Journal of Combinatorics
Volume 2, Number 1, 111–137, 2011

Packing a binary pattern in compositions

Ragnar Freij and Toufik Mansour

In this article we generalize packing density problems from permutations and words to compositions. We are able to find the packing density for some classes of subsequence and generalized patterns and all the three letter binary patterns.

AMS 2000 subject classifications: Primary 05A15; secondary 05A16.

1. Introduction

Let $\pi = \pi_1 \cdots \pi_m$ and $\tau = \tau_1 \cdots \tau_\ell$ be two words. An *occurrence* of τ in π is a subsequence $1 \leq i_1 < i_2 < \cdots < i_\ell \leq m$ such that $\pi_{i_1}, \ldots, \pi_{i_\ell}$ is order-isomorphic to τ; in such a context, τ is usually called a *pattern*.

Recently, much attention has been paid to the problem of counting the number of permutations of length n (k-ary words of length n, compositions of n) containing a given number $r \geq 0$ of occurrences of a certain pattern τ. Most of the authors consider only the case $r = 0$, thus studying permutations (k-ary words, compositions) *avoiding* a given pattern, see [4, 9]. There is considerably less research on other aspects of pattern containment, specifically, on packing patterns into words over a totally ordered alphabet, but see [1, 7, 8, 10, 11, 12] for the permutation case and [2, 5, 6, 13, 14, 15] for the more general pattern case.

While several of the above cited papers have defined packing density on the set of permutations and on the set of k-ary words, in this paper we take the first systematic step in studying the packing density on the set of compositions. This generalization to compositions follows the current interest in compositions which have been studied from different aspects in the literature, see [9] and references therein. The results in this paper add a facet to this research.

2. Notation

Let $\mathbb{N} = \{1, 2, \ldots\}$ be the set of positive integers. A composition $\pi = \pi_1 \cdots \pi_m$ of n in $\mathbb{N}$ is an ordered collection of one or more positive integers whose sum is n. We will call n the *size* of π, and denote it $n = |\pi|$. The number of summands or letters, namely m, is called the number of parts of the composition. Clearly, the number of compositions of n is given by 2^{n-1} and the number of compositions of n with m parts is given by $\binom{n-1}{m-1}$, for all $n \geq m \geq 1$. We denote the set of compositions of n by C_n and we denote the set of compositions of n with m parts by $C_{n,m}$. We define a composition to be *reduced* if its letters are the k first integers for some k. For example, 131 is not reduced, but 121 is. Clearly, every pattern is equivalent (order-isomorphic) to a reduced one.

Given a composition $\pi = \pi_1 \cdots \pi_m$ of n in $\mathbb{N}$ and a pattern $\tau = \tau_1 \cdots \tau_\ell$, let $\nu(\tau, \pi)$ be the total number of occurrences of the pattern τ in π. Define

$$\begin{aligned} \mu(\tau, n) &= \max\{\nu(\tau, \pi) \mid \pi \in C_n\}, \\ d(\tau, \pi) &= \frac{\nu(\tau, \pi)}{\binom{n}{\ell}}, \\ \delta(\tau, n) &= \frac{\mu(\tau, n)}{\binom{n}{\ell}} = \max\{d(\tau, \pi) \mid \pi \in C_n\}. \end{aligned}$$

Let τ be a pattern. We will say that a composition λ of n is τ-optimal if $d(\tau, \lambda) \geq d(\tau, \eta)$ for every composition η of n. If we let $\bar{\sigma}$ denote the reversal of the word σ, it is clear that $\nu(\bar{\tau}, \bar{\pi}) = \nu(\tau, \pi)$, so packing densities are invariant under reversal.

Note that we compare $\nu(\tau, \pi)$ to the largest number of subwords of length ℓ in a word of size n which is $\binom{n}{\ell}$. This should be compared with the case of packing into words, studied in [2, 6], where we instead normalize with the maximal number of subwords of length ℓ in a word with m letters, i.e. $\binom{m}{\ell}$.

The main reason for this definition is finding that the number $\delta(\tau, n)$ for large n converges to a real number, see our results below. Thus, we denote $\lim_{n\to\infty} \delta(\tau, n)$ by $\delta(\tau)$ when it exists, and $\delta(\tau)$ is said to be the *packing density* of the pattern τ. Our interest is to study the asymptotic behavior of $\delta(\tau, n)$ as $n \to \infty$, that is, finding $\delta(\tau)$.

We will use a minimal amount of asymptotic notation, but we will let $f(n) \approx g(n)$ mean that $\frac{f(n)}{g(n)} \to 1$ as $n \to \infty$. Throughout, we will let x^a denote the word $xx \cdots x$ of length a.

3. Subsequence patterns

3.1. Letter reduction

First, we will show how some very natural operations on a composition gives a denser packing of our pattern

Lemma 3.1. *Let $\tau = \tau_1\tau_2\cdots\tau_\ell$ be a reduced pattern, and let n be any integer. Then there is a τ-optimal composition of n that is reduced.*

Proof. Let $\lambda = \lambda_1\cdots\lambda_j$ be a composition of n, and suppose there is some $m \in \mathbb{N}$ such that $m \in \{\lambda_i\}$, but $m-1 \notin \{\lambda_i\}$. Then we can replace every occurrence of m in λ by $1, m-1$. This operation can only increase the number of occurrences of τ in λ. Repeating this operation, we will obtain a reduced composition of the same size, without destroying τ-optimality. $\square$

Lemma 3.2. *If τ is a monotone non-decreasing pattern, then there is a τ-optimal composition that is also monotone non-decreasing.*

Proof. Let π be any composition of n, and let π' be the composition obtained by sorting π increasingly. Since τ is non-decreasing, any occurrence of τ in π will give an occurrence in π'. Thus, if π were τ-optimal, then so is π'. $\square$

Lemma 3.3. *Let $\tau = 1\cdots\ell$ be a reduced and strictly increasing pattern. Then there is a τ-optimal composition using only the letters $1, \ldots, \binom{\ell+1}{2} - 1$.*

Proof. Suppose that π is a τ-optimal composition using a letter $k \geq \binom{\tau_\ell+1}{2}$. By Lemma 3.2 we can assume $\pi = 1^{a_1}2^{a_2}\cdots k^{a_k}$. Construct π' from π by replacing one letter k by $1, 2, \cdots, \ell-1, k-\binom{\ell}{2}$, and sorting the letters increasingly. Note that $k - \binom{\ell}{2} \geq \ell$, so we have added ℓ distinct letters. Consider any occurrence of τ in π. If it does not use the letter k, it is still an occurrence in π'. If it does use the letter k, however, it can not use all the letters $1, 2, \cdots, \ell-1, k-\binom{\ell}{2}$, because the pattern has only ℓ letters. Replace the letter k in the occurrence by the first of the added letters that were not already in the occurrence. This gives an occurrence of τ in π'. This shows that $\nu(\tau, \pi') \geq \nu(\tau, \pi)$, while π' has strictly fewer letters k than π does. Thus there is some τ-optimal composition using no letters $k \geq \binom{\ell+1}{2}$. $\square$

We say that the pattern $\tau = \tau_1\cdots\tau_k$ is *unimodal* if there is some i for which

$$\tau_1 \leq \cdots \leq \tau_i \geq \cdots \geq \tau_k.$$

If there is some i for which

$$\tau_1 \geq \cdots \geq \tau_i \leq \cdots \leq \tau_k,$$

then τ will be called *anti-unimodal.*

Then we have the following proposition.

Proposition 3.4. *If τ is a unimodal (anti-unimodal) pattern, then there is a unimodal (anti-unimodal) τ-optimal composition of every size.*

Proof. We prove the unimodal case. The arguments in the anti-unimodal case are exactly the same.

In fact, we will prove the following, stronger, statement: Let A be a finite multiset of letters. Among the words on the letters of A, that have the maximal number of occurrences of τ, there is a unimodal one. We will let r denote the number of distinct letters in A.

We will prove the statement by joint induction, first on the length k of τ, and then on r. Note that $\nu(1,\pi) = |A|$ for every word π on the letters in A. Moreover, if $r = 1$ we have $\nu(\tau,\pi) = 0$ if τ is not constant, and $\nu(\tau,\pi) = \binom{|A|}{k}$ if $\tau = 1^k$. So the statement is true if $k = 1$ and if $r = 1$.

Suppose it holds if $k < k_0$, and if $k = k_0$, $r < r_0$. Let τ be a reduced unimodal pattern of length k_0, and A a multiset on r_0 distinct letters, the smallest of which is m. Let π be a word on the letters of A, that maximizes $\nu(\tau,\pi)$. We want to prove that π can be chosen with every occurrence of m in its beginning or end. Then the statement will follow by induction.

If $\tau_k \neq 1$, it is clear that all occurrences of m in π can be put in the left end of π, without decreasing $\nu(\tau,\pi)$. So now we assume $\tau_1 = \tau_k = 1$. Let $\tau = 1^a\tau'1^b$, where τ' contains no letter 1.

Write $\pi = m^c\rho m\sigma$, where ρ contains no letters m. Let d be the number of letters m in σ. Without loss of generality we assume that $\frac{c+1}{a} \leq \frac{d+1}{b}$ – otherwise consider $\bar{\tau}$ and $\bar{\pi}$ instead.

If π contained some letter m apart from the first and last ones, i.e. if σ is not all ms, we construct $\pi' = m^{a+1}\rho\sigma$ by moving the first such letter to the beginning of the composition.

This operation destroys

$$\binom{c}{a}\binom{d}{b-1}\nu(\tau',\sigma)$$

occurrences of τ. It also adds

$$\binom{c}{a-1}\binom{d}{b}\nu(\tau',\sigma)$$

occurrences. Since $\frac{c+1}{a} \geq \frac{d+1}{b}$ we get

$$\binom{c}{a}\binom{d}{b-1} \leq \binom{c}{a-1}\binom{d}{b},$$

so $\nu(\tau,\pi') \geq \nu(\tau,\pi)$.

Thus we can move all letters m to the far ends of π, without destroying optimality. So by induction, there is a unimodal τ-optimal ordering of every multiset A. Maximizing the occurrence number over all multisets, we hence obtain a unimodal τ-optimal composition. $\square$

In order to state our next result, we need the following notation. Let $\tau = \tau_1\tau_2\cdots\tau_\ell$ be any reduced pattern, we denote the *normal form* of τ as $\tau = \tau_{i_1}^{k_1}\tau_{i_2}^{k_2}\cdots\tau_{i_s}^{k_s}$ such that $\tau_{i_j} \neq \tau_{i_{j+1}}$ for all $j = 1, 2, \ldots, s-1$. For example, the normal form of the pattern 112133224 is $1^2 2^1 1^1 3^2 2^2 4^1$. Theorem 3.5 reduces, in many cases, the problem of finding the packing density to that of proving that an optimal packing has "the simplest form possible".

Theorem 3.5. *Let τ be any reduced pattern with the normal form*

$$\tau_{i_1}^{k_1}\tau_{i_2}^{k_2}\cdots\tau_{i_s}^{k_s},$$

and length $\ell = \sum_i k_i$. Assume the τ-optimal composition of n has the form

$$\pi = \tau_{i_1}^{a_1}\tau_{i_2}^{a_2}\cdots\tau_{i_s}^{a_s},$$

for all n. Then

$$\delta(\tau) = \frac{\ell!}{\ell^\ell}\prod_{j=1}^{s}\frac{(k_j/\tau_{i_j})^{k_j}}{k_j!}.$$

Proof. Let $a_j = n\alpha_j$ for all j. The condition $\sum_{j=1}^{s} a_j\tau_{i_j} = n$ is equivalent to that $\sum_{j=1}^{s}\tau_{i_j}\alpha_j = 1$, which implies that

$$d(\tau,\pi) = \prod_{j=1}^{s}\binom{a_j}{k_j} = \prod_{j=1}^{s}\frac{n\alpha_j(n\alpha_j-1)\cdots(n\alpha_j-k_j+1)}{k_j!}.$$

From the definitions we have that $1 \leq k_j \leq \ell$ and $0 < \alpha_j < 1$, so

$$(1)\qquad \prod_{j=1}^{s}\binom{n\alpha_j}{k_j} \approx \prod_{j=1}^{s}\frac{(n\alpha_j)^{k_j}}{k_j!} = n^{k_1+\cdots+k_s}\prod_{j=1}^{s}\frac{\alpha_j^{k_j}}{k_j!} = n^\ell\prod_{j=1}^{s}\frac{\alpha_j^{k_j}}{k_j!},$$

which gives that

$$\mu(\tau,n) \approx \frac{n^\ell}{k_1!\cdots k_s!}\max_{\tau_{i_1}\alpha_1+\cdots+\tau_{i_s}\alpha_s=1}\alpha_1^{k_1}\cdots\alpha_s^{k_s}.$$

When $\sum \beta_i = 1$, the function $\beta_1^{k_1} \cdots \beta_s^{k_s}$ is maximized at $\beta_i = k_i/\ell$. Using this, we obtain that

$$\mu(\tau, n) \approx \frac{n^\ell}{k_1! \cdots k_s!} \prod_{j=1}^{s} \left(\frac{k_j/\tau_{i_j}}{\ell} \right)^{k_j}.$$

Hence, from the definition of the packing density and $k_1 + \cdots + k_s = \ell$ we get the desired result. □

As a trivial example, if $\tau = 1^\ell$, then the τ-optimal composition of n is 1^n. Hence, Theorem 3.5 gives

$$\delta(\tau) = \frac{\ell!}{\ell^\ell} \cdot \frac{\ell^\ell}{\ell!} = 1.$$

3.2. Binary patterns

In this subsection we find the packing density for all monotone binary patterns, and for all three letter patterns with repeated letters.

Theorem 3.6. *For any positive integers x and y, we have*

$$\delta(1^x 2^y) = \delta(2^y 1^x) = \binom{x+y}{x} \frac{x^x y^y}{2^y (x+y)^{x+y}}.$$

Proof. We consider only the case $\tau = 1^x 2^y$, and let the case $\bar{\tau} = 2^x 1^y$ follow by reversal.

By Lemma 3.2, a τ-optimal composition of n is increasing, and has the normal form

$$\pi = 1^{a_1} 2^{a_2} \cdots (k-1)^{a_{k-1}} k^{a_k}$$

Assume $k \geq 3$. Then define

$$\pi' = 1^{a_1 + a_k} 2^{a_2} \cdots (k-2)^{a_{k-2}} (k-1)^{a_{k-1} + a_k}.$$

We think of π' as obtained from π by replacing every letter k by one letter $k-1$, and adding a_k letters 1 in the beginning. This deletes

$$\binom{a_{k-1}}{x} \binom{a_k}{y}$$

occurrences of τ, namely the subsequences $(k-1)^x k^y$ in π. Notice that all subsequences $j^x k^y$ in π, with $j < k-1$, yield a subsequence $j^x (k-1)^y$ in π'.

On the other hand, we have added at least

$$\binom{a_1+a_k}{x}\binom{a_{k-1}+a_k}{y}-\binom{a_1}{x}\left[\binom{a_{k-1}}{y}+\binom{a_k}{y}\right]$$

occurrences of τ. These are the subsequences $1^x(k-1)^y$ in π' that are not induced by subsequences $1^x(k-1)^y$ or 1^xk^y in π. We claim that

$$\binom{a_1+a_k}{x}\binom{a_{k-1}+a_k}{y}-\binom{a_1}{x}\left[\binom{a_{k-1}}{y}+\binom{a_k}{y}\right]-\binom{a_{k-1}}{x}\binom{a_k}{y}\geq 0.$$

Indeed, to prove the inequality, we distinguish three different cases. Note that the left-hand side of the inequality is at least

$$\binom{a_k}{x}\binom{a_{k-1}+a_k}{y}-\binom{a_{k-1}}{x}\binom{a_k}{y},$$

which is ≥ 0 if $y\geq x$, or if $x>y$ and $a_k\geq a_{k-1}$.

So assume $x>y$ and $a_{k-1}>a_k$. Then rewrite the left-hand side as

$$\sum_{(i,j)\notin\{(0,0),(0,y)\}}\binom{a_1}{x-i}\binom{a_k}{i}\binom{a_{k-1}}{y-j}\binom{a_k}{j}-\binom{a_{k-1}}{x}\binom{a_k}{y}.$$

If $y\geq 2$, then the term $i=0$, $j=1$ is greater than $\binom{a_{k-1}}{x}\binom{a_k}{y}$. Finally, if $x>y=1$, the term $i=1$, $j=0$ is greater than $\binom{a_{k-1}}{x}\binom{a_k}{y}$. This completes the proof of the claim. Thus, in any case we have added more occurrences than we deleted, so π was not optimal.

Hence we can delete every letter ≥ 3, so we have shown that any 1^x2^y-optimal composition has the form 1^a2^b. By Theorem 3.5 and reversal, we get

$$\delta(1^x2^y)=\delta(2^y1^x)=\frac{(x+y)!}{(x+y)^{x+y}}\cdot\frac{x^x}{x!}\cdot\frac{(y/2)^y}{y!}=\binom{x+y}{x}\frac{x^xy^y}{2^y(x+y)^{x+y}}.$$

□

We are now ready to determine the packing densities of all three letter binary words, into compositions. Again we will prove that an optimal composition is also binary.

Theorem 3.7. *We have*

$$\delta(112)=\delta(211)=\frac{2}{9},\quad \delta(121)=\frac{1}{9},\quad \delta(212)=\frac{1}{18},\quad \delta(221)=\delta(122)=\frac{1}{9}.$$

Proof. The monotone patterns in the theorem follow from Theorem 3.6, so we need only consider the patterns 121 and 212.

Let π be a 121-optimal composition. By Proposition 3.4, π has the structure

$$\pi = 1^{a_1}2^{a_2}\cdots k^{a_k+b_k}\cdots 2^{b_2}1^{b_1}.$$

We enumerate the occurrences by summing over the repeated letter in the occurrence, and get

$$\nu(121,\pi) = \sum_{i=1}^{k-1} a_i b_i (a_{i+1} + b_{i+1} + \cdots + a_k + b_k).$$

For fixed $a_i + b_i$, the expression is maximized over the reals when $a_i = b_i$, and over the naturals when $|a_i - b_i| \leq 1$. Hence, $b_i < b_{i+1}$ would imply $a_i \leq a_{i+1}$. If this were the case, the number of occurrences would increase, while n would decrease, by interchanging (a_i, b_i) and (a_{i+1}, b_{i+1}). So for π to be optimal, we must have $b_1 \geq \cdots \geq b_k$.

Now if $k > 2$, we could construct

$$\pi' = 1^{a_1+1}2^{a_2}\cdots(k-1)^{a_{k-1}+1}k^{a_k-1+b_k}\cdots 1^{b_1}$$

by splitting one occurrence of k into one occurrence of $k-1$ and one occurrence of 1, in the first half of the word. Then

$$\begin{aligned}
&\nu(121,\pi') - \nu(121,\pi)\\
&\quad = b_1 \cdot (a_2 + b_2 + \cdots + a_k + b_k) + b_{k-1}\cdot(a_k - 1 + b_k) - a_{k-1}b_{k-1}\\
&\quad > a_{k-1}b_1 - a_{k-1}b_{k-1} \geq 0
\end{aligned}$$

contradicting the assumption that π was optimal.

So a 121-optimal composition has only letters 1 and 2, so we can use Theorem 3.5 to conclude that $\delta(121) = \frac{1}{9}$.

For the pattern 212, consider an optimal composition π using the letters $1 \ldots k$. By Proposition 3.4, π has the form $k^{a_k}\cdots 1^{a_1+b_1}\cdots k^{b_k}$.

Again, we enumerate the occurrences by summing over the repeated letter in the occurrence. This time, we get

$$\nu(212,\pi) = \sum_{i=2}^{k} a_i b_i (a_1 + b_1 + \cdots + a_{i-1} + b_{i-1}),$$

and fixing $a_i + b_i$ this is maximized when $a_i = b_i$. If we agree to maximize the function over the half-integers rather than over the integers, we may assume $a_i = b_i$.

If $k > 3$, we want to contradict the assumption that π was optimal, by different reductions. If $a_k \le a_i$ for all $i \ne k$, we construct

$$\pi' = (k-1)^{a_{k-1}} \cdots 2^{a_2} 1^{2(a_1 + k a_k)} 2^{a_2} \cdots (k-1)^{a_{k-1}}$$

by replacing each occurrence of k by k occurrences of 1. This destroys $2a_k^2(a_1 + \cdots + a_{k-1})$ occurrences of 212, and adds $2ka_k(a_2^2 + \cdots + a_{k-1}^2) \ge 2ka_k^2(a_1 + \cdots + a_{k-1})$ occurrences. Thus π was not optimal.

So assume $a_k \ge a_i$ for some i. Then there is some i for which $a_i \le a_{i+1}$. Construct

$$\pi'' = k^{a_k} \cdots (i+1)^{a_i}\, i^{a_{i+1}} \cdots 1^{2(a_1 + a_{i+1} - a_i)} \cdots i^{a_{i+1}}\, (i+1)^{a_i} \cdots k^{a_k}.$$

In words, we interchange the block lengths of i and $i+1$. This allows us to insert $2(a_{i+1} - a_i)$ letters 1 in the bottom.

So going from π to π'', the number of subsequences $(i+1)i(i+1)$ is reduced from $2a_{i-1}a_i^2$ to $2a_{i-1}^2 a_i$, so has decreased by $2a_{i-1}a_i(a_i - a_{i-1})$. The number of subsequences $j1j$ has increased by $2(a_i - a_{i-1})(a_2^2 + \cdots + a_k^2) \ge 2a_i^2(a_i - a_{i-1})$. All other occurrences of 212 are naturally preserved. Thus, we have added more occurrences than we have destroyed, so π was not optimal.

To conclude, π is not 212-optimal if $k \ge 3$, so any 212-optimal composition uses only the letters 1 and 2. By Theorem 3.5, $\delta(212) = \frac{1}{18}$. $\square$

3.3. Three letter permutation patterns

In order to complete the list of packing densities of three letter patterns among compositions, we have to find $\delta(\tau)$, where τ is a permutation pattern of length three. By the reversal operation, we can reduce these 6 cases to three, namely, $\delta(123)$, $\delta(132)$ and $\delta(213)$.

Proposition 3.8. *We have* $\delta(213) = \frac{1}{27}$.

Proof. Let $\pi = \pi_1 \cdots \pi_m$ be any 213-optimal composition for $n \ge 6$. By Proposition 3.4, we can assume that π has the form $\pi' 1^a \pi''$, where each letter of π' and π'' is at least 2, π' is decreasing and π'' is increasing. Replacing a letter $x \ge 3$ in π' by 21^{x-2}, we increase the number occurrences of 213 which gives a contradiction to the 213-optimality of π. Thus π can be written as $2^b 1^a \pi''$, where each letter of π'' is at least 2. If we have in π'' a letter $x \ge 4$, then changing x to $1^s 3^t$ where t maximal, we get a contradiction of the optimality of π. If π'' contains a letter 2 then moving it on the left side of the first occurrence of the letter 1 we increase the number occurrence of 213. We thus obtain that π has the form $2^b1^a3^c$. The number of occurrences of 213 in π is now the maximal abc, for which $a+2b+3c=n$, which implies that

$$\delta(213)=\lim_{n\to\infty}\frac{\frac{n}{3}\frac{n}{6}\frac{n}{9}}{n^3/6}=\frac{1}{27},$$

as claimed. $\square$

We proceed by finding the packing densities of the patterns 123 and 132. It is easy to see that 123-optimal composition has the form $1^{a_1}2^{a_2}\cdots k^{a_k}$ and the 132-optimal composition has the form $1^{a_1}k^{a_k}\cdots 2^{a_2}$. Indeed, by Proposition 3.4, the optimal compositions are unimodal, and moving every letter $j<k$ to the appropriate side of k^{a_k} does not reduce the number of occurrences. In both cases we have $\sum_i a_i = N$, and the number of occurrences is

$$\sum_{1\le i<j<l\le k} a_ia_ja_l.$$

In particular, this shows that $\delta(123)=\delta(132)$.

Proposition 3.9. *The packing densities* $\delta(123)$ *and* $\delta(132)$ *both equal the unique positive root of* $17496x^4+38070x^3+2610x^2-100x-3$*, which is approximately* 0.041126.

Proof. We already argued that it suffices to consider the pattern 123. By Lemma 3.3 it suffices to consider compositions of the form $1^{a_1}2^{a_2}3^{a_3}4^{a_4}5^{a_5}$, and maximize

$$\sum_{1\le i<j<l\le 5} a_ia_ja_l$$

subject to $\sum_i a_i = n$. Turning to the corresponding real optimization problem we get

$$\delta(123)=6\max\left\{\sum_{1\le i<j<l\le 5}\alpha_i\alpha_j\alpha_l : \mathbf{0}\le\alpha, \sum i\alpha_i=1\right\}.$$

Differentiation yields no interior extreme points, so we conclude $\alpha_5=0$ for the optimum. Differentiating again, and setting the derivatives to zero, we get the equations

$$\begin{aligned}
0 &= \alpha_3 + \alpha_2 - 8\alpha_4\alpha_3 - 8\alpha_4\alpha_2 - 8\alpha_3\alpha_2 - 3\alpha_3^2 - 2\alpha_2^2 \\
0 &= \alpha_4 + \alpha_2 - 6\alpha_4\alpha_3 - 8\alpha_4\alpha_2 - 6\alpha_3\alpha_2 - 4\alpha_4^2 - 2\alpha_2^2 \\
0 &= \alpha_4 + \alpha_3 - 8\alpha_4\alpha_3 - 4\alpha_4\alpha_2 - 4\alpha_3\alpha_2 - 4\alpha_4^2 - 3\alpha_3^2 \\
\alpha_1 &= 1 - (5\alpha_5 - 4\alpha_4 - 3\alpha_3 - 2\alpha_2)
\end{aligned}$$

This system of equations has a unique solution with all α_i in $(0,1)$, as can be easily seen with any computer algebra system. Indeed, α_i will be a root of the polynomial

$$\begin{aligned}
P_4 &= -1 + 34t - 216t^2 + 252t^3 + 432t^4 && \alpha_4 \approx 0.03828 \\
P_3 &= -1 + 13t + 6t^2 - 234t^3 + 162t^4 && \alpha_3 \approx 0.08361 \\
P_2 &= -1 + 2t + 56t^2 - 156t^3 + 72t^4 && \alpha_2 \approx 0.14245,
\end{aligned}$$

so $\alpha_1 = 1 - (4\alpha_4 - 3\alpha_3 - 2\alpha_2) \approx 0.31116$. Now we find

$$\delta(123) = 6(\alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2\alpha_4 + \alpha_1\alpha_3\alpha_4 + \alpha_2\alpha_3\alpha_4) \approx 0.041126,$$

to be the (unique) positive root of $17496x^4 + 38070x^3 + 2610x^2 - 100x - 3$. $\square$

4. Subword patterns

In this section, we will consider another notion of patterns. This time, an *occurrence* of τ in π is a subword $\pi_i\pi_{i+1}\dots\pi_{i+\ell}$ that is order-isomorphic to τ. This is equivalent to generalized patterns with no dash, as defined in [3]. We will let $\nu_w(\tau,\pi)$ denote the number of occurrences of τ as a subword in π. We will also define

$$\begin{aligned}
\mu_w(\tau,n) &= \max\{\nu_w(\tau,\pi) \mid \pi \in C_n\}, \\
d_w(\tau,\pi) &= \frac{\nu_w(\tau,\pi)}{n}, \\
\delta_w(\tau,n) &= \frac{\mu_w(\tau,n)}{n} = \max\{d_w(\tau,\pi) \mid \pi \in C_n\}, \\
\delta_w(\tau) &= \lim_{n\to\infty} \delta_w(\tau,n)
\end{aligned}$$

When no confusion can arise, we may omit the subscript w. Again, $\nu(\bar\tau,\bar\pi) = \nu(\tau,\pi)$, so packing densities are invariant under reversals. This is why we can restrict attention to patterns $\tau = \tau_1 \cdots \tau_k$ where $\tau_1 \le \tau_k$.

For any word τ, we define its m:th power τ^m to be the concatenation of m copies of m.

It is clear that $\nu_w(\tau, \tau^k) \geq k$, and $\tau^k \in C_{kn}$ if $\tau \in C_n$. On the other hand, each letter of π can start at most one occurrence of τ, and the length of π is no greater than its size $|\pi|$. So for any pattern we have $\frac{1}{|\tau|} \leq \delta_w(\tau) \leq 1$.

We define a word $\tau = \tau_1 \cdots \tau_k$ to be *i-overlapping*, $1 < i < k-1$ if the initial segment $\tau_1 \cdots \tau_i$ of τ is order-isomorphic to the final segment $\tau_{k+1-i} \cdots \tau_k$. If τ is not i-overlapping for any $1 < i < k-1$, then τ is said to be *non-overlapping.*

The classes of i-overlapping and non-overlapping patterns are clearly closed under reversal.

Example 4.1. 1324 is 2-overlapping, since $13 \cong 24$.

1432 is non-overlapping, since every initial sequence starts with an ascent, while no final segment does.

The non-overlapping patterns form a nice class, since they can only be packed in a very limited number of ways. The following definition gives the crucial operation

Definition 4.1. If σ has length k, τ has length ℓ, and $\sigma_k - \tau_1 = j \geq 0$, define the *gluing* of σ and τ to be

$$\langle \sigma, \tau \rangle = \sigma_1 \cdots \sigma_k \tau_2^* \cdots \tau_\ell^*,$$

where

$$\tau_i^* = \begin{cases} \tau_i & \text{if } \tau_i < \tau_1 \\ \tau_i + j & \text{otherwise} \end{cases}$$

We note that gluing is associative, and define the *m:th glued power* of τ in the natural way, by

$$\tau^{\langle m \rangle} = \begin{cases} \tau & \text{if } m = 1 \\ \langle \tau^{\langle m-1 \rangle}, \tau \rangle & \text{otherwise} \end{cases}$$

These definitions are most naturally thought of when $\tau_1 = 1$, in which case we just shift the second pattern up by j, so that we can identify the last letter of the first pattern with the first letter of the latter.

If τ has length k, we note that $\tau^{\langle m \rangle}$ has an occurrence of τ starting at position $(k-1)i + 1$ for $i = 0, \ldots, m-1$. As an example, $132^{\langle 3 \rangle} = 1324354$ has three occurrences of 132, namely 132, 243 and 354.

The glued powers, and powers of these, happen to be optimal for non-overlapping patterns, and this allows us to determine their packing density exactly.

Theorem 4.1. *Let τ be a non-overlapping reduced pattern of length k, and let s be the number of letters in τ that are greater than or equal to $min(\tau_1, \tau_k)$.*

If $s\tau_1 \le (s-1)\tau_k$, then $\delta_\tau = \frac{1}{|\tau|}$.

In general,

$$\delta_\tau = \frac{p}{p|\tau| - (p-1)\tau_1 + (s-1)(\tau_k - \tau_1)\binom{p}{2}},$$

where $p = \lceil \frac{\tau_1}{(s-1)(\tau_k - \tau_1)} \rceil$.

Proof. Assume without loss of generality that $\tau_1 \le \tau_k$, so s is the number of letters in τ that are greater than or equal to τ_1.

Let π_m be a composition of minimal size in which there are m occurrences of τ. The essential part of the proof will be finding the asymptotics of the sequence $\{|\pi_m|\}$.

It is clear that π_m can not contain any superfluous letters, i.e. every letter in π_m must appear in some occurrence of τ. Since τ is non-overlapping, no letter in any word can appear in more than two subword-occurrences of τ. Moreover, if a letter appears in two occurrences, it appears as τ_1 in one, and as τ_k in the other.

So we can construct π_{m+1} from π_m by either concatenation $\pi_m\tau$, or by identifying the first letter of the next occurrence of τ with the last letter of π_m. Note that the word with the smallest size that is order-isomorphic to τ and starts with $\tau_k = \tau_1 + j$ is $\tau*$, defined as before by

$$\tau_i^* = \begin{cases} \tau_i & \text{if } \tau_i < \tau_1 \\ \tau_i + j & \text{otherwise.} \end{cases}$$

Thus, either $\pi_{m+1} = \pi_m\tau$ or $\pi_{m+1} = \langle \pi_m, \tau \rangle$. Moreover, which of the two cases holds only depends on the last letter of π_m, since this determines which of the two compositions is smaller. Indeed, let t be the last letter of π_m. Then $|\pi_m\tau| = |\pi_m| + |\tau|$, whereas

$$|\langle \pi_m, \tau \rangle| = |\pi_m| + |\tau| - \tau_1 + (t - \tau_1)(s-1).$$

So we choose $\pi_{m+1} = \langle \pi_m, \tau \rangle$ if $(t-\tau_1)(s-1) \le \tau_1$, i.e. if $t \le \frac{s\tau_1}{s-1}$, and $\pi_{m+1} = \pi_m\tau$ otherwise.

By induction, we thus see that π_m has the form $(\tau^{\langle p \rangle})^{\lfloor \frac{m}{p} \rfloor}\tau^{\langle q \rangle}$ for some p and q. We also know that $\tau^{\langle p \rangle}$ is the smallest glued power of τ whose last letter is greater than $\frac{s\tau_1}{s-1}$. But the last letter of $\tau^{\langle p\rangle}$ is $\tau_1 + p(\tau_k - \tau_1)$, so p is the smallest integer such that

$$p(\tau_k - \tau_1) > \frac{s\tau_1}{s-1} - \tau_1 = \frac{\tau_1}{s-1}.$$

Now the asymptotic packing density in π_m is clearly seen to equal that in $\tau^{\langle p\rangle}$, which is

$$\frac{p}{|\tau^{\langle p\rangle}|} = \frac{p}{p|\tau| - (p-1)\tau_1 + (s-1)(\tau_k - \tau_1)\binom{p}{2}}.$$

□

Another class of subword patterns that is easy to study are the strictly increasing ones.

Proposition 4.2. *Let $\tau_k = 12\cdots k$. Then $\delta(\tau_k) = \frac{1}{2k-1}$.*

Proof. The letter 1 can only appear as the first letter in an occurrence of τ_k, so every time the letter 1 is repeated in π, that splits the occurrences of τ_k in π into "before" and "after". For π to be optimal, both parts of π must be optimal, so we can write $\pi = \pi^1 \cdots \pi^m$ for some m, where each π^i is optimal and only contains one letter 1. Now each π^i must be strictly increasing, otherwise we could reduce letters and get another 1. So the maximal packing density of τ_k is really obtained in some τ_ℓ, $\ell \geq k$.

We will say that the price of an occurrence is the sum of the letters that we have to add to obtain that occurrence. So the price of the first occurrence of τ_k is $\binom{k+1}{2}$, and the price of the ith occurrence is $i + k - 1$ when $i \geq 2$. The optimal packing is obtained when we can no longer add an occurrence without increasing the average price.

The average price of the j first occurrences is

$$\frac{\sum_{i=1}^{k+j-1} i}{j} = \frac{\binom{k+j}{2}}{j} = \frac{(k+j)(k+j-1)}{2j}.$$

The price of the next occurrence is $j + k$, and this is more expensive than the average if $j \leq k$. Equality holds if $j = k-1$, so both τ_{2k-1} and τ_{2k-2} are optimal packings of τ_k.

The packing density of τ_k is thus

$$d(\tau_k, \tau_{2k-1}) = \frac{k}{\binom{2k}{2}} = \frac{1}{2k-1}.$$

□

Table 1: Packing densities of 3-letter subwords

τ	$\delta(\tau)$	τ	$\delta(\tau)$	τ	$\delta(\tau)$
111	1	$121, 212$	$\frac{1}{3}$	112	$\frac{1}{4}$
$123, 122$	$\frac{1}{5}$	213	$\frac{2}{11}$	132	$\frac{1}{6}$

Table 2: Packing densities of 4-letter subwords

τ	$\delta(\tau)$
1111	1
1212	$\frac{1}{3}$
$1121, 2112$	$\frac{1}{4}$
$1112, 1221, 2122$	$\frac{1}{5}$
$1122, 1213, 1231, 1312, 2132, 3123$	$\frac{1}{6}$
2113	$\frac{2}{13}$
$1234, 1123, 1132, 1222$	$\frac{1}{7}$
$1232, 1322, 2123, 2213$	$\frac{1}{8}$
1223	$\frac{2}{17}$
$1324, 2413, 3124, 3214, 1233, 1323, 1332, 2133, 2313$	$\frac{1}{9}$
$1243, 1342, 1423, 1432, 2134, 2143, 2314$	$\frac{1}{10}$

Note that every pattern with three letters is either non-overlapping, trivial or monotone. Thus we can find their packing densities using the results above, and these are found in Table 1.

We can also tabulate the packing densities of patterns with four letters, as in Table 2. To do this requires some care, because it is not true that every four letter pattern is either non-overlapping, trivial or monotone. So to add another occurrence, it does not suffice to determine “whether to glue or to start a new word”, but also “how many letters to glue”.

For a given pattern, it is an easy task to (let your computer) compare the corresponding packing densities, but a general formula seems harder to get.

For the four letter patterns (see Table 2), however, it can be verified case by case that the optimal packing is always “glued as tightly as possible”, meaning that when packing a k-overlapping pattern, we always identify either k letters or none.

5. Generalized patterns

We will now consider packing generalized patterns, defined in [3]. A generalized pattern is a word τ with dashes – between some letters. An occurrence of τ in π is an occurrence of τ as an ordinary pattern, where the letters corresponding to τ_i and τ_j must be consecutive, unless there is a dash between τ_i and τ_j. For example, the subsequence 243 in 2413 is an occurrence of 13–2, but is not an occurrence of 1–32. We call the dash-free parts of the pattern "blocks" or "subwords" interchangeably.

An occurrence in π of a generalized pattern τ with i dashes is thus given by $i+1$ disjoint subwords, and specifically by $i+1$ indices in π. So it makes sense to define the packing density of τ in π as

$$d(\tau,\pi) = \frac{\nu(\tau,\pi)}{\binom{n}{i+1}}.$$

As usual, we define

$$\delta(\tau) = \lim_{n\to\infty} \max\{d(\tau_\pi) \mid \pi \in C_n\}.$$

This definition agrees with that of δ_w when τ has no dash, and with our first definition if τ is a classical pattern (so has dashes between every letter). As further evidence that the definition is a reasonable one, we prove that the packing density is always strictly positive.

Theorem 5.1. *Let τ be a generalized pattern with i dashes and $|\tau| = k$. Then $k^{-(i+1)} \le \delta(\tau) \le 1$.*

Proof. An occurrence of τ in π is given by the indices where the $i+1$ subwords start. Since $\pi \in C_n$ has at most n letters, we have $\nu(\tau,\pi) \le \binom{n}{i+1}$, so $d(\tau,\pi) \le 1$ for every composition π. Thus $\delta(\tau) \le 1$.

Now let τ^n be the word constructed by concatenating n copies of τ, and deleting all dashes. Then $\tau^n \in C_{nk}$. For every choice of $1 \le a_1 \le \cdots \le a_{i+1} \le n$, we get an occurrence of τ in τ^n, by picking the j:th subword in τ from the a_jth copy in τ^n. But there are $\binom{n+i}{i+1}$ such sequences $\{a_i\}$. Thus, we have

$$\nu(\tau,\tau^n) \ge \binom{n+i}{i+1},$$

which gives

$$\delta(\tau) \ge \lim_{n\to\infty} d(\tau,\tau^n) \ge \lim_{n\to\infty} \frac{\binom{n+i}{i+1}}{\binom{nk}{i+1}} = k^{-(i+1)}. \qquad \square$$

In this section, we mainly consider generalized patterns on three letters. We have already dealt with classical patterns and with subwords, so we will now try to find the packing densities of patterns with one dash. In some cases, we get results for more general classes of patterns for free. As before, we will see that the task is fairly easy in the binary case, and much harder if the pattern is a permutation.

By reflection invariance, we can restrict attention to patterns *ab–c*.

Proposition 5.2. *A generalized pattern τ has packing density* 1 *if and only if all its letters are equal.*

Proof. We let k denote the length of the pattern, and i the number of dashes. If $\sigma = 1^{a_1}–1^{a_2}–\cdots–1^{a_{i+1}}$, where $\sum a_j = k$, then an occurrence of σ in 1^n is given by the spots b_j where the subwords start, with the restriction that the subwords do not overlap, so $b_{j+1} - b_j \geq a_j$ for each j. Such a choice can be made in $\binom{n-k}{i+1}$ ways, so

$$d(\sigma, 1^n) = \frac{\binom{n-k}{i+1}}{\binom{n}{i+1}} \longrightarrow 1,$$

as $n \to \infty$. Thus, $\delta(\tau) = 1$.

Now let τ be any pattern with $\delta(\tau) = 1$, and let a be the number of blocks in τ that are not constantly 1. We want to show that $a = 0$. Let π_n be a τ-optimal composition of n, that has a_n letters that are not 1. So π_n has length at most $n - a_n$, whence $\nu(\tau, \pi_n) \leq \binom{n-a_n}{i+1}$. We get

$$\frac{\binom{n-a_n}{i+1}}{\binom{n}{i+1}} \geq d(\tau, \pi_n) \longrightarrow 1,$$

so $\frac{n-a_n}{n} \to 1$, $\frac{a_n}{n} \to 0$.

On the other hand, non-minimal elements in τ must correspond to non-minimal elements in π_n, so

$$\begin{aligned} d(\tau, \pi_n) &= \frac{\nu(\tau, \pi_n)}{\binom{n}{i+1}} \leq \frac{\binom{a_n}{a}\binom{n-a_n}{i+1-a}}{\binom{n}{i+1}} \\ &\leq \frac{a_n^a (n-a_n)^{i+1-a}\binom{i+1}{a}}{n^{i+1}} \leq \binom{i+1}{a}\left(\frac{a_n}{n}\right)^a \longrightarrow 0, \end{aligned}$$

if $a \neq 0$. This is a contradiction, so $\delta(\tau) = 1$ implies that $a = 0$, which means that τ is constant. $\square$

The next thing to consider would be monotone patterns whose every dash-free subword is constant. But even this seemingly easy case is surprisingly hard. Focusing on binary patterns, we will instead continue with the case $\tau = 1^{x+1}\text{–}2^{y+1}$.

Proposition 5.3. *Let $\tau = 1^{x+1}\text{–}2^{y+1}$, where $x, y \geq 0$. Then $\delta(\tau) = \frac{1}{4}$.*

Proof. Let π be any composition. Sorting the letters in π increasingly does not reduce the number of occurrences of τ. Hence, a τ-optimal composition has the form $\pi = 1^{a_1} \cdots k^{a_k}$, and summing over the letters in the occurrence, we get

$$\nu(\tau, \pi) = \sum_{1 \leq i < j \leq k} (a_i - x)^+ (a_j - y)^+,$$

where we use the notation $x^+ = \max(x, 0)$.

If $k \geq 3$, construct $\pi' = 1^{a_1+a_k} 2^{a_2} \cdots (k-2)^{a_{k-2}} (k-1)^{a_{k-1}+a_k}$, by replacing each letter k by $k-1$ and inserting a_k letters 1 in the beginning. We may certainly assume that $a_k \geq y$, otherwise the letters k would be of no use in π. Then $(a_{k-1} - x)^+ \cdot (a_k - y)$ occurrences have been destroyed when going from π to π', while all others are naturally preserved. But we have also added $a_k(a_k + a_{k-1} - y) \geq a_k a_{k-1}$ occurrences of the form $1^{x+1}(k-1)^{y+1}$, so we have added more occurrences than we have destroyed. This shows that π was not optimal.

So an optimal composition has the form $\pi = 1^{n-2a} 2^a$, and

$$\nu(\tau, \pi) = (n - 2a - x)(a - y) = -2a^2 + (2y + n - x)a + xy.$$

For fixed n, x and y, this is maximized when $a = \frac{2y+n-x}{4} \approx \frac{n}{4}$ for big n. Then

$$d(\tau, \pi) \approx \frac{\frac{n}{2} \cdot \frac{n}{4}}{\binom{n}{2}} \longrightarrow \frac{1}{4},$$

when $n \to \infty$. So $\delta(\tau) = \frac{1}{4}$, independently of x and y. $\square$

In the next two propositions, we take a joint look at the remaining binary three letter patterns, with one dash. These patterns have the common property that they have a subword using both letters 1 and 2.

Proposition 5.4. *The patterns* 12–2*,* 21–2*,* 2–12 *and* 2–21 *all have packing density* 1/8*.*

Proof. By reversal we may consider only the first two cases. Write $\tau = w\text{–}2$, where w is 12 or 21. Let π be any composition, and let π' be the composition obtained by replacing every letter in π that is ≥ 3, with the two-letter subword w of τ. Then every occurrence of τ in π gives an occurrence in π'. Indeed, subsequences w–2 are preserved, while every pair 3–3 in π gives an occurrence w–2 as a subsequence of w–w in π'.

This shows that a τ-maximal composition uses only the letters 1 and 2. Moreover, every letter 1 must be in a subword w, i.e preceded (if $w = 21$) or succeeded (if $w = 12$) by a letter 2. Hence we can write the optimum as a concatenation of words w and '2'. An occurrence of τ comes from a subsequence w–2 or w–w, so the number of occurrences increases if every w is moved to the beginning of the word.

So an optimal composition has the form $\pi = w^x 2^y$, with $3x + 2y = n$ and

$$\nu(\tau, \pi) = \binom{x}{2} + xy \approx \frac{x^2}{2} + \frac{x(n-3x)}{2} = \frac{x(n-2x)}{2}.$$

This is maximized over the reals (and hence asymptotically) when $x = n/4$, so

$$\delta(12\text{–}2) = \delta(21\text{–}2) = \lim_{n\to\infty} \frac{\frac{n}{4} \cdot \frac{n}{2}}{2\binom{n}{2}} = \frac{1}{8}.$$

$\square$

When the isolated letter in the pattern is 1, we need a slight twist of the argument, but the big picture is the same.

Proposition 5.5. *The patterns* 12–1, 21–1, 1–12 *and* 1–21 *all have packing density* 1/5.

Proof. By reversal we may consider only the first two cases. Write $\tau = w$–1, where w is 12 or 21. Let π be any composition. If a letter 2 occurs in π, but not preceded (if $w = 12$) or succeeded (if $w = 21$) by a letter 1, then it can only be used as a '1' in an occurrence of τ. So we can change every such letter into a letter 1, without destroying occurrences.

Now, we can assume that every letter 2 is preceded (succeeded) by a letter 1 in π. But this allows us to construct π' by replacing every letter ≥ 3 with the word w, without destroying occurrences. Indeed, every pair of letters ≥ 3 in π will give an occurrence of τ in π'. Also, every subsequence 13–1 (31–1) in π, gives a subsequence w–1 in π'. Finally, for each subsequence 23–2 (32–2) in π, the second letter 2 comes with a letter 1 next to it, which gives an occurrence of w–1 in π'.

This shows that a τ-maximal composition uses only the letters 1 and 2, and that every letter 2 is contained in a subword w. An occurrence of τ comes from a subsequence w–1 or w–w, so the number of occurrences increases if every w is moved to the beginning of the word.

So an optimal composition has the form $\pi = w^x 1^y$, with $3x + y = n$ and

$$\nu(\tau, \pi) = \binom{x}{2} + xy \approx \frac{x^2}{2} + x(n - 3x) = \frac{x(2n - 5x)}{2}.$$

This is maximized over the reals (and hence asymptotically) when $x = n/5$, so

$$\delta(12\text{–}1) = \delta(21\text{–}1) = \lim_{n\to\infty} \frac{\frac{n}{5} \cdot n}{2\binom{n}{2}} = \frac{1}{5}.$$

□

6. Patterns of the form xy–z where xyz is a permutation pattern

This section will be a collection of exact packing densities for the six remaining three letter patterns with one dash. The patterns come naturally in pairs with the same packing densities.

Proposition 6.1. *We have* $\delta(32\text{–}1) = \delta(23\text{–}1) = \frac{4}{35}$.

Proof. We start by considering the pattern 32–1. Assume that π is an optimal composition for 32–1. If there is i with $\pi_i = \pi_{i+1} > 1$ then we can remove the letter π_i and put π_i ones at the end of π, without decreasing the occurrence number. Also any letter 1 can be moved to the end of the word without destroying occurrences. Thus we can write $\pi = w^1 w^2 \dots w^s 111\dots 1$ where w^j is strictly decreasing, and the first letter in w^{i+1} is strictly larger than the last letter in w^i. (Observe that the upper indices, for the time being, do not denote exponents.)

If there is any letter in w^j, say w^j_m, satisfying $w^i_a > w^j_m > w^i_{a+1}$ where $i < j$ then moving the letter w^j_m to the position between w^i_a and w^i_{a+1} only increases the number of occurrences of 32–1. If $w^j_m > w^i_1$, the same holds when moving w^j_m to the beginning of w_i.

Since the optimal composition can be assumed to be reduced, this means that we can write $\pi = (a_1(a_1-1)\cdots 2)(a_2(a_2-1)\cdots 2)\cdots(a_s(a_s-1)\cdots 2)1^t$ where $a_1 > a_2 > \cdots > a_s \geq 3$. (We are now back to using exponents, so 1^t is the word with t letters 1.)

Now, we want to bound the size of the letters of an optimal composition. For this purpose, let $k \geq 5$, and suppose

$$\hat{d}(32\text{–}1, \sigma) := \frac{\nu(32\text{–}1, \sigma)}{2|\sigma|^2} \leq \frac{4}{35} \tag{2}$$

whenever σ uses only the letters $1,\ldots,k-1$. We want to prove that the same holds if we allow the letter k. Note that this would imply that if we can get density above $4/35$, we can do so only with the letters $1,\ldots,4$, since $\hat{d}(32\text{–}1,\sigma)\approx d(32\text{–}1,\sigma)$ for large $|\sigma|$.

Let $\pi=(k\cdots 2)^a\sigma$, where σ uses no letters $1,\ldots,k-1$. Let $\sigma'=(k-1\cdots 2)^a\sigma$. Then $|\pi|=|\sigma'|+ka$ and

$$\nu(32\text{–}1,\pi)=\nu(32\text{–}1,\sigma')+\binom{a}{2}(k-3)+ax,$$

where x is the number of letters $\le k-2$ in σ. Assume for a contradiction that

$$\frac{4}{35}\le\hat{d}(32\text{–}1,\pi)=\frac{2\nu(32\text{–}1,\pi)}{|\pi|^2}.$$

That would imply that

$$2(|\sigma'|+ka)^2=2|\pi|^2\le 35(\nu(32\text{–}1,\sigma')+\binom{a}{2}(k-3)+ax).$$

But (2) says that

$$2|\sigma'|^2\ge 35\nu(32\text{–}1,\sigma').$$

We thus would get

$$2k^2a^2+4ka|\sigma'|\le 35\binom{a}{2}(k-3)+35ax.$$

Division by a yields

$$2k\left[a+4|\sigma'|\right]\le\frac{35(a+1)}{2}(k-3)+35x.$$

Observing that $|\sigma'|\ge x+\binom{k+1}{2}-1$, and using $k\ge 5$, we obtain a contradiction.

The above shows that an optimal composition is $(432)^a(32)^b1^c$, where $9a+5b+c=N$ and $\nu=\binom{a}{2}+ab+2ac+bc$. Now we let $(\alpha,\beta,\gamma)=\frac{1}{N}(a,b,c)$, and relax the problem to the reals, with the restriction $(\alpha,\beta,\gamma)\ge\mathbf{0}$. We see that an optimum is obtained at $(\alpha,\beta,\gamma)=\frac{1}{35}(2,0,17)$.

Thus $\pi_N=(432)^{\frac{2N}{35}}1^{\frac{17N}{35}}$ is a 32–1-optimal composition. It follows that

$$\delta(32\text{–}1)=\lim_{n\to\infty}d(32\text{–}1,\pi_n)=\lim_{n\to\infty}\frac{\frac{2n}{35}^2+4\frac{2n}{35}\frac{17n}{35}}{n^2}=4/35.$$

We now turn to the pattern 23–1. Let $\tau = (234)^a 1^b$, and note that $\nu(23\text{–}1, \tau) = 2ab + \binom{a}{2}$. If we let $a = \frac{2n}{35}$ and $b = \frac{17n}{35}$, we get $d(23\text{–}1, \tau) \to \frac{4}{35}$ as $n \to \infty$. Thus we have shown that $\delta(23\text{–}1) \geq \frac{4}{35}$.

On the other hand, suppose $\delta(23\text{–}1) > \frac{4}{35}$, and let τ_n be a sequence of compositions with $|\tau_n| = n$ and $d(23\text{–}1, \tau_n) \to \delta(23\text{–}1)$. Decompose $\tau = w^1 \dots w^s$ into maximal non-decreasing subwords as before. Consider the composition $\tau' = \bar{w^1} \dots \bar{w^s}$ obtained by reversing each of the maximal non-decreasing subwords, into (not necessarily maximal) non-increasing ones. Any occurrence of 23–1 in τ induces an occurrence of 32–1 in τ', so $\nu(32\text{–}1, \tau') \geq \nu(23\text{–}1, \tau)$.

This is a contradiction, knowing that $\delta(32\text{–}1) = \frac{4}{35}$. Thus we have $\delta(23\text{–}1) = \frac{4}{35}$. □

Proposition 6.2. *We have* $\delta(13\text{–}2) = \delta(31\text{–}2) = \frac{2}{31}$.

Proof. Let $\pi = \pi_1 \cdots \pi_m$ be any 13–2-optimal composition for $n \geq 6$, and assume π has an increasing subword $\pi_{i-1} \leq \pi_i \leq \pi_{i+1}$, where at least one of the inequalities is strict. Then construct $\pi' = \pi_1 \cdots \pi_{i-1}\pi_{i+1} \cdots \pi_m \pi_i$, by moving the middle letter of the subword to the end of π. If $j \geq i$ and either of $\pi_{i-1}\pi_i\pi_j$ and $\pi_i\pi_{i+1}\pi_j$ is an occurrence of 13–2 in π, and then $\pi_{i-1}\pi_{i+1}\pi_j$ is an occurrence in π'. So $\nu(13\text{–}2, \pi) \leq \nu(13\text{–}2, \pi')$, so π' is also optimal.

Similarly, if $\pi_{i-1} \geq \pi_i \geq \pi_{i+1}$, then π_i can never be used in the subword 13 of an occurrence of 13–2, so it can be moved to the end of the word.

Thus we can assume that π has the form $\pi = x_1 y_1 x_2 y_2 \cdots x_d y_d \rho$ where $x_i + 1 < y_i > x_{i+1}$ and ρ is constant. Also we can assume that $1 = x_1 \leq x_2 \leq \cdots \leq x_d$ and $y_1 \geq y_2 \geq \cdots \geq y_d$.

Assume $x_1 = \cdots = x_l = 1 \neq x_{l+1}$. Then we can replace x_j by $x_j - 1$ and replace y_j by $y_j - 1$ whenever $l < j \leq d$. This allows us to add $d - l$ letters 2 in the end of the word, and this procedure does not decrease the number of occurrences, so we can assume that $x_i = 1$ for $1 \leq i \leq d$. This means that $\pi = 1y_1 1y_2 \cdots 1y_d\rho$, where $y_i \geq 3$ and ρ is constant. Reducing letters without destroying order type, we can write $\pi = (1k)^{a_k} \cdots (13)^{a_3} 2^b$, with $a_i > 0$ for $3 \leq i \leq k$.

Let $\hat{\pi}$ be the word where we have replaced $(1i)$ by $i + 1$. An occurrence of 13–2 in π corresponds exactly to an occurrence of 2–1 in $\hat{\pi}$, and $\hat{\pi}$ is 2–1-optimal among compositions on the letters in $\mathbb{N} \setminus \{1, 3\}$.

The same arguments as above, changing increasing subsequences to decreasing ones, and with $(i1)$ instead of $(1i)$, show that a 31–2-optimal composition has the form $\pi = (k1)^{a_k} \cdots (31)^{a_3} 2^b$. Replacing $(i1)$ by $i + 1$, we see that this is also equivalent to the same $\hat{\pi}$.

So we need to maximize the function

$$\nu(2\text{–}1,\hat{\pi}) = \sum_{i\neq j,\ i,j\notin\{1,3\}} b_i b_j,$$

subject to $\sum_j j b_j = n$.

First, we consider $2b_2 + 4b_4 + 5b_5 + 6b_6 = p \leq n$ fixed. Dividing by p and relaxing the problem to the reals, we want to optimize $\alpha\beta + \alpha\gamma + \alpha\delta + \beta\gamma + \beta\delta + \gamma\delta$ subject to $2\alpha + 4\beta + 5\gamma + 6\delta = 1$. Differentiation yields no interior extreme point, so in an optimal solution $b_6 = 0$, and thus of course $b_j = 0$ for all $j \geq 6$, so $p = n$.

We want to optimize the function $\alpha\beta + \alpha\gamma + \beta\gamma$ subject to $2\alpha + 4\beta + 5\gamma = 1$. Differentiation gives the maximum $(\alpha, \beta, \gamma) = \frac{1}{31}(7, 3, 1)$.

So a 31–2-optimal composition is $(41)^a(31)^{3a}2^{7a}$, and a 13–2-optimal composition is $(14)^a(13)^{3a}2^{7a}$, for any a. These have packing densities

$$\delta(13\text{–}2) = \delta(31\text{–}2) = \frac{a^2(1\cdot 3 + 1\cdot 7 + 3\cdot 7)}{\binom{31a}{2}} \to \frac{2}{31},$$

as required. □

Proposition 6.3. *We have $\delta(21\text{–}3) = \frac{1}{18}$.*

Proof. First, let $\pi_k = (21)^k 3^k$. Note that $\nu(21\text{–}3, \pi_k) = k^2$ and $|\pi_k| = 6k$, so $d(21\text{–}3, \pi_k) \to \frac{1}{18}$. This shows that $\delta(21\text{–}3) \geq \frac{1}{18}$. Since $\binom{|\pi|}{2} \approx \frac{|\pi|^2}{2}$, we are done if we can show that

$$\frac{2\nu(21\text{–}3,\sigma)}{|\sigma|^2} \leq \frac{1}{18} \tag{3}$$

holds for every composition σ.

We proceed by restricting the problem to a finite alphabet. Assume that $k \geq 6$ is the largest letter in τ, and that (3) holds for any composition σ using only letters $1, \ldots, k-1$. We will conclude that (3) holds for τ as well.

Since k is the largest letter in τ, it can only be used as the last letter in an occurrence of 21–3. We may therefore assume that all the letters k come in the end of τ, so $\tau = \sigma k^p$, where σ only uses the letters $1, \ldots, k-1$, so (3) holds for σ.

Assume for a contradiction that

$$\frac{2\nu(21\text{–}3,\tau)}{|\tau|^2} > \frac{1}{18}. \tag{4}$$

Observe that $|\tau| = |\sigma| + kp \geq |\sigma| + 6p$ and

$$\nu(21\text{–}3, \tau) = \nu(21\text{–}3, \sigma) + p\nu(21, \sigma),$$

so (4) implies that

$$36\left(\nu(21\text{–}3, \sigma) + p\nu(21, \sigma)\right) > |\tau|^2 \geq |\sigma|^2 + 36p^2 + 12p|\sigma|.$$

But (3) holds for σ, so $36\nu(21\text{–}3, \sigma) \leq |\sigma|^2$, which gives the inequality

$$36p\nu(21, \sigma) > 36p^2 + 12p|\sigma|.$$

But the subword pattern 21 has packing density 1/3, so $36\nu(21, \sigma) \leq 12|\sigma|$, which gives a contradiction.

We have now shown that if there is a a composition π that violates (3), then π can be chosen to use only the letters $1, \ldots, 5$.

Decompose $\pi = \alpha_1 \cdots \alpha_s$ into maximal non-increasing subwords. By letter reduction, we can assume $\alpha_1 = 21$. Moving the largest letters to the right does not decrease the occurrence number. Hence we may assume that α_s is constant, and all the other blocks are strictly decreasing.

Suppose that p is the greatest letter in the word $\alpha_1 \cdots \alpha_j$. If α_{j+1} has some letter $\geq p + 2$, then we can replace this letter by $(p + 1, 1)$ without deleting any occurrences of 21–3. So we may assume that the maximal letter is increased by at most one for every subword α.

Any letter 2 must be followed by a letter 1, otherwise we could reduce it, since it can only be used as 1 or 2 in an occurrence of 21–3. Finally, for every block, move its first letter k to the right-most block that has the first letter $\leq k - 1$. Move any letter that is not the largest in its block to the left-most block in which it is still not largest. These operations do not destroy occurrences.

This leaves us with three different possible forms of an optimal composition:

- $\pi_3 = (21)^a 3^b$, with $|\pi_3| = 3a + 3b$ and $\nu = ab$.
- $\pi_4 = (21)^a (321)^b (31)^c 4^d$, with $|\pi_4| = 3a + 6b + 4c + 4d$ and $\nu = ab + ac + ad + \binom{b}{2} + bc + 2bd + cd$.
- $\pi_5 = (21)^a (321)^b (31)^c (4321)^d (41)^e 5^f$ with $|\pi_5| = 3a + 6b + 4c + 10d + 5e + 5f$ and $\nu = a(b+c+2d+e+f)+b(c+3d+2e+2f)+c(d+e+f)+d(2e+3f)+ef$.

Relaxing the variables to the reals, dividing by $\frac{|\pi|^2}{2}$ to get the asymptotic density, and differentiating, we obtain the optimal parameters in each of the cases. The overall maximum is obtained by π_3 when $a = b = n/6$, in which case $d(21\text{–}3, \pi_3) = 1/18$. Hence $\delta(21\text{–}3) = 1/18$. $\square$

Proposition 6.4. *We have* $\delta(12\text{–}3) = \frac{1}{18}$.

Proof. By copying the arguments from the case 32–1, but looking at increasing subwords rather than decreasing ones, we see that an optimal packing has the form

$$\pi = (12)^{a_2}(123)^{a_3} \cdots (1 \cdots k-1)^{a_{k-1}} k^{a_k},$$

for some k, and $a_i \geq 0$. Our next step is to bound the size of the alphabet. If $\frac{2\nu(12\text{–}3,\sigma)}{|\sigma|^2} \leq \frac{1}{18}$ holds whenever σ uses only the letters $1, \ldots, k-1$, we want to show that the same inequality holds for τ, which is allowed to use the letter k as well. If $k \geq 6$, this is proven by the exact same argument as for the case 21–3.

Thus an optimal composition has either of the forms

- $\pi_3 = (12)^a 3^b$, with $|\pi_3| = 3a + 3b$ and $\nu = ab$.
- $\pi_4 = (12)^a (123)^b 4^c$, with $|\pi_4| = 3a + 6b + 4c$ and $\nu = ab + ac + \binom{b+1}{2} + 2bc$.
- $\pi_5 = (12)^a (123)^b (1234)^c 5^d$ with $|\pi_5| = 3a + 6b + 10c + 5d$ and $\nu = a(b + 2c + d) + \binom{b+1}{2} + b(3c + 2d) + 3\binom{c+1}{2} + 3cd$.

Relaxing the variables to the reals, dividing by $\frac{|\pi|^2}{2}$ to get the asymptotic density, and differentiating, we obtain the optimal parameters in each of the cases. The overall maximum is obtained by π_3 when $a = b = n/6$, in which case $d(12\text{–}3, \pi_3) = 1/18$. Hence $\delta(12\text{–}3) = 1/18$. $\square$

Thus we have found the packing densities of all three letter patterns with one dash.

7. Open problems

Quite disturbingly, the monotone (classical) pattern $1^{a_1}2^{a_2}\cdots k^{a_k}$ remain unsolved when $k \geq 3$, even if we would let all the a_i equal 1. The same is true for the generalized pattern $1^{a_1}\text{–}2^{a_2}\text{–}\cdots\text{–}k^{a_k}$. In both these cases, it is clear that an optimal composition has the form $1^{x_1}\cdots \ell^{x_\ell}$ but to find an optimum we must optimize both ℓ and $\{x_i\}$ simultaneously.

One could consider the following generalized problem: for each letter i, assign a cost c_i (subject to some technical constraints). Define $\|\pi\|$ to be the sum of the costs for the letters in π, and let the packing density of τ in π be $\frac{\nu(\tau,\pi)}{\binom{\|\pi\|}{i+1}}$, where τ has i dashes.

Then pattern packing in words is the special case where all $c_i = 1$, and pattern packing in compositions is the case $c_i = i$. Some of our methods rely on that the c_i are ordered in the same way as the letters i. Maybe looking at the generalized version might give some better insight in the structural essence of pattern packing problems.

Acknowledgements

The authors wish to thank an anonymous referee for substantial contributions on the proofs of Lemma 3.3 and Proposition 3.9.

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** (2002) #R5. MR1887086

[2] R.W. Barton, Packing densities of patterns, *Electron. J. Combin.* **11** (2004) #R80. MR2114184

[3] E. Babson and E. Steingrimsson, Generalized permutation patterns and a classification of the Mahonian statistics, *Semi. Loth. de Combinatoire*, **B44b** (2000). MR1758852

[4] M. Bóna, Combinatorics of Permutations, CRC Press, Boca Raton, 2004. MR2078910

[5] M. Bóna, B.E. Sagan and V.R. Vatter, Pattern frequency sequences and internal zeros, *Adv. in Appl. Math.* **28** (2002) 395–420. MR1900002

[6] A. Burstein, P. Hästö and T. Mansour, Packing patterns into words, *Electron. J. Combin.* **9**(2) (2002–2003) #R20. MR2028289

[7] A. Burstein and P. Hästö, Packing sets of patterns, *European J. Combin.* **31**(1) (2010) 241–253. MR2552604

[8] P.A. Hästö, The packing density of other layered permutations, *Electron. J. Combin.* **9**(2) (2002–2003) #R1. MR2028271

[9] S. Heubach and T. Mansour, Combinatorics of Compositions and Words, *Discrete Mathematics and its Applications* (Boca Raton), CRC Press, Boca Raton, FL, 2010. MR2531482

[10] M. Hildebrandt, B. Sagan and V. Vatter, Bounding quantities related to the packing density of $1(m+1)m\cdots 2$, *Adv. Appl. Math.* **33** (2004) 633–653. MR2081046

[11] A. Price, Packing densities of layered patterns, Ph.D. thesis, University of Pennsylvania, 1997. MR2695616

[12] W. Stromquist, Packing layered posets into posets, manuscript, unpublished, 1993.

[13] D. Warren, Optimal packing behavior of some two-block patterns, *Ann. Comb.* **8** (3) (2004) 355–367. MR2161643

[14] D. Warren, A mild generalization of two results of Hästö, preprint, 2005.

[15] D. Warren, Packing densities of more 2-block patterns, *Adv. Appl. Math.* **36** (2) (2006) 202–211. MR2199989

Ragnar Freij
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
S-412 96 Göteborg
Sweden
E-mail address: freij@chalmers.se

Toufik Mansour
Department of Mathematics
University of Haifa
31905 Haifa
Israel
E-mail address: toufik@math.haifa.ac.il

Received January 13, 2010