τ be a pattern of length k. Then the total number of τ-patterns contained in all permutations in $[n!]$ is given by*

$$f(\tau, [n!]) = \frac{n!}{k!}\binom{n}{k}$$

Proof. There are a total of $n!$ permutations contained within $[n!]$ and each permutations contains $\binom{n}{k}$ length-k subsequences. The probability of a randomly selected subsequence being isomorphic to τ is one in $k!$, the total number of length-k patterns. Thus the number of τ-isomorphic subsequences is the total number of length-k subsequences within $[n!]$ multiplied by the expectation value that one of these subsequences is τ-isomorphic. Therefore $f(\tau, [n!]) = \frac{n!}{k!}\binom{n}{k}$. $\square$

As an example consider the pattern $\tau = 132$. Given a relatively short permutation

$$\pi = 35421$$

the quantity $f(132, 35421)$ can be counted directly. Because $n = 5$ there are a total of $\binom{5}{3} = 10$ subsequences. They are given by

354	352	351	342	341
321	542	541	521	421

Clearly the only length-3 subsequence that is isomorphic to 132 is the subsequence 354. Thus $P(132, 35421) = 1$ and 35421 contains one 132 pattern.

Suppose now it is desired to know the total number of 132 patterns contained in the set of all length-4 permutations. Again for small n these can be enumerated directly by examining the 24 permutations of length n and counting the number of 132 patterns:

π	$f(132,\pi)$	π	$f(132,\pi)$	π	$f(132,\pi)$	π	$f(132,\pi)$
1234	0	1243	2	1324	1	1342	2
1423	2	1432	3	2134	0	2143	2
2314	0	2341	0	2413	1	2431	1
3124	0	3142	1	3214	0	3241	0
3412	0	3421	0	4123	0	4132	1
4213	0	4231	0	4312	0	4321	0

By summing all of the patterns it is clear that $f(132,[4!]) = 16 = \frac{4!}{3!}\binom{4}{3}$.

Before going through a review of the literature, let us illustrate how the occurrences of a simple pattern can be found using a generating function.

Definition 2.3. Let n be any positive integer. The q-analog of n, denoted $[n]_q$, is defined as follows:

$$[n]_q = \frac{1-q^n}{1-q}$$

What makes this a q-analog is the fact that $[n]_q \to n$ as $q \to 1$. From this definition it is natural to define the q-factorial:

$$[n]_q! = \prod_{i=1}^{n} \frac{1-q^i}{1-q}$$

which approaches $n!$ as q tends to one. If we expand this out as a polynomial in q we obtain:

$$[n]_q! = 1q^0 + (n-1)q^1 + \frac{(n-2)(n+1)}{2}q^2 + \ldots + 1q^{\binom{n}{2}}$$

The coefficients in this expansion have a combinatorial interpretation. The coefficient of q^i is the number of permutations of length n containing i 12 patterns[2]. For instance, if $n = 3$ we have that there are 2 permutations containing one 12 pattern. These two patterns are 231 and 312.

Let us give a brief combinatorial (number theoretic) proof. Notice that $[n]_q!$ is multiplying (simplified) polynomials of the form $1 + q + \ldots + q^{i-1}$. The $i = 1$ polynomial represents the first letter placed in the polynomial. The next polynomial is $q^0 + q^1$. This represents the choice of adding one 12 pattern (by letting this letter be greater than the first) or continuing to have no 12 patterns (by letting this letter be less than the first). For the $i = 3$ polynomial $q^0 + q^1 + q^2$ we have three options: add no 21 patterns (q^0) by making this letter less than the previous two, add one 21 patterns (q^1) by making this letter have numerical value between the previous two, or add two 21 patterns (q^2) by making this letter larger than the previous two. Continuing for all $1 \leq i \leq n$ gives the desired result.

In this way we can think of permutations in a slightly different way: as a sequence of numbers, $a_0a_1 \ldots a_{n-1}$, for which $0 \leq a_i \leq i$ shows how many 12 patterns are added by each letter. For example, we can write the permutation 146352 as 012131. 1 adds no 12 patterns, 4 adds one, 6 adds two, 3 adds one, 5 adds three, and 2 adds one.

[2]Classically this is the number of inversions in the permutation. An inversion is a 21 pattern which has the same counting sequence as the pattern 12.

Using this analysis we see that there is exactly one permutation of length n that avoids 12 and exactly one optimal permutation. Unfortunately this type of analysis does not work for more complex permutations, and we must stick to asking more specific questions.

We will define two bijective maps on $[n!]$ that will be very useful. They are defined as follows:

Definition 2.4. Let $\pi = \pi(1)\pi(2)\ldots\pi(n)$ be a permutation in $[n!]$. Define the reversal of π as

$$r(\pi) = \pi(n)\pi(n-1)\ldots\pi(1)$$

Similarly define the conjugate of π as

$$c(\pi) = (n-\pi(1)+1)(n-\pi(2)+1)\ldots(n-\pi(n)+1)$$

Let us give a couple examples. We have:

$$r(132) = 231$$

$$r(15243) = 34251$$

$$c(132) = 312$$

$$c(15243) = 51423$$

Proposition 2.5. *Let* $[n!]$ *be the set of permutations of length* n*. Then the reversal and conjugate are bijective maps on* $[n!]$ *with the property that for any* $\pi \in [n!]$ *and pattern* $\tau \in [k!]$*:*

$$f(\tau, \pi) = f(r(\tau), r(\pi)) = f(c(\tau), c(\pi))$$

Proof. Consider any subsequence π^* of π that is isomorphic to τ. Because they are isomorphic, the ordered adjacency matrices of π^* and τ are identical. Now the reversal corresponds to flipping the matrix horizontally and vertically, while the conjugate corresponds to the transpose of the matrix. Performing these operations on the matrix for π results in performing the same operations on π^*. This leaves the matrix for $r(\pi^*)$ equivalent to $r(\tau)$ and $c(\pi^*)$ equivalent to $c(\tau)$. Similarly any subsequence that is not isomorphic to τ will not be isomorphic to $r(\tau)$ or $c(\tau)$ after reversal or conjugation. □

Notice this proposition implies that $\delta(\tau) = \delta(r(\tau)) = \delta(c(\tau))$. We thus have classes of patterns that are, in a packing since, equivalent. Another interesting observation is that $r(123) = c(123)$. This is not true in general, and defines a class of permutations which we call symmetric permutations. Here is an interesting result:

Proposition 2.6. *There are* $\lfloor \frac{n}{2} \rfloor ! 2^{\lfloor \frac{n}{2} \rfloor}$ *symmetric permutations of length* n*.*

Proof. For any symmetric permutation $\pi \in [n!]$, we have that $\pi(n - i + 1) = n - \pi(i) + 1$ for each $1 \leq i \leq n$. Thus $\pi(i) + \pi(n - i + 1) = n + 1$, so that each letter in π is paired with its reflected element to sum to $n-1$. This means that after selecting the first $\lfloor \frac{n}{2} \rfloor$ letters in π, the rest of the permutation is determined. These elements can be rearranged ($\lfloor \frac{n}{2} \rfloor!$ ways) and switched with their reflected element ($2^{\lfloor \frac{n}{2} \rfloor}$ ways). □

2.1. **Permutation Pattern Avoidance.** In this section we give a brief overview of the key results for permutation subsequence patterns. The majority of the early results are on pattern avoidance. Not until the last couple of decades has pattern packing been studied more extensively.

Recall that a permutation $\pi \in [n!]$ avoids a pattern $\tau \in [k!]$ if $f(\tau, \pi) = 0$. We will be most interested in the case that $k \leq n$, since otherwise there is nothing to study.

Definition 2.7. Let $[n!]$ be the set of permutations of length n. For a fixed pattern $\tau \in [k!]$, define the subset $Av_n(\tau) \subseteq [n!]$ as follows:

$$Av_n(\tau) = \{\pi \in [n!] : f(\tau, \pi) = 0\}$$

For convenience let $av_n(\tau) = |Av_n(\tau)|$.

Much of the early work on pattern avoidance involved finding the counting sequence for $av_n(\tau)$ for various patterns. We have already seen that $av_n(12) = av_n(21) = 1$. Note this equivalence follows from $r(12) = 21$. It is also of interest to find patterns that have the same sequence $av_n(\tau)$ that are not related by reversal or conjugation.

Definition 2.8. Let $\tau, \sigma \in [k!]$ be two patterns. If $av_n(\tau) = av_n(\sigma)$ for all $n \geq 1$ then τ and σ are said to be Wilf-equivalent.

We see that all patterns of length one and two are trivially Wilf-equivalent. We now turn our attention to the patterns of length three.

Proposition 2.9. *For the pattern* 132 *(and all equivalents) we have* $av_n(132) = \mathcal{C}_n$*, where* $\mathcal{C}_n$ *are the Catalan numbers defined as follows:*

$$\mathcal{C}_n = \frac{1}{n+1}\binom{2n}{n}$$

Proof. Consider any permutation of length $n+1$, $n \geq 0$. We will define a recursion based on the location of the largest element $n+1$. If $n+1$ is located at location i, then there is a length $i-1$ 132 avoiding permutation to the left of $n+1$ and a length $n+1-i$ 132 avoiding permutation to the right of $n+1$. Thus summing over all i:

$$av_{n+1}(132) = \sum_{i=1}^{n+1} av_{i-1}(132)av_{n-i}(132), \quad n \geq 0 \tag{2.4}$$

where $av_0(132) = 1$. We can determine the generating function $C(x) = \sum_{n=0}^{\infty} av_n(132)x^n$ by multiplying each side by x^n and summing over all n:

$$
\begin{aligned}
\sum_n av_{n+1}(132)x^n &= \sum_n \sum_{i=1}^{n+1} av_{i-1}(132)av_{n+1-i}(132)x^n \\
\frac{1}{x}\sum_n av_{n+1}(132)x^{n+1} &= \sum_n \sum_{i=0}^{n} av_i(132)av_{n-i}(132)x^n \\
\frac{1}{x}\left(\sum_n av_n(132)x^n - 1\right) &= \sum_n av_n(132)x^n \sum_n av_n(132)x^n \\
\frac{1}{x}\left(C(x)-1\right) &= C(x)^2 \\
C(x) &= 1 + xC(x)^2
\end{aligned}
$$

Using the quadratic formula we find $C(x)$ implicitly:

$$C(x) = \frac{1 \pm \sqrt{1-4x}}{2x}$$

Now we must have that $C(x) \to 1$ as $x \to 0^+$ so that

$$C(x) = \frac{1-\sqrt{1-4x}}{2x} = \frac{2}{1+\sqrt{1+4x}}$$

Recall the generalized binomial theorem:

$$(x+y)^r = \sum_{k=0}^{\infty} \frac{(r)_k}{k!} x^i y^{r-i}$$

where $(r)_k = r(r-1)(r-2)\dots(r-k+1)$ is the falling factorial. Thus

$$\sqrt{1-4x} = \sum_{k=0}^{\infty} \frac{(\frac{1}{2})_k}{k!}(-4x)^n$$

Furthermore,

$$\begin{aligned}\frac{(\frac{1}{2})_k}{k!} &= \frac{1}{k!}\prod_{i=0}^{k-1}\frac{(-1)(2k-1)}{2} \\ &= \frac{(-1)^{k+1}}{2^k k!}(2k-3)!!\end{aligned}$$

where $n!! = n(n-2)\ldots 2$ if n is even, $n!! = n(n-2)\ldots 1$ if n is odd. For an odd integer,

$$(2n-1)!! = \frac{(2n)!}{(2n)!!} = \frac{(2n)!}{2^n n!}$$

so that we now have

$$\begin{aligned}\frac{(\frac{1}{2})_k}{k!} &= \frac{(-1)^{k+1}(2k)!}{(k!)^2 4^k (2k-1)} \\ &= \frac{(-1)^{k+1}}{(2k-1)4^k}\binom{2k}{k}\end{aligned}$$

All together now

$$
\begin{aligned}
C(x) &= \frac{1}{2x}\left(1-\sum_{k=0}^{\infty}\frac{(-1)^{k+1}}{4^k(2k-1)}\binom{2k}{k}(-4x)^k\right)\\
&= \frac{1}{2x}\left(1-1-\sum_{k=1}^{\infty}\frac{(-1)^{2k+1}}{2k-1}\binom{2k}{k}x^k\right)\\
&= \frac{1}{2x}\sum_{k=0}^{\infty}\frac{1}{2k+1}\binom{2k+2}{k+1}x^{k+1}\\
&= \sum_{k=0}^{\infty}\frac{1}{k+1}\binom{2k}{k}x^k
\end{aligned}
$$

Comparing coefficients we have the result. □

Since the patterns $132, 213, 231, 312$ are equivalent by reversal or conjugation, we need only consider the pattern 321 (equivalently (123). This question was first considered by MacMahon [9], though in a slightly different context.

Proposition 2.10. *All patterns of length* 3 *are Wilf-equivalent.*

Proof. We will show that the set of 321-avoiding permutations is also counted by the Catalan numbers by showing $av_n(321) = \mathcal{C}_n$. A bijection between permutations and standard Young tableaus will be used, first presented by Schensted [12].

Consider a permutation $\pi = \pi(1)\ldots\pi(n)$ written in two-line notation

$$
\begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}
$$

Create a tableau in the following way. Place the elements $p(1), p(2), \ldots p(j-1)$ in the first row, where j is the first index such that $p(j) < p(j-1)$. For instance take the permutation 23154. The tableau will start out like

$$\begin{array}{|c|c|} \hline 2 & 3 \\ \hline \end{array}$$

Next take the smallest $p(i)$, $1 \leq i \leq j-1$, such that $p(j) < p(i)$. Replace $p(i)$ with $p(j)$ and move $p(i)$ to the second row and first column of the tableau. In our example we then have

$$\begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$$

Continue adding the elements $p(j+1)\ldots$, repeating the procedure until all elements are placed. The final tableau for 23154 is given by

$$\begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 5 & \multicolumn{1}{c}{} \\ \cline{1-2} \end{array}$$

Now to ensure the uniqueness of this tableau we need to record the order for which the squares appeared. Thus we have for 23154 a second tableau called the insertion tableau:

1	2	4
3	5	

If we denote the set of all standard Young tableaus of shape λ by f_λ, then we have a mapping from

$$[n!] \to \bigcup_{\lambda} (f_\lambda \times f_\lambda)$$

It can be shown that this is a bijective map, and that any permutation avoiding the permutation $(k+1)k \ldots 1$ will be mapped to a tableau with at most k rows.

Thus to count the number of 321 avoiding permutations, it suffices to count pairs of standard Young tableaus of a fixed shape and size n with at most two rows.

To do this we use the following one-to-one mapping based on the lattice permutations proposed by MacMahon [9]. Start with the first tableau for a permutation π. Create a word generated by the alphabet $\{\alpha, \beta\}$ where letter i is α if i is in the first row or β if it is in the second. Append to this word a second word generated by the same alphabet, however assigning β to letter $n-j+1$ if j is in the second row if the insertion tableau, and α otherwise.

Consider again the example of 23154. We then have the word $\alpha\beta\alpha\alpha\beta\alpha\beta\alpha\beta\beta$. It can be verified that that these words have n αs and n βs, and that any initial subword has at least as many αs as βs.

A word of this type will satisfy the following recursion. Let $q(m, n)$ denote the number of words of length $m+n$, $m > n$, m giving the number of αs, n giving the number of βs, and any initial subword having more αs than βs. If we remove the last letter of any of these words, we obtain any of the words with either $m-1$ αs or $n-1$ βs. Thus

$$q(m, n) = q(m, n-1) + q(m-1, n)$$

It is easily verified that a general solution of this difference equation is given by

$$\frac{(m+n)!}{(m+s)(n-s)!}$$

where s is an integer. Taking a linear combination of these solutions we have

$$q(m, n) = \sum_{s=0}^{n} a_s \frac{(m+n)!}{(m+s)!(n-s)!}$$

If $m = 0$, then this reduces to

$$q(0,n) = \sum_{s=0}^{n} a_s \binom{n}{s}$$

Now if $n = 0$ then $q(0,0) = 1$ and thus

$$1 = a_0$$

Furthermore, if $n = 1$ then $q(0,1) = 0$ (and for larger values of n), thus

$$0 = a_0 + a_1$$

which shows that $a_1 = -1$. Taking any larger value of n yields that $a_n = 0$, therefore

$$\begin{aligned} q(m,n) &= \frac{(m+n)!}{m!n!} - \frac{(m+n)!}{(m+1)!(n-1)!} \\ &= \frac{(m+1)(m+n)! - n(m+n)!}{(m+1)!n!} \\ &= \frac{(m+n)!(m-n+1)}{(m+1)!n!} \end{aligned}$$

Finally setting $m = n$ shows that $av_n(321) = \mathcal{C}_n$. $\square$

Now that we have shown the Wilf-equivalence of patterns of length-3, we will show some other results. Our attention is now turned to length-4 patterns. The following gives examples of patterns that are nontrivially Wilf-equivalent.

Stankova [14] showed first that the two length four patterns 1342 and 2413 are Wilf-equivalent. Later with West [13], he showed that any length-4 pattern beginning with a 231 pattern is Wilf-equivalent to a length-4 pattern beginning with a 312 pattern. Backelin, West and Xin [2] showed a similar Wilf-equivalence for length-4 patterns beginning with 123 and 321 patterns.

In total there are $4! = 24$ length four patterns. Considering the Wilf-equivalence classes formed by the symmetries of the patterns and the three results stated above, there are only three length-4 Wilf-equivalence classes:

$$\{1234, 4321, 2341, 3214, 1432, 4123, 1243, 3421, 2134, 4312, 2143, 3412\}$$

$$\{1342, 2413, 4213, 3124, 2413, 3142, 2314, 4132, 1423, 3241\}$$

$$\{1324, 4231\}$$

Thus there are only three values to enumerate, however only two have been found exactly. The number of permutations avoiding 1234 (and all Wilf-equivalent patterns) was found by Gessel [5]. The number of permutations avoiding 1342 (and thus all Wilf-equivalent patterns) was enumerated by Bóna [3]. The exact enumeration for $av_n(1324)$ is an open question.

2.2. **Layered Permutations.** We will now introduce a type of permutation, called a layered permutation, that is essential to the study of permutation patterns. We first define a direct sum of two permutations:

Definition 2.11. Let $\pi \in [n!]$ and $\sigma \in [k!]$ be two permutations. The direct sum of π and σ is given by

$$\pi \oplus \sigma = \pi(1)\pi(2)\dots\pi(n)(\sigma(1)+n)(\sigma(2)+n)\dots(\sigma(k)+n)$$

A few important properties of the direct sum are summarized next.

Proposition 2.12. *The direct sum is an associative, non-commutative binary operation on* $\cup_i[i!]$.

Proof. Let $\pi \in [n!]$, $\sigma \in [k!]$, and $\phi \in [\ell!]$. Clearly we have $\pi \oplus \sigma \in [(n+k)!]$. Then

$$\begin{aligned}
(\pi \oplus \sigma) \oplus \phi &= \pi(\sigma(1)+n)\dots(\sigma(k)+n) \oplus \phi \\
&= \pi(\sigma(1)+n)\dots(\sigma(k)+n)(\phi(1)+k+n)\dots(\phi(\ell)+k+n) \\
&= \pi \oplus \sigma(1)\dots\sigma(k)(\phi(1)+k)(\phi(\ell)+k) \\
&= \pi \oplus (\sigma \oplus \phi)
\end{aligned}$$

and thus $\oplus$ is associative. Next consider $123 \oplus 132 = 123465$. Since $132 \oplus 123 = 132456 \neq 123465$, $\oplus$ is not commutative. $\square$

Denote a decreasing permutation by $\bar{n} := n(n-1)\dots 1$. Recall that for pattern avoidance and pattern packing, the nicest results were for such decreasing patterns. We consider now direct sums of decreasing permutations.

Definition 2.13. Let $\{\bar{n}_i\}_{1 \leq i \leq k}$ be a set of decreasing permutations. A layered permutation is of the form

$$\bigoplus_{i=1}^{k} \bar{n}_i$$

Let $\mathcal{L}_n$ denote the set of all layered permutations of length n.

A graph of the layered permutation $1 \oplus 21 \oplus 321 = 132654$ is shown in Figure 4. There is a useful way to characterize layered permutations in terms of pattern avoidance.

Proposition 2.14. *Let $Av_n(\{312, 231\})$ be the subset of permutations of length n that avoid both the patterns 312 and 231. Then $Av_n(\{312, 231\}) = \mathcal{L}_n$.*

Proof. We will prove this by double inclusion. First consider $\subseteq$. For any permutation in $\pi \in Av_n(\{312, 231\})$, we can partition π into consecutive decreasing subwords $\pi = \pi_1\pi_2 \dots \pi_k$. For instance, let $\pi = 2147653$. Then we can write $\pi = 21|4|765|3$. Now take the last element of some π_i, x, and the first element of π_{i+1}, y.. We argue that y must be larger than x. Assume

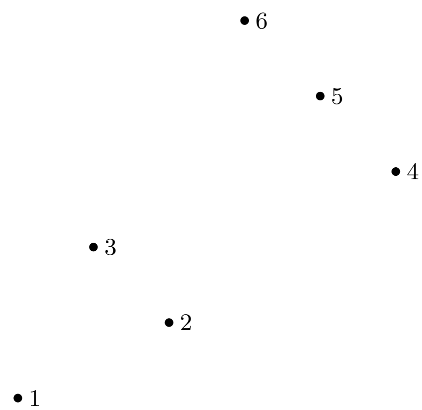


Figure 4. Graph of the layered permutation 132654.

y is smaller than x. Then there must be some z such that $y < x < z$, since otherwise π_i and π_{i+1} would be a single consecutively decreasing subword. If z is located before x, then zxy is a 231 pattern. If z is located after y, then xyz is a 312 pattern. Thus y is larger than x and π consists of collectively increasing, consecutively decreasing subsequences. Thus π is layered.

For the other inclusion $\supseteq$, suppose π is layered. Clearly π avoids 312 and 231 patterns. Thus we have equivalence. $\square$

The consecutive subwords π_i are called the layers of a layered permutation π. Viewing the permutations in $\mathcal{L}_n$ it is straightforward to enumerate $|\mathcal{L}_n|$.

Proposition 2.15. *Let $\mathcal{L}_n$ be the set of all layered permutations of length n. Furthermore, let $\mathcal{L}_n(k)$ be the subset of $\mathcal{L}_n$ consisting of the layered permutations with k layers, $1 \leq k \leq n$. Then*

$$|\mathcal{L}_n(k)| = \binom{n-1}{k-1}$$

It follows then that there are 2^{n-1} *layered permutations of length* n*.*

Proof. For any layered permutation $\pi \in \mathcal{L}_n$ with exactly k layers, consider the largest element n. Either n is in a layer by itself (the last layer) or it is part of a longer layer. Thus if we remove n we are left with one of the permutations in $\mathcal{L}_{n-1}(k)$ or $\mathcal{L}_{n-1}(k-1)$. Then denoting $\ell(n,k) = |\mathcal{L}_n(k)|$ we have the recursion

$$\ell(n,k) = \ell(n-1,k) + \ell(n-1,k-1)$$

It is easy to verify that

$$\ell(n,k) = \binom{n}{k}$$

is a general solution to this recursion. If we apply the boundary condition $\ell(n,1) = 1$, then we have

$$\ell(n,k) = \binom{n-1}{k-1}$$

Now summing over all k and applying the binomial theorem we find $|\mathcal{L}_n|$:

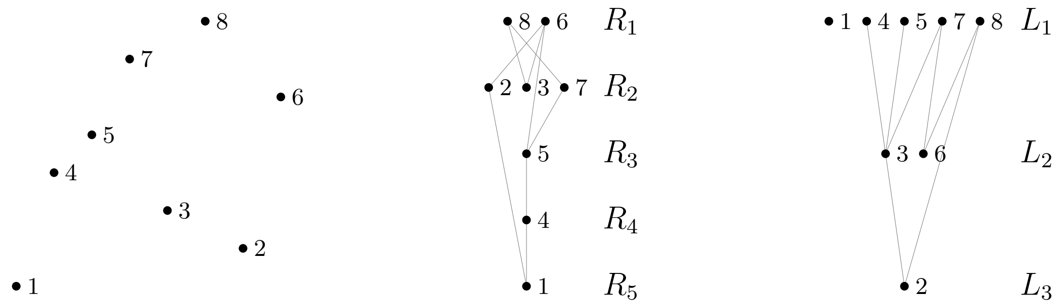
$$|\mathcal{L}_n| = \sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1}$$

□

Layered permutations are essential for the study of packing densities. In fact, there are no non-layered patterns for which the packing density is known (up to equivalence). That does not mean the packing density of every layered pattern is known, as calculating the packing density of layered patterns turns out to be a nontrivial optimization question. We first introduce a way to decompose permutations into partially ordered sets, used by Walter Stromquist [15] to prove an important theorem regarding packing densities of layered patterns.

Definition 2.16. Let $\pi = \pi(1)\ldots\pi(n)$ be a length-n permutation. Define a relation $\prec$ between two elements, defined $\pi(j) \prec \pi(k)$ if $j < k$ and $\pi(j) < \pi(k)$. An element $\pi(j)$ is a right-to-left maxima if there is no $\pi(k)$ such that $\pi(i) \prec \pi(k)$. The set of all right-to-left maxima will be denoted by R_1. Furthermore, define R_i to be the set of right-to-left maxima of the permutation $\pi \setminus \cup_{r=1}^{i-1} R_r$. We will similarly define the sets L_i to be the left-to-right maxima under $\succ$, defined $\pi(j) \succ \pi(k)$ if $j > k$ and $\pi(j) < \pi(k)$.

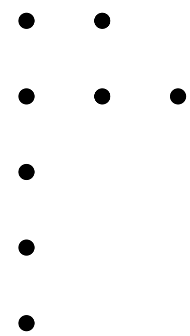
Note that the term "maxima" here is precise, the relation $\prec$ induces a partial ordering on the set of elements in the permutation. This relationship can be seen in Figure 5. Note also that the upper triangles of the adjacency matrices are complements of each other. This is clear since for any two elements $\pi(j)$ and $\pi(k)$, WLOG $\pi(j) < \pi(k)$. Thus either $j < k$ implying $\pi(j) \prec \pi(k)$, or $j > k$ implying $\pi(j) \succ \pi(k)$.



$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

FIGURE 5. Graph of the permutation 14573826 (top left); Graph of the right-to-left maxima (top middle); Graph of the left-to-right maxima (top right); adjacency matrix for $\prec$ (bottom left); adjacency matrix for $\succ$ (bottom right).

We then can map a permutation onto a certain kind of lattice in $\mathbb{Z}^+ \times \mathbb{Z}^+$, which we refer to as a maxima lattice. Here each element $\pi(k) \in L_i, R_j$ is mapped to (i, j). The lattice will be reversed for the second coordinate (as with partitions and matrices). Thus for our previous example of 14573826 we have the lattice



One may ask how many different maxima lattices are mapped to by the permutations in $[n!]$? This question seems to be extremely challenging, however a program is written to generate these lattices. We find the first few terms of the counting sequence to be $1, 2, 5, 15, 51, \ldots$.

One use of these maxima lattices is to provide an immediate proof of the Erdős-Szekeres theorem.

Theorem 2.17 (Erdős-Szekeres [4]). *Any permutation of length* $mn+1$ *contains either a* $1 \ldots (m+1)$ *pattern or an* $(n+1) \ldots 1$ *pattern.*

Proof. Consider any permutation of length $mn + 1$. Then we can map this permutation to a maxima lattice λ. It is straightforward to see that the number of rows in λ gives the length of the longest increasing subsequence, and the number of columns gives the length of the longest decreasing subsequence. By the pigeon hole principle, λ must have at least $n+1$ columns or $m+1$ rows. $\square$

Now we can define a "gap" in such a lattice as follows. Let $\lambda \subseteq \mathbb{Z}^+ \times \mathbb{Z}^+$ be a maxima lattice of a permutation π. A point (i, j) is a horizontal gap in λ

if $(i, j) \notin \lambda$ but $(i+1, j) \in \lambda$. Similarly (i, j) is a vertical gap if $(i, j) \notin \lambda$ but $(i, j+1) \in \lambda$.

Proposition 2.18. *There is a one-to-one correspondence between layered permutations and maxima lattices with no horizontal gaps.*

Proof. Start with any maxima lattice λ with rows $\lambda_1\lambda_2\ldots\lambda_k$, n nodes, and no horizontal gaps. Clearly the length-n layered permutation $\pi = \pi_1\pi_2\ldots\pi_k$ with $|\pi_i| = |\lambda_i|$ will map to this maxima lattice. For the inverse of this map, start with the maxima lattice $\lambda_1\lambda_2\ldots\lambda_k$. Label the nodes in the following manner. Let $(1,1) = n$, $(1,2) = n-1$, up to $(1, |\lambda_1|) = n - \lambda_1 + 1$. Continue labeling the rows consecutively in this way. Writing down the label of the nodes starting bottom to top from row λ_k, reading left to right yields the layered permutation $\pi = \pi_1\pi_2\ldots\pi_k$. $\square$

The equal enumeration of these two sets is clear when thought of in this way. The maxima lattices that avoid horizontal gaps are just compositions of n, counted by 2^{n-1}.

We see here that associated with every layered permutations is a maxima lattice. Now in general there are other permutations that are not layered that will map to this same maxima lattice. If we consider just the graph of the right-to-left maxima, we can think of layered permutations in a slightly different way. Let R_i be a set of right-to-left maxima, $i \geq 1$. If for any $y \in R_i$, $x \prec y$ for each $x \in R_j$, $j < i$, then R_i is layered.

Definition 2.19. If for a permutation π, R_1 is layered, then π is layered on top (LOT).

It follows immediately that if each R_i is layered, then π is layered.

2.3. **Packing Densities of Layered Patterns.** Let us now turn our attention to finding packing densities of layered patterns. It is a well known result that for any layered pattern τ, there exists a layered permutation that is τ-optimal of every length[3]. We state a slightly more general result here, proven by Albert, et al. [1].

Theorem 2.20. *Let $\Pi = \{\tau_1 \ldots \tau_m\}$ be a multiset of permutations that are all layered. A Π-optimal permutation $\hat{\pi}$ satisfies*

$$\sum_{i=1}^{m} f(\tau_i, \hat{\pi}) \geq \sum_{i=1}^{m} f(\tau_i, \pi)$$

for any other permutation $\pi \in [n!]$. Then among all Π-optimizers, there exists one that is layered.

Note that as a special case we have for any layered pattern τ

$$f_n(\tau) = f(\tau, \pi)$$

for some $\pi \in \mathcal{L}_n$. We refer the reader to the preceding reference for the proof.

[3]This result was first proven by Stromquist [15]

We next illustrate the idea of the proof on some specific patterns. First consider an increasing pattern $k^+ := 12\dots k$. It is trivial to show the packing density of k^+ is equal to one, however we generalize the analysis as follows.

Let $[n!; m, \infty]$ be the set of permutations that map to a maxima lattice with m rows; i.e. contains a longest increasing subsequence of length m. By searching in this restricted set, we can ask a slightly more interesting question. What is the maximum number of occurrences of k^+ in any one permutation in $[n!; m, \infty]$?

Proposition 2.21. *Let $\delta(\tau; m, \infty)$ be the packing density of τ when restricting the set of permutations to $[n!; m, \infty]$. Then*

$$\delta(k^+; m, \infty) = \frac{k!}{(m)^k}\binom{m}{k}$$

Proof. For any permutation π with right-to-left maxima sets $R_1, R_2, \dots, R_m$, consider any chain $x_1 \prec x_2 \prec \dots \prec x_s$ where $x_i \in R_{j_i}$ $(j_i < j_{i+1})$. Within this chain there are $\binom{s}{k}$ k^+ patterns. Now there are at most

$$\prod_{j=1}^{m} R_j$$

chains of length m. This quantity is obtainable by any layered permutation $\pi = \pi_1\pi_2\dots\pi_k$. Thus if we denote $|R_j| = r_j$, we have an optimization question

$$f_n(k^+; m, \infty) = \max_{r_1 + \ldots r_m = n} \left\{ \sum_{\substack{p \subseteq [m] \\ |p| = k}} \prod_{i=1}^{m} r_{p_i} \right\} \tag{2.5}$$

where $f_n(\tau; m, \infty) = f(\tau, \pi)$, π being a τ optimal permutation among permutations in $[n!; m, \infty]$.

We will show that 2.5 is achieved when $r_1 = r_2 = \ldots = r_m$. We proceed by induction on m and k, walking through the first several cases for illustrative purposes. Let $p(m, k)$ be the statement that 2.5 is achieved when all m variables are equal. Our base case is $p(m, 1)$, which is trivial. Next we will show that $p(m, m)$ is true. The expression to optimize is

$$x_1 x_2 \ldots x_m$$

subject to $x_1 + x_2 + \ldots + x_m = n$. By the arithmetic/geometric mean inequality

$$\sqrt[m]{x_1 x_2 \ldots x_m} \leq \frac{x_1 + x_2 + \ldots + x_m}{m}$$
$$x_1 x_2 \ldots x_m \leq \frac{n^m}{m^m}$$

Equality holds when $x_1 = x_2 = \ldots = x_m$, thus $p(m, m)$ is true. The first case not covered by these two is $p(3, 2)$. In this case we need to maximize the expression

$$xy + xz + yz$$

subject to $x + y + z = n$. Now if we factor out the x from the first two expressions we obtain

$$x(y + z) + yz$$

The term $y + z$ is equal to $n - x$, but more specifically it is the case $p(2, 1)$. The term yz is the case $p(2, 2)$. Thus we can say, by induction, that $y = z$ maximizes both of these expressions. Thus we have

$$x(2y) + y^2 = x(n - x) + \frac{(n - x)^2}{4}$$

All that is left to do is to find the optimal value of x. Taking the derivative and setting it equal to zero:

$$\begin{aligned} n - 2x + \frac{x - n}{2} &= 0 \\ 2n - 4x + x - n &= 0 \\ x &= \frac{n}{3} \end{aligned}$$

Thus $y = z = \frac{n-x}{2} = \frac{n}{3} = x$ and $p(3,2)$ is true. Now we will show the general case, $p(m,k)$, is true. We need only assume $1 < k < m$. The expression to maximize is

$$x_1 \ldots x_k + \ldots + x_1 \ldots x_m + \ldots + x_{m-k+1} \ldots x_m$$

subject to $x_1 + x_2 + \ldots + x_m = n$. Exploiting the same trick by factoring out x_1 we obtain

$$x_1(x_2 \ldots x_k + \ldots + x_{m-k+2} \ldots x_m) + \ldots + x_{m-k+1} \ldots x_m$$

By $p(m-1, k-1)$, the sum in parenthesis is maximized when all $m-1$ variables are equal. By $p(m-1, k)$ the remaining sum of terms is also maximized when the same $m-1$ variables are equal. Thus for any fixed value of x_1, the whole expression reaches a maximum when $x_2 = x_3 = \ldots = x_m$. In this case we have

$$x_1 \binom{m-1}{k-1} \frac{(n-x_1)^{k-1}}{(m-1)^{k-1}} + \binom{m-1}{k} \frac{(n-x_1)^k}{(m-1)^k}$$

Now we solve for the optimal size of x_1 by taking the derivative and setting it equal to zero. For simplicity let $x = x_1$.

$$\frac{d}{dx}\left[\frac{(m-1)!x(n-x)^{k-1}}{(k-1)!(m-k)!(m-1)^{k-1}}+\frac{(m-1)!(n-x)^{k}}{k!(m-k-1)!(m-1)^{k}}\right]=0$$

$$\frac{(m-1)![(n-x)^{k-1}-x(k-1)(n-x)^{k-2}]}{(k-1)!(m-k)!(m-1)^{k-1}}-\frac{(m-1)!k(n-x)^{k-1}}{k!(m-k-1)!(m-1)^{k}}=0$$

$$(m-1)[(n-x)-x(k-1)]-(m-k)(n-x)=0$$

$$(m-1)(n-kx)-(m-k)(n-x)=0$$

$$(mn-mkx-n+kx)-(mn-mx-kn+kx)=0$$

$$mx(1-k)+n(k-1)=0$$

$$mx=n$$

$$x=\frac{n}{m}$$

Now we can solve for the rest of the variables. For $i \geq 2$,

$$x_i=\frac{n-x}{m-1}=\frac{nm-n}{m(m-1)}=\frac{n}{m}=x$$

Thus we have $x_1 = x_2 = \ldots = x_m$. Thus 2.5 becomes

$$f_n(k^+;m,\infty)=\binom{m}{k}\left(\frac{n}{m}\right)^k$$

The packing density is then given by

$$\begin{aligned}\delta(k^+;m,\infty) &= \lim_{n\to\infty} \frac{\binom{m}{k}\left(\frac{n}{m}\right)^k}{\binom{n}{k}} \\ &= \frac{k!}{m^k}\binom{m}{k}\end{aligned}$$

$\square$

We see immediately from this analysis that as $m \to \infty$, $\delta(k^+;m,\infty) \to \delta(k^+) = 1$. It is natural to assume this is true for any pattern, an idea we will explore in the next section.

In general we can restrict ourselves to $[n!;m,\ell]$, the set of permutations of length n with a maxima lattice containing m rows and ℓ columns. In general it is very difficult to find the dependence of an optimal permutation on m and ℓ. In fact, for m and ℓ both fixed we inherently bound n. More specifically,

$$m + \ell - 1 \le n \le m\ell$$

Equality holds for more than one permutation. For the lower bound, we have 132 and 231 with $m = \ell = 2$ and $n = m + \ell - 1 = 3$. As for the upper bound, consider 3142 and 2413. We have $m = \ell = 2$ and $n = m\ell = 4$.

We use these facts to compute the packing densities of a couple nontrivial layered patterns. For patterns of length-3 there is only one nontrivial pattern (up to equivalence), the pattern 132. The packing density for 132 was first calculated by Galvin under the assumption of Theorem 2.20. A specific case of Theorem 2.20 was then proven independently by Kleitman and Stromquist. This account was recorded by Price [10].

We will prove a Lemma next that allows us to calculate the packing density of 132. Theorem 2.20 also follows from this Lemma.

Lemma 2.22. *Let τ be a layered pattern of length $k \leq n$. Let $[n!; m, \infty, LOT]$ be the set of all permutations of length n that are layered on top (LOT). Then among τ-optimizers there exists one that is in $[n!; m, \infty, LOT]$.*

Proof. We will prove by contradiction. Assume then that there is no $\pi \in [n!; m, \infty, LOT]$ such that $f(\tau, \pi) = f_n(\tau; m, \infty)$. We then select another τ-optimizer $\sigma \in [n!; m, \infty]$. By our assumption $\sigma \notin [n!; m, \infty, LOT]$. Thus if we consider the right-to-left maxima R_1 of σ we can find two elements $x, y \in R_1$ satisfying the following. First x and y are consecutive elements in R_1. That is, if $x = \sigma(j)$ and $y = \sigma(k)$ with $j < k$, there does not exist a $z = \sigma(i) \in R_1$ such that $j < i < k$. Next define the following sets:

$$A = \{\sigma(a) : a < j \ \wedge \ \sigma(k) < \sigma(a) < \sigma(j)\}$$

$$B = \{\sigma(b) : j < b < k \ \wedge \ \sigma(b) < \sigma(k)\}$$

Figure 6 gives an illustration of these two sets. We can assume finally that either $A \neq \varnothing$ or $B \neq \varnothing$.

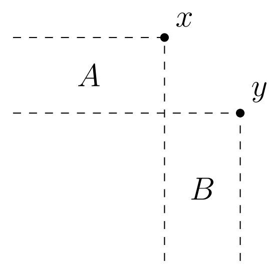


FIGURE 6.

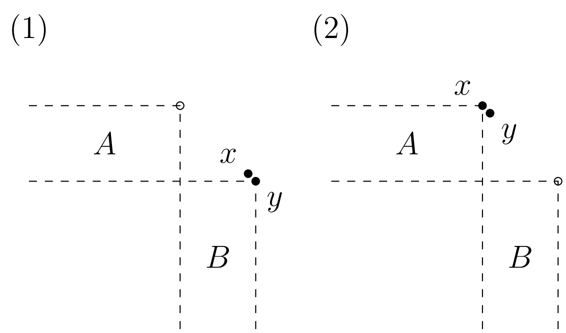


FIGURE 7. New permutations: (1) formed by moving x next to y; (2) formed by moving y next to x.

Now we will perform two different operations on σ, showing that one of them will not decrease the number of occurrences of τ.

(1) In the graph of σ, x is at location $(j, \sigma(j))$. Create a new permutation by moving x to location $(k-1, \sigma(k)+1)$, shifting the other elements in σ accordingly.

(2) In the graph of σ, y is at location $(k, \sigma(k)$. Create a new premutation by moving y to location $(j+1, \sigma(j)-1)$, shifting the other elements in σ accordingly.

These two new words are shown in Figure 7.

We can now count the change in τ patterns for these two words. In σ, let $f(\tau,\sigma;x,\neg y)$ denote the τ patterns in σ formed using x and not y, and let $f(\tau,\sigma;y,\neg x)$ denote the τ patterns in σ formed using y and not x. Then the change in τ patterns from σ to (1) is given by

$$f(\tau,(1)) - f(\tau,\sigma) = f(\tau,\sigma;y,\neg x) - f(\tau,\sigma;x,\neg y)$$

Similarly the change in τ patterns from σ to (2) is given by

$$f(\tau,(2)) - f(\tau,\sigma) = f(\tau,\sigma;x,\neg y) - f(\tau,\sigma;y,\neg x)$$

Conveniently one of these two terms must be nonnegative. Thus either (1) or (2) has at least as many τ patterns as σ. This is a contradiction since σ not being LOT requires such a gap between two elements in R_1. Thus there must exist a τ optimal permutation that is layered on top. $\square$

This proof never required there being only one layered pattern τ being packed. If instead we had considered a set of layered patterns $\{\tau_i\}_{1\leq i\leq s}$, it follows that there is still a LOT permutation π that maximizes the sum

$$\sum_{i=1}^{s} f(\tau_i,\pi)$$

2.4. **Bounded and Unbounded Permutations.** We saw for the simple pattern k^+ that as the number right-to-left maxima sets (for layered patterns corresponding to the number of layers) increased, the packing density increased as well. This is not true in general. It may be the case that there is an m^* such that

$$\delta(\tau; m^*, \infty) \geq \delta(\tau; m, \infty)$$

for any $m > m^*$. For a layered pattern, this means there is an optimal number of layers for which any more layers will not increase the packing density.

Definition 2.23. Given a pattern τ with restricted packing density $\delta(\tau; m, \infty)$. Let $m(\tau)$ be the value of m such that

$$\delta(\tau; m, \infty)$$

is a maximum. If $m(\tau) < \infty$ then τ is said to be bounded or of the bounded type. If $m(\tau) = \infty$ then τ is said to be unbounded or of the unbounded type.

The first result regarding bounded/unbounded permutations is again due to Price [10].

Proposition 2.24. *Let $\tau = 1 \oplus \bar{k}$ where again $\bar{k} = k(k-1)\dots 1$. Then τ is of the unbounded type.*

Proof. Assume that there exists an $m < \infty$ such that $\delta(\tau; m, \infty)$ is maximized. Thus there exists an N such that $f(\tau, \pi; m, \infty) = f_N(\tau)$ for some permutation $\pi \in [N!; m, \infty]$. Since τ is layered we can assume π is also layered by Theorem 2.20. If we consider the last right-to-left maxima set R_m, the size of this set is some fraction of N. Thus we can write $|R_m| = aN$ for some $a \in (0,1)$. Now for every $n > N$, there is some optimal layered permutation in $[n!; m, \infty]$ with $|R_m| = an$. If we increase the size of our permutation to be of size $n = \frac{N+k+1}{a}$, then the our set R_m contains at least $k+2$ elements. The number of τ patterns formed using elements in R_m is then

$$|R_m| \sum_{i=1}^{m-1} \binom{|R_i|}{k}$$

Now because this permutation is layered, R_m is just a layer containing $\pi(|R_m|) = 1$. If we move this element to the beginning of the permutation, we form a permutation with $m+1$ layers containing

$$|R_m| \sum_{i=1}^{m-1} \binom{|R_i|}{k} + 1$$

τ patterns. This contradicts m maximizing $\delta(\tau; m, \infty)$. Thus τ is unbounded.

□

We will now calculate the packing density of the pattern 132.

Proposition 2.25. *For the pattern* 132*, we have*

$$\delta(132) = 2\sqrt{3} - 3$$

Proof. Since 132 is a layered pattern, we can assume there is a 132-optimal permutation $\pi \in [n!; m, \infty, LOT]$ for any m. If we denote the top layer of π by π_1 and $|\pi_1| = x$, then we can write

$$\pi = \pi^* \pi_1$$

where π^* is a 132-optimal permutation of length $n - x$ in $[(n-x)!; m-1, \infty]$. For instance,

$$f_n(132; 2, \infty) = \max_{x \in (0,n)} \left\{ (n-x) \binom{x}{2} \right\}$$

It is straightforward to find the optimal x:

$$\begin{aligned}
\frac{d}{dx}\left[(n-x)\frac{x^2-x}{2}\right] &= 0 \\
\frac{1}{2}\big((x-x^2)+(n-x)(2x-1)\big) &= 0 \\
x - x^2 + 2nx - n - 2x^2 + x &= 0 \\
3x^2 - 2(1+n)x + n &= 0
\end{aligned}$$

$$\begin{aligned}
x &= \frac{2(1+n) \pm \sqrt{4(1+n)^2 - 12n}}{6} \\
&= \frac{2(1+n) \pm \sqrt{4(1-n+n^2)}}{6} \\
&= \frac{1+n \pm \sqrt{1-n-n^2}}{3}
\end{aligned}$$

Now as $n \to \infty$ we see that $x \sim \frac{2n}{3}$. Thus

$$\delta(132; 2, \infty) = \lim_{n\to\infty} \frac{\frac{n(4n^2-6n)}{54}}{\binom{n}{3}} = \frac{4}{9}$$

For $m \geq 2$ we can write

$$f_n(132; m, \infty) = \max_{x\in(0,n)} \left\{ f_{n-x}(132; m-1, \infty) + (n-x)\binom{x}{2} \right\}$$

Recall

$$f_{n-x}(132; m-1, \infty) = \delta_{n-x}(132; m-1, \infty)\frac{(n-x)^3}{6}$$

Then we solve for x in the expression

$$\frac{d}{dx}\left[\delta_{n-x}(132; m-1, \infty)\frac{(n-x)^3}{6} + (n-x)\binom{x}{2}\right] = 0$$

We then obtain

$$x = \frac{-\sqrt{-n^2 + \delta_{n-x}[m-1]n + n^2 - n + 1} + \delta_{n-x}[m-1]n + n + 1}{\delta_{n-x}[m-1] + 3}$$

For simplicity let $\delta_{n-x}(132; m-1, \infty) = \delta_{n-x}[m]$. Here we must make an important note. Though $\delta_{n-x}[m]$ implicitly depends on x, we use the fact that

$\delta_n[m] \to \delta[m]$ as $n \to \infty$. Thus we must argue that $(n-x) \to \infty$ as $n \to \infty$ for the optimal value of x (which inherently depends on n). Let $x[m]$ be the optimal x for any given m as $n \to \infty$. We can show that $x[\infty] = an$ for some $a > 0$. Consider

$$f_n(132) = \max_{x \in (0,n)} \left\{ f_{n-x}(132) + (n-x)\binom{x}{2} \right\}$$

This recursion can be solved explicitly, giving

$$f_n(132) = \max_{a \in (0,1)} \left\{ \frac{a(1-n)\left(a^2n^2 + 2a^2n + 2a^2 + an + a - n^2\right)}{2(a+1)\left(a^2+a+1\right)} \right\}$$

Solving for a we obtain

$$a \sim \frac{1}{2}(3-\sqrt{3})$$

Therefore $x[\infty] = \frac{1}{2}(3-\sqrt{3})n$ as claimed. It is now easy to see that each $x[m] > x[\infty]$, so that

$$n \geq n - x[m] \geq n - x[\infty] = \frac{1}{2}(\sqrt{3}-1)n$$

and thus $n - x[m] \to \infty$ for every m.

Substituting the expression for the optimal x into the expression for $f_n(132; m, \infty)$, dividing by $\binom{n}{3}$, and letting $n \to \infty$ (using the fact that $\delta_n(\tau) \to \delta(\tau)$ as $n \to \infty$) we have

$$(2.6) \qquad \delta[m] = \frac{2\left(2 - \sqrt{1 - \delta[m-1]}\right)\left(\delta[m-1] + \sqrt{1 - \delta[m-1]} + 1\right)}{(\delta[m-1] + 3)^2}$$

where again $\delta[m] = \delta(132; m, \infty)$. Now because 132 is of the unbounded type, it follows that

$$\delta[m+1] \geq \delta[m]$$

Since this sequence is also bounded above (by 1), there is a unique fixed point of this sequence which is easily found:

$$\delta = \frac{2\left(2 - \sqrt{1 - \delta}\right)\left(\delta + \sqrt{1 - \delta} + 1\right)}{(\delta + 3)^2}$$

and we find that $\delta = \delta(132) = 2\sqrt{3} - 3$. $\square$

Note that using (2.6) we have found $\delta(132; m, \infty)$ for every m recursively. Setting $m = 1$ we have the base case $\delta(132; 1, \infty) = 0$. For $m = 2$ we have $\delta(132; 2, \infty) = \frac{4}{9}$ as we have seen.

Analysis of this type works for any pattern of the form $1 \oplus \bar{k}$. Unfortunately for other patterns with other structure we must use other methods. In general this can be almost impossible, particularly if the pattern is of the bounded type. In fact, in many cases $m(\tau)$ may not even be known.

Let us look at a simple pattern of the bounded type. This result was first shown by [1].

Proposition 2.26. *The length-4 pattern $\tau = 2143$ is of the bounded type. Furthermore, $m(\tau) = 2$ and*

$$\delta(2143) = \frac{3}{8}$$

Proof. We will first show that $m(\tau) = 2$. It then follows immediately the 2143 is of the bounded type, and that $\delta(2143) = \frac{3}{8}$. It is trivial to see $m(\tau) \geq 2$ since otherwise there would be no 2143 patterns. Also any τ-optimal permutation may be assumed to be layered. Now suppose there is a permutation $\pi \in [n!; m, \infty]$ for some $m > 2$ such that $f(\tau, \pi) > f_n(2143; 2, \infty)$. Now if we enumerate the number of 2143 patterns in π with layeres $\pi_1, \pi_2, \ldots, \pi_r$, we have

$$f(\tau, \pi) = \binom{|\pi_1|}{2}\binom{|\pi_2|}{2} + \binom{|\pi_1|}{2}\binom{|\pi_3|}{2} + \ldots + \binom{|\pi_{s-1}|}{2}\binom{|\pi_s|}{2}$$

We will first factor out $\binom{|\pi_1|}{2}$ from each term that contains one,

$$f(\tau, \pi) = \binom{|\pi_1|}{2}\left(\binom{|\pi_2|}{2} + \ldots + \binom{|\pi_s|}{2}\right) + \ldots + \binom{|\pi_{s-1}|}{2}\binom{|\pi_s|}{2}$$

The term in parenthesis is a convex function. We have

$$\frac{\sum_{i=1}^{s} |\pi_i|^2 - n}{2}$$

with Hessian $\delta_{i,j}$, the Kronecker delta defined as

$$\delta_{i,j} = \begin{cases} 1 & i = j \\ \\ 0 & i \neq j \end{cases}$$

so all the eigenvalues of the Hessian are 1. This implies the Hessian is positive definite, i.e. the function is convex. Because it is convex the maximum value will occur on the boundary. Now consider the case when $m = 3$. We have

$$f(\tau, \pi) = \binom{|\pi_1|}{2}\left(\binom{|\pi_2|}{2} + \binom{|\pi_3|}{2}\right) + \binom{|\pi_2|}{2}\binom{|\pi_3|}{2}$$

For simplicity, let $|\pi_i| = x_i$. Then using Lagrange multipliers we have

$$\left(x_1 - \frac{1}{2}\right)\left(\binom{x_2}{2} + \binom{x_3}{2}\right) = 1$$
$$\left(x_2 - \frac{1}{2}\right)\left(\binom{x_1}{2} + \binom{x_3}{2}\right) = 1$$
$$\left(x_3 - \frac{1}{2}\right)\left(\binom{x_1}{2} + \binom{x_2}{2}\right) = 1$$

With some algebra we see there is a unique local maximum on the interior at $x_1 = x_2 = x_3 = \frac{n}{3}$. Looking on the boundary (WLOG $(x_1, x_2, 0)$, since our function is symmetric) we see the maximum is located at $x_1 = x_2 = \frac{n}{2}$. We can find the values of n for which one of these terms is larger:

$$3\binom{\frac{n}{3}}{2}^2 \leq \binom{\frac{n}{2}}{2}^2$$

for $n \geq \frac{6}{11}\left(1 + 2\sqrt{3}\right) \approx 2.43$. Since we are only interested in permutations of length ≥ 4, the point on the boundary is the absolute maximum.

Having a base case, we can show by induction that

$$\binom{\frac{n}{2}}{2}^2 \geq f_n(2143; m, \infty)$$

for any $m > 2$. Thus $m(\tau) = 2$. $\square$

We conclude by giving a brief table of packing densities of well known patterns.

τ	$\delta(\tau)$	$m(\tau)$	Reference
$1\dots k$	1.000	∞	—
$k\dots 1$	1.000	∞	—
132	0.464	∞	[15]
1432	0.424	∞	[10]
2143	0.375	2	[10]
1243	0.375	∞	[1]
1324	≈ 0.244	∞	[10]
1342	≈ 0.197	∞	Conjecture
2413	≈ 0.105	∞	Conjecture

FIGURE 8.

3. Colored Permutations

We now turn our attention to permutations that are colored. That is, start with a permutation $\pi \in [n!]$. Let $C \neq \varnothing$ be a set of distinct yet unrelated elements, called colors. A colored permutation is any element $\chi \in [n!] \times (C^n)$. The sequence $c = c(1)c(2)\ldots c(n) \in C^n$ of a colored permutation $\chi = (\pi, c)$ is called the coloring of π. We give a simple definition here:

Definition 3.1. An *m-colored permutation* χ of *length* n is a permutation of length n in which each element is assigned one of m distinct colors.

For example, let $\chi = 3_a 2_a 5_b 1_a 4_b$ be a two-colored permutation where 3,2 and 1 have color a while 5 and 4 have color b. Analogously to the case of non-colored patterns the colored pattern $\phi = 2_a 1_a 3_b$ occurs in χ as the subsequences $3_a 1_a 4_b$ and $3_a 2_a 4_b$.

Colored permutations are similar to permutations of a multi-set. The question of pattern avoidance on multi-sets has been studied in past years (see, for instance, [11]). In this section we focus on pattern packing in colored permutations.

We first define *colored blocks*, which are central to our study. Colored blocks are analogous to layers in non-colored permutations.

Definition 3.2. In a colored permutation χ, a *colored block* is a maximal monochromatic segment $\chi_i^{(a)}$ in which each element has color a and every element not in $\chi_i^{(a)}$ is either larger than or smaller than all elements in $\chi_i^{(a)}$.

Remark 3.3. Note that every entry in a colored permutation is in exactly one of its colored blocks.

In other words, a colored block is a monochromatic segment of elements with consecutive numerical values. For instance, the permutation $\chi = 3_a 1_a 2_a 6_a 5_b 4_b$ contains three colored blocks, $\chi_1^{(a)} = 3_a 1_a 2_a$, $\chi_2^{(a)} = 6_a$, and $\chi_3^{(b)} = 5_b 4_b$. A graphical representation of colored blocks is shown in Figure 9. An important note is that colored blocks are both numerically and chromatically disjoint.

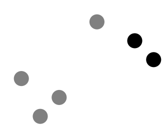


FIGURE 9. A Colored Permutation $3_a 1_a 2_a 6_a 5_b 4_b$

In what follows we provide some observations on the optimal colored permutations for colored patterns which contain either two or three colored blocks. For convenience we will reuse the notation $f(\phi, \chi)$ to represent the number of occurrences of the colored pattern ϕ in the colored permutation χ. Colored blocks will often be denoted simply by their color and/or location, i.e. $\chi_1^{(a)} = A_1$, $\chi_2^{(b)} = B_2$, $\chi_3^{(a)} = A_3$, etc. Similarly for colored patterns $\phi_1^{(a)} = \alpha_1$, $\phi_2^{(b)} = \beta_2$, etc. The collection of all colored blocks in χ of color a (b) will be denoted by χ_A (χ_B).

3.1. **Two Colored Blocks.** Note that a single-colored (or non-colored) permutation has exactly one colored block (namely the permutation itself). We then assume the permutations/patterns under consideration to be at least two-colored.

Since the number of colored blocks is at least the number of colors, we may assume that a colored pattern ϕ with two colored blocks $\phi_1^{(r)} = \rho$ and $\phi_2^{(b)} = \beta$ to have exactly two colors and is of the form $\rho\beta$. Furthermore, we may assume without loss of generality that all elements in ρ are less than all elements of β (similarly to layers denoted $\alpha < \beta$). For the remainder of this section color r will be referred to as "red" and b will be referred to as "blue".

Theorem 3.4. *For a pattern ϕ with two blocks of the form $\rho\beta$ with $\rho < \beta$, there is an optimal length-n permutation $\hat{\chi}$ of the form RB with $R < B$.*

Remark 3.5. The proof below follows the simple idea that sliding all of the red entries to the left and all of the blue entries to the right leaves every instance of a ϕ-pattern intact.

Proof. Let χ be an optimal permutation of length n with colored blocks $\chi_1\chi_2\dots\chi_k$. First we claim that $\chi_1 = R_1$ is red. If χ_1 were blue,

$$f(\phi, \chi) = f(\rho, \chi_1) \cdot f(\beta, \chi_{>1},) + f(\phi, \chi_{>1})$$

where $\chi_{>i_0}$ ($\chi_{<i_0}$) is the collection of all colored blocks in χ after (before) χ_{i_0}. Clearly the first term in the sum is zero. By recoloring χ_1 red, this term is replaced with

$$f(\rho, R_1) \cdot f(\beta, \chi_{>1}) \geq 0,$$

and thus $f(\phi, \chi)$ will only increase. Similarly, we may assume $\chi_k = B_k$ is blue. Along the same lines we claim there is no blue block immediately preceding a red block. Otherwise, let $\chi_j = B_j$ and $\chi_{j+1} = R_{j+1}$ in χ, we have

$$f(\phi, \chi) = f(\rho, \chi_{<j}) \cdot f(\beta, B_j) + f(\beta, \chi_{>j+1}) \cdot f(\rho, R_{j+1}) + f(\phi, \chi_{<j}\chi_{>j+1}).$$

Let χ' be obtained from χ by switching χ_j and χ_{j+1}, we have

$$f(\phi, \chi') = f(\phi, \chi) + f(\rho, R_{j+1}) \cdot f(\beta, B_j).$$

Thus $f(\phi, \chi)$ may only increase.

Consequently, we now have an optimal permutation χ of the form $\chi = \chi_R \chi_B$. Because any $\phi = \rho\beta$ pattern occurring in χ must consist of a ρ pattern from χ_R and a β pattern from χ_B, we have

$$f(\phi, \chi) \leq f(\rho, \chi_R) \cdot f(\beta, \chi_B)$$

with equality if $\chi_R < \chi_B$. Hence there is an optimal permutation $\hat{\chi} = RB$ with $R < B$. $\square$

For example, to pack the pattern $\chi = 2_r 1_r 3_b 4_b$ an optimal permutation of length n consists of a decreasing sequence of the elements $\lfloor \frac{n}{2} \rfloor \ldots 1$ colored red followed by an increasing sequence of the elements $(\lfloor \frac{n}{2} \rfloor + 1) \ldots n$ colored blue. More detailed applications will be discussed in Section 3.3.

3.2. **Three Colored Blocks.** We next consider patterns with three colored blocks through several different cases. Some arguments are similar to those in the previous subsection and we omit some details.

First consider the case when the pattern has three distinct colors. Assume without loss of generality $\phi_1^{(r)} = \rho$, $\phi_2^{(b)} = \beta$, and $\phi_3^{(g)} = \gamma$ (with colors red, blue and green) and thus the colored pattern has the form $\phi = \rho\beta\gamma$.

Theorem 3.6. *Given a pattern ϕ with three colored blocks of distinct colors of the form $\rho\beta\gamma$, there is an optimal permutation $\hat{\chi}$ (of length n) of the form RBG with the same numerical ordering as $\rho\beta\gamma$.*

Remark 3.7. For instance, given a pattern ϕ of the form $\rho\beta\gamma$ with $\rho < \gamma < \beta$, there is an optimal permutation of the form RBG such that $R < G < B$.

Proof. Following the same arguments as Theorem 3.4, it is easy to show that the optimal permutation is of the form

$$R_1 \dots R_i B_{i+1} \dots B_j G_{j+1} \dots G_k = \chi_R \chi_B \chi_G.$$

Then

$$f(\phi, \chi) \leq f(\rho, \chi_R) \cdot f(\beta, \chi_B) \cdot f(\gamma, \chi_G)$$

with equality if any elements $a \in \chi_R$, $b \in \chi_B$ and $c \in \chi_G$ assume the same numerical ordering as ρ, β and γ. □

Suppose now that a pattern ϕ has three colored blocks with only two colors. Assume without loss of generality that there is one red block and two blue blocks. First consider the case when the two blue blocks are adjacent, i.e., $\phi = \rho\beta_1\beta_2$. This case is representative of all patterns with two adjacent blue blocks since reversing the permutation/pattern turns all $\rho\beta_1\beta_2$ patterns into $\beta_1\beta_2\rho$ patterns.

Theorem 3.8. *For a pattern ϕ with three colored blocks of the form $\rho\beta_1\beta_2$, there is an optimal permutation $\hat{\chi}$ (of length n) that is also of the form RB_1B_2*

and the numerical ordering of the colored blocks in $\hat{\chi}$ is the same as that of the colored blocks in ϕ.

Proof. First note that since the two blue blocks are numerically disjoint, no element from β_1 may be (numerically) adjacent to an element in β_2. That is, either $\beta_1 < \rho < \beta_2$ or $\beta_2 < \rho < \beta_1$. Without loss of generality we will assume the former.

Once again arguments from Theorem 3.4 yield that all red blocks in an optimal permutation χ can be placed before any blue blocks. That is, an optimal permutation is of the form

$$\chi = R_1 \dots R_i B_{i+1} \dots B_k = \chi_R \chi_B.$$

Note that any $\rho\beta_1\beta_2$ pattern is a result of a ρ pattern in χ_R and a $\beta_1\beta_2$ pattern in χ_B. For any particular pattern ρ in χ_R, let $\chi_{B_{<\rho}}$ be the set of all blue blocks less than ρ and $\chi_{B_{>\rho}}$ be the set of all blue blocks greater than ρ. Since $\rho\beta_1\beta_2$ patterns are only formed using β_1 patterns from $\chi_{B_{<\rho}}$ and β_2 patterns from $\chi_{B_{>\rho}}$, the contribution from this ρ pattern to $f(\phi, \chi)$ is at most

$$f(\beta_1, \chi_{B_{<\rho}}) \cdot f(\beta_2, \chi_{B_{>\rho}}).$$

This can be achieved (regardless of the choice of ρ) by putting the blue blocks in increasing order. Under this assumption, let $\chi_{B_{<j}}$ ($\chi_{B_{>j}}$) denote the collection of blue blocks before (after) B_j in χ_B and j_0 be such that

$$f(\beta_1, \chi_{B_{<j_0+1}}) \cdot f(\beta_2, \chi_{B_{>j_0}}) \geq f(\beta_1, \chi_{B_{<j+1}}) \cdot f(\beta_2, \chi_{B_{>j}})$$

for any $i+1 \leq j \leq k-1$, we now have

$$f(\phi, \chi) \leq f(\rho, \chi_R) \cdot f(\beta_1, \chi_{B_{<j_0+1}}) \cdot f(\beta_2, \chi_{B_{>j_0}}).$$

Equality holds if

$$\chi_{B_{<j_0+1}} < \chi_R < \chi_{B_{>j_0}}.$$

Consequently each of $R = \chi_R$, $B_1 = \chi_{B_{<j_0+1}}$ and $B_2 = \chi_{B_{>j_0}}$ is a single block and the optimal permutation is of the form RB_1B_2 with $B_1 < R < B_2$. $\square$

Lastly we consider the case when the pattern is of the form $\phi = \beta_1\rho\beta_2$.

Theorem 3.9. *For a pattern ϕ with three colored blocks of the form $\beta_1\rho\beta_2$, there is an optimal permutation $\hat{\chi}$ that is of the form B_1RB_2 with same numerical ordering as those in ϕ.*

Proof. First we may assume (following the same argument as before), that in an optimal permutation χ, the first and last blocks are blue, i.e.,

$$\chi = \chi_1^{(b)} \dots \chi_i^{(c_i)} \dots \chi_k^{(b)}$$

where $c_i \in \{r, b\}$. Consider any ρ pattern formed in the sequence of the s red blocks $\chi_R := \chi_{j_1}^{(r)} \chi_{j_2}^{(r)} \dots \chi_{j_s}^{(r)}$. A $\beta_1\rho\beta_2$ pattern can only be formed by a β_1 pattern in the the sequence of blue blocks before $\chi_{j_1}^{(r)}$ and a β_2 pattern in the sequence of blue blocks after $\chi_{j_s}^{(r)}$. Thus the number of ϕ patterns formed from this particular ρ pattern is at most

$$f(\beta_1, \chi_{B_{<j_1}}) \cdot f(\beta_2, \chi_{B_{>j_s}})$$

where $\chi_{B_{<j_1}}$ denotes the sequence of blue blocks before block j_1 and $\chi_{B_{>j_s}}$ denotes the sequence of blue blocks after block j_s.

Let j_0 (not necessarily unique) be a value such that $c_{j_0} = r$ and

$$f(\beta_1, \chi_{B_{<j_0}}) \cdot f(\beta_2, \chi_{B_{>j_0}}) \geq f(\beta_1, \chi_{B_{<j}}) \cdot f(\beta_2, \chi_{B_{>j}})$$

for all j, then

$$f(\beta_1\rho\beta_2, \chi) \leq f(\beta_1, \chi_{B_{<j_0}}) \cdot f(\rho, \chi_R) \cdot f(\beta_2, \chi_{B_{>j_0}})$$

with equality if all red blocks are located in between the first blue block immediately preceding and following $\chi_{j_0}^{(r)}$. Consequently χ is an optimal permutation of the form

$$\chi_1^{(b)} \cdots \chi_\ell^{(b)} \chi_{\ell+1}^{(r)} \cdots \chi_m^{(r)} \chi_{m+1}^{(b)} \cdots \chi_k^{(b)} = \chi_{B_1}\chi_R\chi_{B_2}$$

with colored blocks $\chi_{B_1} := \chi_1^{(b)} \cdots \chi_\ell^{(b)}$, $\chi_R := \chi_{\ell+1}^{(r)} \cdots \chi_m^{(r)}$, and $\chi_{B_2} := \chi_{m+1}^{(b)} \cdots \chi_k^{(b)}$.

Through arguments similar to those of Theorem 3.8 one can see that the numerical ordering of $\chi_{B_1}\chi_R\chi_{B_2}$ is the same as $\beta_1\rho\beta_2$. $\square$

3.3. **Applications to Specific Patterns.** In this section, we apply our findings to some specific colored patterns and obtain their corresponding packing densities. Recall $f_n(\phi)$ denotes the specific number of occurrences of ϕ in an optimal permutation of length n.

3.3.1. *Patterns of Length 2.* For non-colored patterns of length 2, the packing density is trivially equal to one.

In the colored case, Theorem 3.4 implies that the optimal permutation of length n of the colored pattern 1_r2_b (or equivalently 2_r1_b) is of the form RB with $R < B$. Then

$$f_n(1_r2_b) = f(1,R) \cdot f(1,B) = |R| \cdot |B|.$$

Given $|R| + |B| = n$, it is easy to see that

$$f_n(1_r2_b) = \left\lfloor \frac{n^2}{4} \right\rfloor = \frac{2n^2 - 1 + (-1)^n}{8}.$$

Therefore the packing density of all length two colored patterns (in which two distinct colors occur) is given by

$$\delta(1_r2_b) = \lim_{n\to\infty} \frac{f_n(1_r2_b)}{\binom{n}{2}} = \frac{1}{2}.$$

For non-colored patterns of length 3, the packing densities for the decreasing and increasing patterns are trivial. The layered pattern 132 has packing density $2\sqrt{3} - 3$ as established in [15].

3.3.2. *The pattern $2_r1_b3_b$ (and equivalents).* Theorem 3.8 implies that the optimal permutation of length n of the colored pattern $2_r1_b3_b$ is of the form

RB_1B_2 with $B_1 < R < B_2$, then

$$f_n(2_r1_b3_b) = f(1,R)\cdot f(1,B_1)\cdot f(1,B_2) = |R|\cdot|B_1|\cdot|B_2|.$$

With $|R|+|B_1|+|B_1| = n$, it is easily shown that

$$f_n(2_r1_b3_b) = \begin{cases} \frac{n^3}{27}, & \text{if } n\equiv 0 \mod 3\\ \frac{(n-1)^2(n+2)}{27}, & \text{if } n\equiv 1 \mod 3\\ \frac{(n+1)^2(n-2)}{27}, & \text{if } n\equiv 2 \mod 3\end{cases}$$

Consequently the packing density of $2_r1_b3_b$ (and equivalent patterns) is

$$\delta(2_r1_b3_b) = \lim_{n\to\infty}\frac{f_n(2_r1_b3_b)}{\binom{n}{3}} = \frac{2}{9}.$$

3.3.3. *The pattern $1_r3_b2_b$ (and equivalents).* Theorem 3.4 implies that the optimal permutation of length n of the colored pattern $1_r3_b2_b$ is of the form RB in which $R < B$, then

$$f_n(1_r3_b2_b) = f(1,R)\cdot f(21,B) = |R|\binom{|B|}{2}.$$

Let $|B| = k$, then

$$f_n(1_r3_b2_b) = \max_{1\le k\le n}\left\{(n-k)\binom{k}{2}\right\},$$

achieved when $k\sim\frac{2n}{3}$. Consequently $\delta(1_r3_b2_b) = \frac{4}{9}$.

Patterns of length over three may also be studied so long as they contain no more than three colored blocks and each colored block is equivalent to a non-colored pattern with known packing density.

3.3.4. *Pattern* $1_b3_r4_r2_b$. Theorem 3.9 implies that the optimal colored permutation $\hat{\chi}$ is of the form B_1RB_2 with $B_1 < B_2 < R$, then

$$f_n(1_b3_r4_r2_b) = f(1, B_1) \cdot f(12, R) \cdot f(1, B_2).$$

For convenience, let $|B_1| = x$, $|R| = y$, and $|B_2| = z$. Then $x + y + z = n$ and thus for any fixed y

$$f(1, B_1) \cdot f(1, B_2) = x \cdot z \leq \frac{2(n-y)^2 - 1 + (-1)^{n-y}}{8}$$

with equality when $|x - z| \leq 1$. Consequently

$$f_n(1_b3_r4_r2_b) = \max_{2 \leq y \leq n-2} \left\{ \frac{2(n-y)^2 - 1 + (-1)^{n-y}}{8} \cdot \binom{y}{2} \right\},$$

achieved when $y \sim \frac{n}{2}$. Hence $\delta(1_b3_r4_r2_b) = \frac{3}{16}$.

3.3.5. *Pattern* $3_b2_b4_r6_r5_r1_b$. Theorem 3.8 implies that the optimal permutation $\hat{\chi}$ is of the form B_1RB_2 with $B_2 < B_1 < R$. Hence

$$f_n(3_b2_b4_r6_r5_r1_b) = f(21, B_1) \cdot f(132, R) \cdot f(1, B_2).$$

Letting $|B_1| = x$, $|R| = y$, $|B_2| = z$ and fixing y again, we have $x + z = n - y$ and

$$f(21, B_1) \cdot f(1, B_2) \leq \binom{x}{2} \cdot z.$$

This expression is maximized when $x \sim \frac{2(n-y)}{3}$ and $z \sim \frac{n-y}{3}$. From [15] we have $f_y(132) \sim (2\sqrt{3}-3)\frac{y^3}{6}$, hence

$$\begin{aligned} &f_n(3_b 2_b 4_r 6_r 5_r 1_b) \\ &\sim \max_{3 \le y \le n-3} \left\{ \frac{2(n-y)}{6} \binom{2(n-y)-3}{6} \cdot \binom{2(n-y)}{3} \cdot (2\sqrt{3}-3)\frac{y^3}{6} \right\}, \end{aligned}$$

achieved when $y \sim \frac{n}{2}$. Thus $\delta(3_b 2_b 4_r 6_r 5_r 1_b) = \frac{5}{9}(2\sqrt{3}-3)$.

Here we considered the question of packing colored patterns into colored permutations. We summarize our results in the following Theorem:

Theorem 3.1 ([6]). *For a colored pattern with at most three colored blocks, the optimal colored permutation with respect to a given colored pattern always shares the same number and arrangement of the colored blocks as those of the pattern.*

It is worth noting that, with our characterizations of the optimal permutations, the colored version of the pattern packing question is in some sense easier than the non-colored version and encompasses a wider range of patterns. For instance, the optimal permutation for the colored pattern $6_r 1_r 3_r 2_r 5_b 4_b$ can indeed be characterized since it contains only three colored blocks and each block is a layered pattern. However, the non-colored pattern 613254 is not layered and its optimal permutation is much more difficult to characterize.

It is natural to conjecture that the following holds in general, which we post as a question.

Question 3.10. *Is it true that the optimal colored permutation with respect to a given colored pattern always shares the same number and arrangement of the colored blocks as those of the pattern?*

Finally we will give some more details on calculating the packing densities of colored patterns. In addition to listing computational results, we also provide some combinatorial explanations for some identities of packing densities in Section 3.4.

3.4. **Evaluation of Packing Densities.** First note that for any optimal permutation $\hat{\chi}$ with respect to a pattern ϕ, we have

$$f_n(\phi) \sim \delta(\phi)\binom{n}{|\phi|}$$

as $n \to \infty$.

Let ϕ be a pattern (of length k) of the form $\phi_1\phi_2\phi_3$ where each ϕ_i $(i = 1, 2, 3)$ is a colored block of ϕ. Theorem 3.1 implies that the optimal permutation $\hat{\chi}$ of length n must be of the form $\chi_1\chi_2\chi_3$ such that χ_i is of the same color as ϕ_i for $i = 1, 2, 3$ and the numerical orderings of ϕ_1, ϕ_2, ϕ_3 and χ_1, χ_2, χ_3 are the same.

Since each ϕ pattern in $\hat{\chi}$ is a result of a ϕ_i pattern in χ_i for $i = 1, 2, 3$, each χ_i must be optimal with respect to ϕ_i. Hence

$$\begin{aligned}
&f(\phi, \hat{\chi}) \\
=&f(\phi_1, \chi_1) f(\phi_2, \chi_2) f(\phi_3, \chi_3) \\
\sim&\left(\delta(\phi_1)\binom{|\chi_1|}{|\phi_1|}\right)\left(\delta(\phi_2)\binom{|\chi_2|}{|\phi_2|}\right)\left(\delta(\phi_2)\binom{|\chi_2|}{|\phi_2|}\right) \\
\sim&\delta(\phi_1)\delta(\phi_2)\delta(\phi_3)\left(\frac{|\chi_1|^{|\phi_1|}}{|\phi_1|!}\frac{|\chi_2|^{|\phi_2|}}{|\phi_2|!}\frac{|\chi_3|^{|\phi_3|}}{|\phi_3|!}\right)
\end{aligned}$$

as $|\chi_i| \to \infty$ for $i = 1, 2, 3$.

Given that $|\chi_1| + |\chi_2| + |\chi_3| = n$, $|\phi_1| + |\phi_2| + |\phi_3| = k$ and let $n_i = |\chi_i|$, $a_i = |\phi_i|$ for any i,

$$|\chi_1|^{|\phi_1|}|\chi_2|^{|\phi_2|}|\chi_3|^{|\phi_3|} = n_1^{a_1} n_2^{a_2}(n - n_1 - n_2)^{a_3}.$$

With fixed n_1, $f(x) := x^{a_2}((n - n_1) - x)^{a_3}$ is maximized when

$$f'(x) = 0 \Leftrightarrow x \sim \frac{a_2 \cdot (n - n_1)}{k - a_1}.$$

Following similar standard calculations, we have the maximum

$$|\chi_1|^{|\phi_1|}|\chi_2|^{|\phi_2|}|\chi_3|^{|\phi_3|}$$

when

$$|\chi_i| \sim \frac{|\phi_i|}{k} \cdot n$$

for $i = 1, 2, 3$, as $n \to \infty$.

Consequently

$$
\begin{aligned}
&\delta(\phi)\\
&= \lim_{n\to\infty} \frac{f(\phi,\hat{\chi})}{\binom{n}{k}}\\
&= \lim_{n\to\infty} \frac{1}{\binom{n}{k}} \frac{1}{|\phi_1|!|\phi_2|!|\phi_3|!}\delta(\phi_1)\delta(\phi_2)\delta(\phi_3)\cdot\\
&\qquad\qquad\qquad\cdot \left(\frac{|\phi_1|}{k}\cdot n\right)^{|\phi_1|}\left(\frac{|\phi_2|}{k}\cdot n\right)^{|\phi_2|}\left(\frac{|\phi_3|}{k}\cdot n\right)^{|\phi_3|}\\
&= \lim_{n\to\infty} \frac{1}{n^k} \frac{k!}{|\phi_1|!|\phi_2|!|\phi_3|!}\delta(\phi_1)\delta(\phi_2)\delta(\phi_3)\cdot\\
&\qquad\qquad\qquad\cdot \frac{|\phi_1|^{|\phi_1|}|\phi_2|^{|\phi_2|}|\phi_3|^{|\phi_3|}n^{|\phi_1|+|\phi_2|+|\phi_3|}}{k^{|\phi_1|+|\phi_2|+|\phi_3|}}\\
&= \binom{k}{|\phi_1|,|\phi_2|,|\phi_3|}\delta(\phi_1)\delta(\phi_2)\delta(\phi_3)\frac{|\phi_1|^{|\phi_1|}|\phi_2|^{|\phi_2|}|\phi_3|^{|\phi_3|}}{k^k}
\end{aligned}
$$

and the following theorem follows.

Theorem 3.2. *Given a pattern ϕ (of length k) of the form $\phi_1\phi_2\phi_3$ with colored blocks ϕ_i ($i = 1, 2, 3$), we have*

$$\delta(\phi) = \binom{k}{|\phi_1|,|\phi_2|,|\phi_3|}\delta(\phi_1)\delta(\phi_2)\delta(\phi_3)\frac{|\phi_1|^{|\phi_1|}|\phi_2|^{|\phi_2|}|\phi_3|^{|\phi_3|}}{k^k}.$$

Of course, similar arguments can be applied to a pattern ϕ (of length k) with two colored blocks, of the form $\phi_1\phi_2$.

Theorem 3.3. *Given a pattern $\phi = \phi_1\phi_2$ with two colored blocks ϕ_1 and ϕ_2, we have*

$$\delta(\phi) = \binom{k}{|\phi_1|,|\phi_2|}\delta(\phi_1)\delta(\phi_2)\frac{|\phi_1|^{|\phi_1|}|\phi_2|^{|\phi_2|}}{k^k}.$$

Remark 3.11. As a simple but important consequence of Theorems 3.2 and 3.3, the packing density of a colored pattern with two or three colored blocks is entirely decided by the subpattern in each block regardless of the color or order of the blocks.

Following Theorem 3.2 and 3.3, it is easy to calculate the packing densities of many small patterns. Figure 10 lists some of the representative patterns and their packing densities. Note that, from Remark 3.11, only two colors are needed to consider all patterns with up to three colored blocks, in what follows we use a and a' to denote a number a with different colors.

Many interesting observations can be made through Figure 10. It is easy to see that $\delta(\phi) = 1$ if ϕ is a single colored pattern $1 \ldots k$ or $k \ldots 1$ for some k. Many other packing densities are obviously shared by different patterns. For instance,

$$\delta(12'3') = \delta(13'2') = \delta(123') = \delta(213') = \delta(31'2') = \delta(321') = ...$$

follows directly from Theorem 3.3 and the equivalence between subpatterns formed by a colored block.

Other less trivial identities include

$$\delta(2143) = \delta(124'3'), \tag{3.1}$$

which follows from the equivalent layers and colored blocks in the corresponding optimal permutations.

ϕ	$\delta(\phi)$	ϕ	$\delta(\phi)$
1	1	$124'3'$	0.375
12	1	$12'34$	0.1875
$12'$	0.5	12345	1
123	1	12543	0.3456
132	0.464...	$12'3'4'5'$	0.4096
$12'3'$	0.444...	$12'5'4'3'$	0.173...
$12'3$	0.222...	$13'2'5'4'$	0.1536
1234	1	$123'4'5'$	0.3456
1432	0.424...	$123'5'4'$	0.160...
2143	0.375	$12'345$	0.1728
$12'3'4'$	0.421875	$12'354$	0.080...
$12'4'3'$	0.196...	$123'45$	0.1536

FIGURE 10.

Following (3.1) it is then easy to see that

$$\delta(13'2'5'4') = \delta(123'45),$$

as the colored block $3'2'5'4'$ (from $13'2'5'4'$) yields the same packing density as two disjoint blocks 32 and 54, which in turn is equivalent to that of 12 and 45 (in $123'45$).

Another interesting observation from Figure 10 is that

$$\delta(124'3') = 2\delta(12'34) = 2\delta(12'43),$$

and

$$\delta(123'4'5') = 2\delta(12'345).$$

It is not difficult to see that the above relations were resulted from the simple fact $\delta(12) = 2\delta(12')$. A more general statement is as follows.

Proposition 3.12. *Let $\phi_1 = 12A'$ and $\phi_2 = 12'A$ for some A with different colors in ϕ_1 and ϕ_2 respectively, then*

$$\delta(\phi_1) = 2\delta(\phi_2).$$

References

[1] M. H. Albert, M. D. Atkinson, C. C. Handley, D. A. Holton, and W. Stromquist. On packing densities of permutations. *Electron. J. Combin*, 9(1), 2002.

[2] J. Backelin, J. West, and G. Xin. Wilf-equivalence for singleton classes. *Advances in Applied Mathematics*, 38(2):133–148, 2007.

[3] M. Bóna. Exact enumeration of 1342-avoiding permutations: a close link with labeled trees and planar maps. *Journal of Combinatorial Theory, Series A*, 80(2):257–272, 1997.

[4] P. Erdös and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.

[5] I. M. Gessel. Symmetric functions and p-recursiveness. *Journal of Combinatorial Theory, Series A*, 53(2):257–285, 1990.

[6] M. Just and H. Wang. Note on packing densities in colored permutations. *to appear in Online Journal of Analytic Combinatorics*, 2016.

[7] D. E. Knuth. *The art of computer programming: sorting and searching*, volume 3. Pearson Education, 1998.

[8] M. Lothaire. *Combinatorics on words*. Cambridge university press, 1997.

[9] P. A. MacMahon. *Combinatory analysis*, volume 2. Courier Corporation, 1916.

[10] A. L. Price. Packing densities of layered* patterns. 1997.

[11] C. D. Savage and H. S. Wilf. Pattern avoidance in compositions and multiset permutations. *Advances in Applied Mathematics*, 36(2):194–201, 2006.

[12] C. Schensted. Longest increasing and decreasing subsequences. *Canad. J. Math*, 13(2):179–191, 1961.

[13] Z. Stankova and J. West. A new class of wilf-equivalent permutations. *Journal of Algebraic Combinatorics*, 15(3):271–290, 2002.

[14] Z. E. Stankova. Forbidden subsequences. *Discrete Mathematics*, 132(1):291–316, 1994.

[15] W. Stromquist. Packing layered posets into posets. *Unpublished typescript*, 1993.

Matthew Just, Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, USA

E-mail address: mj00788@georgiasouthern.edu