Davenport–Schinzel Trees*

Pavel Valtr

Department of Applied Mathematics, Charles University

and

Institute for Theoretical Computer Science (ITI)†, Charles University

Malostranské nám. 25, 118 00 Praha 1, Czech Republic

valtr@kam.mff.cuni.cz

Abstract

In this paper we study labeled–tree analogues of (generalized) Davenport–Schinzel sequences.

We say that two sequences $a_1 \dots a_k$, $b_1 \dots, b_k$ of equal length k are isomorphic, if $a_i = a_j$ iff $b_i = b_j$ (for all i, j). For example, the sequences 11232, 33141 are isomorphic. We investigate the maximum size of a labeled (rooted) tree with each vertex labeled by one of n labels in such a way that, besides some technical conditions, the sequence of labels along any path (starting from the root) contains no subsequence isomorphic to a fixed "forbidden" sequence u.

We study two models of such labeled trees. Each of the models is known to be essentially equivalent also to other models. The labeled paths in a special case of one of our models correspond to classical Davenport–Schinzel sequences.

We investigate, in particular, for which sequences u the labeled tree has at most $O(n)$ vertices. In both models, we answer this question for any forbidden sequence u over a two-element alphabet and also for a large class of other sequences u.

*This research was partially supported by Charles University grants No. 99/158 and 99/159 and by Czech Republic Grant GAČR 201/99/0242.

†Supported by project LN00A056 of The Ministry of Education of the Czech Republic.

1 Introduction

1.1 Basic Definitions and Notation

By a sequence, we always mean a finite sequence (possibly of length 0). Let u be a (finite) sequence. Then $|u|$ denotes the length of u, $L(u)$ denotes the alphabet of u, i.e., the set of elements (letters) appearing in the sequence u, and $||u|| = |L(u)|$ denotes the size of $L(u)$. For example, the sequence $u = 125625$ (for simplicity, we do not separate terms in a sequence by commas) has length $|u| = 6$, its alphabet is $L(u) = \{1, 2, 5, 6\}$, and thus $||u|| = 4$. Elements of $L(u)$ are called *letters*. A contiguous part of a sequence is called *interval*. We say that u is *k-sparse*, if no interval in u of length at most k contains two occurrences of the same letter, i.e., $|I| = ||I||$ for every interval I in u with $|I| \leq k$. Two sequences $u = a_1a_2 \dots a_k$ and $v = b_1b_2 \dots b_k$ of equal length are said to be *isomorphic*, if there is a bijection Ψ between $L(u)$ and $L(v)$ such that $b_i = \Psi(a_i)$ for each $i = 1, \dots, k$. We say that a sequence v is *u-free* if it has no subsequence isomorphic to u. Otherwise we say that *v contains u*.

A sequence $aa \dots a$ consisting of k a-occurrences is shortly denoted by a^k. In a natural way, we write $u = u_1u_2 \dots u_k$, if u is a concatenation of sequences (and/or letters) $u_1, u_2, \dots, u_k$. For a sequence u, $\overline{u}$ denotes the sequence u written in the reversed order.

1.2 Davenport–Schinzel Trees

For an integer n and a sequence u, a u-free $||u||$-sparse sequence over at most n elements (letters) is called an *(n, u)-sequence*. Let $f_u(n)$ be the maximum length of an (n, u)-sequence. For the alternating sequence $abab\dots$ of length $s+2$, $f_u(n)$ is the intensively studied function $\lambda_s(n)$ known from the theory of Davenport–Schinzel sequences (see the book [SA95] devoted to Davenport–Schinzel sequences). For other sequences u, the function $f_u(n)$ has been also studied in a series of papers (e.g., [AKV92, KV94, Kl92, Kl94, Kl95, Kl97, Kl99, V99b]). We mention some of the results on the functions $f_u(n)$ and their applications in Section 9.

In this paper we develop an idea of Klazar [Kl95] by considering another generalization of Davenport–Schinzel sequences, Davenport–Schinzel trees. We use a general notion *Davenport–Schinzel trees* to denote (rooted) trees with vertices labeled by n labels in such a way that, besides some technical conditions, the sequence of labels along any path (from the root) contains no subsequence isomorphic to a fixed sequence u. We give results about the maximum size of Davenport–Schinzel trees.

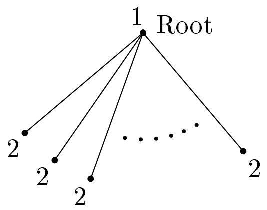
A *labeled tree* $\mathcal{T} = (T, l)$ is a tree $T = (V, E)$ with a labeling $l : V \to \mathbb{N}$, which assigns a label to each vertex. A *rooted labeled tree* is a labeled tree with one vertex specified as the *root*. The set $\{l(v) : v \in V\}$ of labels appearing in $\mathcal{T}$ is denoted by $L(\mathcal{T})$, its size by $||\mathcal{T}||$.

Let u be a sequence. We say that a rooted labeled tree $\mathcal{T} = (T, l)$ is *u-free* if, for every path $v_1 v_2 \dots v_k$ in $\mathcal{T}$ starting in the root ($v_1 =$ root), the sequence $l(v_1)l(v_2)\dots l(v_k)$ of labels along the path is u-free. We say that a labeled tree $\mathcal{T} = (T, l)$ is *$\hat{u}$-free* if, for every path $v_1 v_2 \dots v_k$ in $\mathcal{T}$ (not necessarily starting from the root), the sequence $l(v_1)l(v_2)\dots l(v_k)$ is u-free.

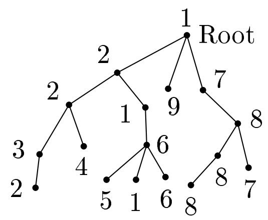
We say that a labeled tree is *k-sparse*, if no two vertices at distance less than k are labeled by the same label. Thus, a labeled tree is 2-sparse, if no two adjacent vertices are labeled by the same label.

Let n be an integer and let u be a sequence. An *(n, u)-tree* (*$(n, \hat{u})$-tree*, respectively) is a u-free ($\hat{u}$-free, respectively) $||u||$-sparse rooted labeled tree whose vertices are labeled by at most n distinct labels.

We consider the problem of estimating the maximum size of an (n, u)-tree ($(n, \hat{u})$-tree, respectively) satisfying still an additional technical condition. An additional technical condition is needed to get reasonable tree analogues of Davenport–Schinzel sequences. For example, if $u = abab$, an additional technical condition is needed to avoid arbitrarily large (n, u)-trees such as the star in Fig. 1.



$(n, abab)$-tree



spider

Figure 1: An arbitrarily large $(n, abab)$-tree (for $n \geq 2$) and an example of a spider.

A labeled rooted tree $\mathcal{T}$ is called *spider*, if for each label $i \in L(\mathcal{T})$, all vertices in $\mathcal{T}$ labeled by i lie on a path starting in the root (see Fig. 1).

If an (n, u)-tree ($(n, \hat{u})$-tree, respectively) is a spider we call it an (n, u)-*spider* ($(n, \hat{u})$-*spider*, respectively) for short.

We investigate the following two functions:

$$\sigma_u(n) = \text{ the maximum size (i.e., number of vertices) of an } (n, u)\text{-spider},$$

$$\hat{\sigma}_u(n) = \text{ the maximum size of an } (n, \hat{u})\text{-spider}.$$

Observe that $f_u(n) \leq \sigma_u(n)$ and $\hat{\sigma}_u(n) \leq \sigma_u(n)$ for each u and n. If u is isomorphic to $\overline{u}$ (u is "symmetric"), then also $f_u(n) \leq \hat{\sigma}_u(n)$.

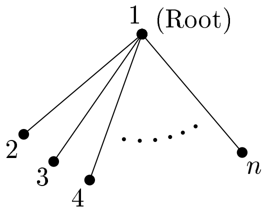


Figure 2: A labeled star giving the lower bound $\sigma_u(n) \geq n$ for any u with $|u| > \min\{||u||, 2\}$.

1.3 Main Results

The star in Fig. 2 gives $\sigma_u(n) \geq n$ (unless $|u| = ||u|| \leq 2$) and $\hat{\sigma}_u(n) \geq n$ (unless $|u| = ||u|| \leq 3$). Our paper is motivated by the following problem:

Problem 1 *Determine sequences u with $\sigma_u(n) = O(n)$ ($\hat{\sigma}_u(n) = O(n)$, respectively).*

Here are the main results of this paper giving a partial solution to Problem 1:

Theorem 1 *Let $\mathcal{L}$ denote the class of all sequences u with $\sigma_u(n) = O(n)$. Then:*

(i) The sequence a^i lies in $\mathcal{L}$ for any $i \geq 1$ (a is a letter).

(ii) (monotone property) If v contains u and $v \in \mathcal{L}$ then $u \in \mathcal{L}$.

(iii) *(letter duplication) Let u_1, u_2 be two sequences, let a be a letter, and let $u_1 a^2 u_2 \in \mathcal{L}$. Then $u_1 a^3 u_2 \in \mathcal{L}$.*

(iv) *(first–letter and last–letter duplications) Let u be a sequence and let a be a letter. Then, if $ua \in \mathcal{L}$ then $ua^2 \in \mathcal{L}$, and if $au \in \mathcal{L}$ then $a^2 u \in \mathcal{L}$.*

(v) *(letter insertions) Let $u = au_1 a^2 u_2$ be a sequence in $\mathcal{L}$, where u_1, u_2 are sequences (possibly empty), a, b are letters, and $b \notin L(u)$. Then the sequence $u' = bau_1 abau_2$ also lies in $\mathcal{L}$.*

(vi) *(sequence insertion) Let u_1, u_2, w be three sequences, let a be a letter, let $u = u_1 a^2 u_2$, and let $L(u) \cap L(w) = \emptyset$. Moreover, let u, w lie in $\mathcal{L}$. Then also $u_1 awau_2$ lies in $\mathcal{L}$.*

Theorem 2

(i) *If u contains at least one of the four sequences $ababa$, aba^2b, $abcabc$, $abcacb$, then*

$$\sigma_u(n) = \Omega(n\alpha(n)).$$

(ii) *If u cannot be written as $u = vw$, $L(v) \cap L(w) = \emptyset$, so that the sequences $\overline{v}, w$ are $ababa$-, aba^2b-, $abcabc$-, and $abcacb$-free then*

$$\hat{\sigma}_u(n) = \Omega(n\alpha(n)).$$

Theorem 1 describes the class of sequences for which we are able to prove $\sigma_u(n) = O(n)$. At the end of Section 4, we derive it from more general results given in Sections 2–4. Theorem 2 is obtained from a construction given in Section 5.

We do not state an analogue of Theorem 1 for the function $\hat{\sigma}_u(n)$. The reason is that all sequences u for which we know $\hat{\sigma}_u(n) = O(n)$ are obtained by combining Theorem 1 with Theorem 6 below.

The following theorems, derived from Theorems 1 and 2, determine the validity of $\sigma_u(n) = O(n)$ and $\hat{\sigma}_u(n) = O(n)$ for a large class of sequences u.

Theorem 3 (two-letter theorem)

(i) *If u is a sequence with $||u|| \leq 2$, then $\sigma_u(n) = O(n)$ if and only if u is $ababa$-free and aba^2b-free.*

(ii) *If u is a sequence with $||u|| \le 2$, then $\hat{\sigma}_u(n) = O(n)$ if and only if u is ababa-free and ab^2a^2b-free.*

Theorem 4 *If u is abab-free, then*

$$\hat{\sigma}_u(n) \le \sigma_u(n) = O(n).$$

Theorem 5 *If $u = a_1^{i_1} a_2^{i_2} \ldots a_k^{i_k}$, where $a_1, a_2, \ldots, a_k$ are letters (not necessarily distinct), $i_1, i_k \ge 1$, and $i_2, i_3, \ldots, i_{k-1} \ge 2$, then*

$$\sigma_u(n) = O(n) \iff \hat{\sigma}_u(n) = O(n) \iff u \text{ is abab-free.}$$

Some more results derived from Theorem 1 and from our construction can be found in Paragraph 10.1 in [V99b].

1.4 Relations between $\boldsymbol{\sigma}$ and $\boldsymbol{\hat{\sigma}}$

Here is an important rule allowing to derive an upper bound on $\hat{\sigma}_u(n)$ from upper bounds on the functions $\sigma_v(n)$:

Theorem 6 (derivation rule) *For any sequence u with $|u| > ||u||$,*

$$\hat{\sigma}_u(n) = O(\min\{\sigma_{\overline{v}}(n) + \sigma_w(n) : u = vw, L(v) \cap L(w) = \emptyset\}).$$

(For the empty sequence ε and for each n, we take $\sigma_\varepsilon(n) = 0$.)

We remark that Theorem 6 holds also for sequences u with $|u| = ||u|| \notin \{2, 4\}$. However, in the case $|u| = ||u||$ we easily get $\hat{\sigma}_u(n) \le 2$ (if $|u| = ||u|| \le 3$) or $\hat{\sigma}_u(n) = n$ (if $|u| = ||u|| > 3$). Therefore any analogues of Theorem 6 are not too interesting in this case.

The bound in Theorem 6 is asymptotically best possible for "symmetric" sequences:

Theorem 7 *If u is isomorphic to $\overline{u}$ (u is "symmetric"), then*

$$\hat{\sigma}_u(n) = \Theta(\min\{\sigma_{\overline{v}}(n) + \sigma_w(n) : u = vw, L(v) \cap L(w) = \emptyset\}).$$

1.5 General Estimates

Klazar proved an almost linear upper bound on $\hat{\sigma}_u(n)$ (originally stated for a somewhat different model — see Section 9):

Theorem 8 (Klazar [Kl99]) *For any sequence u, there is a constant $c = c(u)$ such that*

$$\hat{\sigma}_u(n) = O\left(n 2^{(\alpha(n))^c}\right),$$

where $\alpha(n)$ is the functional inverse to the Ackermann function.

The definition of the Ackermann function and its functional inverse $\alpha(n)$ can be found e.g. in [Ro67, SA95, V99b]. We remark that the term $2^{(\alpha(n))^c}$ appearing in Theorem 8 is a function which grows extremely slowly to infinity. In particular, it grows to infinity slower than the functional inverse of any unbounded primitive recursive function (see [Ro67, SA95]). On the other hand, $\sigma_u(n) \geq n$ for any n and u with $|u| > \min\{||u||, 2\}$ (see Fig. 2).

Theorem 8 is improved in [V00]:

Theorem 9 (Valtr [V00]) *For any sequence u, there is a constant $c = c(u)$ such that*

$$\hat{\sigma}_u(n) \leq \sigma_u(n) = O\left(n 2^{c(\alpha(n))^{|u|+||u||-5}}\right).$$

The following proposition gives more information about the possible asymptotic behavior of the function $\sigma_u(n)$.

Proposition 10 ([V99b]) *For any sequence u, the function $\sigma_u(n)$ behaves asymptotically in one of the following three ways:*

(i) $\sigma_u(n) \leq 1$ *for all* n,

(ii) $\sigma_u(n) = \Theta(n)$, *or*

(iii) $\frac{\sigma_u(n)}{n}$ *tends very slowly to infinity as n tends to infinity (it tends to infinity slower than the functional inverse of any unbounded primitive recursive function).*

An analogous statement holds also for the function $\hat{\sigma}_u(n)$. It is not difficult to see that case (i) in Proposition 10 holds if and only if $|u| = ||u|| \leq 2$.

Once we have Proposition 10 and the upper bound in Theorem 9, it may seem almost needless to work on Problem 1. However, Davenport–Schinzel and generalized Davenport–Schinzel sequences lead to analogous extremal functions $\lambda_s(n)$ and $f_u(n)$, respectively, for which the distinction between a linear and an almost linear bound plays an important role in applications (e.g., see the book [SA95] and the survey paper [V99b]).

The functions $\sigma_u(n)$ and $\hat{\sigma}_u(n)$ might be also interesting for their unusual asymptotic behavior — we show that there are sequences u (e.g., $u = ababa$) with $\sigma_u(n) = \Omega(n\alpha(n))$. It follows that the inverse of the function $\frac{\sigma_u(n)}{n}$ is not primitive recursive. Perhaps, for some sequences u, the inverse of the function $\frac{\sigma_u(n)}{n}$ may grow even much faster than the Ackermann function.

We remark that Klazar [Kl95, Kl97, Kl99] originally considered a somewhat different model of unrooted labeled trees. It has been shown in [V00] that Klazar's model leads for any sequence u to an extremal function of order $\Theta(\hat{\sigma}_u(n))$. Thus, there is no essential difference between Klazar's model and our model of $(n, \hat{u})$-spiders. We study the functions $\sigma_u(n)$ and $\hat{\sigma}_u(n)$ since they seem to be more natural and easier to handle with than the corresponding extremal function in Klazar's model. The other models are described in Section 9 (see also [V99b, Section 8] and [V00] for more details).

1.6 More Definitions and Notation

Throughout the rest of the paper, $|\mathcal{T}|$ denotes the number of vertices in a spider $\mathcal{T}$. *i-vertices* are vertices labeled by i. The path between two vertices u, v is called *$[u, v]$-path*. If u, v are i-vertices, then we say that the $[u, v]$-path is an *i-path*.

In the pictures we always draw a spider with the root on the top. Correspondingly, if u, v are two different vertices in a spider and v lies on the $[\text{root}, u]$-path, then we say that u lies *under* v and v lies *above* u. A vertex u is *the topmost vertex* (*the bottommost vertex*, respectively) of a set $U \subset V(\mathcal{T})$ of vertices of $\mathcal{T}$, if U contains u and all other vertices of U lie under u (above u, respectively).

A spider $\mathcal{T}' = ((V', E'), l')$ is a *subspider* of a spider $\mathcal{T} = ((V, E), l)$, if $V' \subseteq V$, $E' \subseteq E$, $l' = l|_{V'}$, and the topmost (in $\mathcal{T}$) vertex of V' is the root of $\mathcal{T}'$.

If $\mathcal{T} = ((V, E), l)$ is a spider and $U \subseteq V$ is a set of vertices in $\mathcal{T}$ containing the root of $\mathcal{T}$, then the *U-reduction* of $\mathcal{T}$ is defined as the spider on the vertex set U with the root in the root of $\mathcal{T}$ and with two vertices $v_1, v_2 \in U$ connected by an edge if and only if v_1 lies above or under v_2 and the $[v_1, v_2]$-path in $\mathcal{T}$ contains no other vertices of U (see Fig. 3). If $\mathcal{T}'$ is a U-reduction of $\mathcal{T}$ for

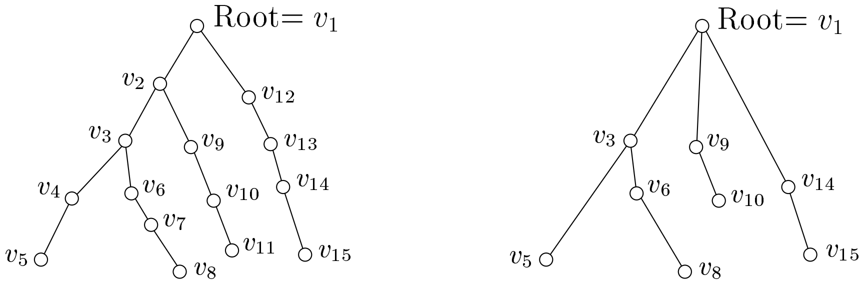


Figure 3: A rooted tree $\mathcal{T}$ and its U-reduction, where $U = \{v_1, v_3, v_5, v_6, v_8, v_9, v_{10}, v_{14}, v_{15}\}$.

some $U \subseteq V$, we say that $\mathcal{T}'$ is a *reduction* of $\mathcal{T}$.

Observation 11

(i) Any reduction of a spider is also a spider,

(ii) any reduction of a u-free (û-free, respectively) spider is also u-free (û-free, respectively).

□

A reduction $\mathcal{T}'$ of a spider $\mathcal{T} = (T, l)$ is called an *induced reduction*, if the subgraph of T induced by the vertices of $\mathcal{T}'$ is a tree. Equivalently, an induced reduction of $\mathcal{T}$ is any subspider containing the root of $\mathcal{T}$.

If $\mathcal{T}$ is an (n, u)-spider of the maximum size $\sigma_u(n)$, then we say that $\mathcal{T}$ is (n, u)*-critical.*

1.7 Organization of the Paper

The paper is organized as follows. In Section 2 we describe and prove auxiliary lemmas which are then used in Sections 3 and 4 to obtain two strong results about the asymptotic behavior of the maximum size of Davenport–Schinzel trees. The proof of Theorem 1 is derived from the results of Sections 2–4 at the end of Section 4. A lower bound construction for Davenport–Schinzel trees giving Theorem 2 is given in Section 5. Theorems 3–5 are derived from Theorems 1 and 2 in Section 6. The derivation rule (Theorem 6) and its corollary (Theorem 7) are proved in Section 7. In Section 8 we discuss conclusions of general results for short sequences. Section 9 contains concluding remarks.

2 Basic Tools

In this section we prove 8 lemmas (Lemmas 12–19) which give the easier parts (i)-(iv) of Theorem 1 and describe some useful tools and properties of the function $\sigma_u(n)$ used in the proofs of the parts (v),(vi), such as monotone and additive properties, weak super- and sub-additivity, and some others.

Lemma 12 (sparsity lemma) *Let u be a sequence, and let $l \geq ||u||$. Then there exists a positive constant $\varepsilon = \varepsilon(u, l) > 0$ such that for any u-free $||u||$-sparse spider $\mathcal{T}$ there is an l-sparse reduction of $\mathcal{T}$ with at least $\varepsilon|\mathcal{T}|$ vertices.*

Before the proof of Lemma 12, we present a simple greedy algorithm $\mathcal{B}(k)$ giving a k-sparse reduction for a given spider.

Let $k \geq 2$, and let $\mathcal{T} = (T, l)$ be a spider. The *algorithm* $\mathcal{B}(k)$ finds a k-sparse reduction $\mathcal{B}(\mathcal{T}, k)$ of $\mathcal{T}$ as follows. Let $V = \{v_1, \ldots, v_{|\mathcal{T}|}\}$ be the vertex set of $\mathcal{T}$ listed so that v_i does not lie below v_j for any $i < j$ (e.g., the vertices may be listed in the order of search–to–depth or search–to–width). We inductively define vertex sets $V_i \subseteq \{v_1, \ldots, v_i\}, i = 1, \ldots, |\mathcal{T}|$, as follows. We set $V_1 = \{v_1\} = \{\text{root}\}$. Further, for $i = 1, \ldots, |\mathcal{T}| - 1$,

$$V_{i+1} = \begin{cases} V_i \cup \{v_{i+1}\}, & \text{if the } (V_i \cup \{v_{i+1}\})\text{-reduction of } \mathcal{T} \text{ is } k\text{-sparse}, \\ V_i, & \text{otherwise.} \end{cases}$$

The $V_{|\mathcal{T}|}$-reduction of $\mathcal{T}$ is taken for $\mathcal{B}(\mathcal{T}, k)$. (see Fig. 4).

Proof of Lemma 12. We obtain Lemma 12 by running the algorithm $\mathcal{B}(l)$ on $\mathcal{T}$. The spider $\mathcal{B}(\mathcal{T}, l)$ is u-free (by Observation 11 (ii)) and l-sparse, so it remains to show the required bound on the size of $\mathcal{B}(\mathcal{T}, l)$.

For a vertex v of $\mathcal{B}(\mathcal{T}, l)$, let $\mathcal{T}_v$ be the maximal subspider of $\mathcal{T}$ rooted in v such that v is the only vertex of $\mathcal{T}_v$ lying in the vertex set of $\mathcal{B}(\mathcal{T}, l)$.

It follows from the definition of $\mathcal{B}(\mathcal{T}, l)$ that the spider $\mathcal{T}_v$ is u-free and its vertices are labeled by at most $l - 1$ labels (namely, by the labels used on the first $l - 1$ vertices on the path in $\mathcal{B}(\mathcal{T}, l)$ from v to the root of $\mathcal{B}(\mathcal{T}, l)$). Thus, $\mathcal{T}_v$ contains at most $\sigma_u(l - 1)$ vertices. Since each vertex of $\mathcal{T}$ lies in

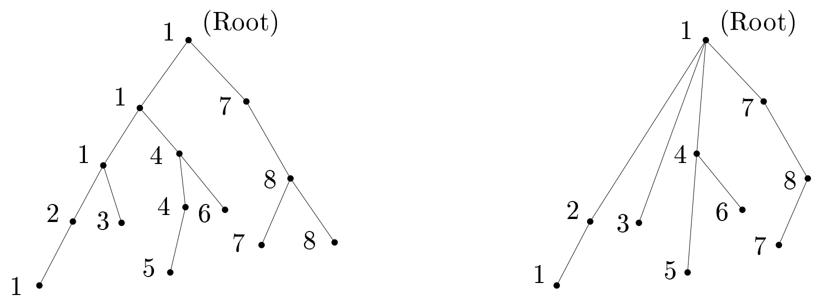


Figure 4: A spider $\mathcal{T}$ and the spider $\mathcal{B}(\mathcal{T}, 2)$ obtained from $\mathcal{T}$ by running the algorithm $\mathcal{B}(2)$.

exactly one tree $\mathcal{T}_v, v \in \mathcal{B}(\mathcal{T}, l)$, we get $|\mathcal{T}| \leq \mathcal{B}(\mathcal{T}, l) \cdot \sigma_u(l-1)$. The lemma with $\varepsilon = \frac{1}{\sigma_u(l-1)}$ follows. $\square$

Lemma 13 (reduction lemma) *For any integer $l \geq 0$ and for any $\varepsilon \in (0,1)$, there is an $\varepsilon' > 0$ with the following property. Let $\mathcal{T}$ be a $\lceil l/\varepsilon \rceil$-sparse spider and let $\mathcal{T}_1$ be a reduction of $\mathcal{T}$ with at least $\varepsilon|\mathcal{T}|$ vertices. Then $\mathcal{T}_1$ has an l-sparse reduction with at least $\varepsilon'|\mathcal{T}|$ vertices.*

Proof. Let $\mathcal{T}$ be a $\lceil l/\varepsilon \rceil$-sparse spider, let V_1 be a set of at least $\varepsilon|\mathcal{T}|$ vertices of $\mathcal{T}$, and let $\mathcal{T}_1$ be the V_1-reduction of $\mathcal{T}$. We obtain the required reduction of $\mathcal{T}_1$ by running the algorithm $\mathcal{B}(l)$ on $\mathcal{T}_1$. Let $\mathcal{T}' = \mathcal{B}(\mathcal{T}_1, l)$ be the reduction obtained by the algorithm. $\mathcal{T}'$ is an l-sparse reduction of $\mathcal{T}_1$, so it remains to show that $\mathcal{T}'$ has at least $\varepsilon'|\mathcal{T}|$ vertices.

For a vertex $u \in V_1$, let $v(u)$ be the first vertex of $\mathcal{T}'$ on the $[u, \text{root}]$-path in $\mathcal{T}$ (or in $\mathcal{T}_1$) (thus, if u lies in $\mathcal{T}'$ then $v(u) = u$). If the distance between u and $v(u)$ in $\mathcal{T}$ is less than $\lceil l/\varepsilon \rceil$, then we say that u is *of type I*; otherwise we say that u is *of type II*.

Let v be any vertex in $\mathcal{T}'$, and let U_v be the set of vertices $u \in V_1$ of type I with $v(u) = v$. By the $\lceil l/\varepsilon \rceil$-sparsity of $\mathcal{T}$, vertices of U_v are labeled by pairwise different labels. Since all vertices of U_v different from v are discarded by the algorithm $\mathcal{B}(l)$, they are labeled only by the (at most) $l-1$ labels used on the first (at most) $l-1$ vertices of the $[v, \text{root}]$-path in $\mathcal{T}'$. It follows that each U_v contains at most $1 + (l-1) = l$ vertices. Consequently, there are at most $l \cdot |\mathcal{T}'|$ vertices of type I.

To estimate the number of vertices of type II, we count in two ways the size of the set Z of pairs (u, x) of vertices such that $u \in V_1$ is a vertex of type II and x equals u or lies above u in $\mathcal{T}$ at distance less than $\lceil l/\varepsilon \rceil$.

Each vertex $u \in V_1$ of type II lies in exactly $\lceil l/\varepsilon \rceil$ pairs $(u, x) \in Z$. Thus, the number of vertices of type II is $\frac{|Z|}{\lceil l/\varepsilon \rceil}$. On the other hand, for a vertex x in $\mathcal{T}$, let $C(x)$ be the set of vertices u such that $(u, x) \in Z$. All vertices of $C(x)$ different from x lie under x at distance less than $\lceil l/\varepsilon \rceil$ from x. Therefore, since $\mathcal{T}$ is $\lceil l/\varepsilon \rceil$-sparse, $C(x)$ contains no pair of vertices labeled by the same label. Let $u \in C(x)$. By the definition of type II, the $[u, x]$-path in $\mathcal{T}$ contains no vertex of $\mathcal{T}'$. Thus, u is discarded by the algorithm $\mathcal{B}(l)$ and is therefore labeled by one of the labels used on the first (at most) $l - 1$ vertices of the $[v(u) = v(x), \mathrm{root}]$-path in $\mathcal{T}'$. It follows that the size of $C(x)$ is at most $l - 1$. Thus, $|Z| \leq (l - 1)|\mathcal{T}|$. Hence, $\mathcal{T}_1$ contains $\frac{|Z|}{\lceil l/\varepsilon \rceil} \leq \frac{(l-1)|\mathcal{T}|}{(l/\varepsilon)}$ vertices of type II.

Altogether, $\mathcal{T}_1$ contains at most

$$l \cdot |\mathcal{T}'| + \frac{(l-1)|\mathcal{T}|}{(l/\varepsilon)} = l \cdot |\mathcal{T}'| + (1 - \frac{1}{l})\varepsilon|\mathcal{T}|$$

vertices. On the other hand, it contains at least $\varepsilon|\mathcal{T}|$ vertices. Comparing these bounds, we get

$$|\mathcal{T}'| \geq \frac{\varepsilon}{l^2}|\mathcal{T}|.$$

□

Lemma 14 (letter duplication) *Let u_1, u_2 be two sequences, and let a be a letter. Then*

$$\sigma_{u_1 a^3 u_2}(n) = \Theta(\sigma_{u_1 a^2 u_2}(n)).$$

We remark that repeated applications of Lemma 14 give

$$\sigma_{u_1 a^i u_2}(n) = \Theta(\sigma_{u_1 a^2 u_2}(n)),$$

for any fixed $i \geq 3$.

Proof of Lemma 14. Clearly, $\sigma_{u_1 a^3 u_2}(n) \geq \sigma_{u_1 a^2 u_2}(n)$. On the other hand, let $\mathcal{T}$ be a $(n, u_1 a^3 u_2)$-critical spider. By the sparsity lemma (Lemma 12), $\mathcal{T}$ contains a $2||u_1 a^2 u_2||$-sparse reduction $\mathcal{T}_0$ of size $\Omega(|\mathcal{T}|)$. Let $\mathcal{T}_1$ be the reduction of $\mathcal{T}_0$ obtained from $\mathcal{T}_0$ by removing every vertex v such that the

$[\mathrm{root}, v]$-path contains an even number of $l(v)$-vertices. Thus, if there are k i-vertices in $\mathcal{T}_0$ (for some $i \in L(\mathcal{T}_0)$) then there are $\lceil \frac{k}{2} \rceil$ i-vertices in $\mathcal{T}_1$. It follows that $|\mathcal{T}_1| \geq \frac{|\mathcal{T}_0|}{2}$. It is easy to see that $\mathcal{T}_1$ is $u_1 a^2 u_2$-free (otherwise $\mathcal{T}_0$ would contain $u_1 a^3 u_2$). By the reduction lemma (Lemma 13, applied with $l = ||u_1 a^2 u_2||, \varepsilon = 1/2$), $\mathcal{T}_1$ has a $||u_1 a^2 u_2||$-sparse reduction $\mathcal{T}_2$ of size $\Omega(|\mathcal{T}_0|) = \Omega(|\mathcal{T}|) = \Omega(\sigma_{u_1 a^3 u_2}(n))$. Therefore,

$$\sigma_{u_1 a^2 u_2}(n) \geq |\mathcal{T}_2| = \Omega(\sigma_{u_1 a^3 u_2}(n)).$$

$\square$

Lemma 15 (first–letter and last–letter duplications) *Let u be a sequence and let a be a letter. Then*

$$\sigma_{a^2 u}(n) = \sigma_{au}(n) + O(n),$$

$$\sigma_{ua^2}(n) = \sigma_{ua}(n) + O(n).$$

We remark that repeated applications of Lemma 15 give

$$\sigma_{a^i u}(n) = \sigma_{au}(n) + O(n) \quad \text{and} \quad \sigma_{ua^i}(n) = \sigma_{ua}(n) + O(n),$$

for any fixed $i \geq 2$.

Proof of Lemma 15. Let $\mathcal{T} = (T, l)$ be a $(n, a^2 u)$-critical spider, and let $d = ||a^2 u||$. Let $V_1 \subseteq V$ be the set of vertices v in $\mathcal{T}$ with the property that no $l(v)$-vertex lies above v. (Thus, V_1 contains just the topmost i-vertex of $\mathcal{T}$ for each $i \in L(\mathcal{T})$.) Hence, $|V_1| = ||\mathcal{T}||$. Let $\mathcal{T}_0$ be the maximum induced reduction of $\mathcal{T}$ such that no path in $\mathcal{T}_0$ from the root to an internal vertex contains d consecutive vertices lying in $V \setminus V_1$. Let V_0 be the vertex set of $\mathcal{T}_0$, and let z be a vertex of $V \setminus (V_0 \cup V_1)$ closest to the root in $\mathcal{T}$ (thus, every vertex above z lies in $V_0 \cup V_1$). We change the label of the root to $l(z)$ and consider the $((V \setminus (V_0 \cup V_1 \cup \{z\})) \cup \{\mathrm{root}\})$-reduction $\mathcal{T}'$ of $\mathcal{T}$. $\mathcal{T}'$ is au-free — otherwise $\mathcal{T}$ would contain $a^2 u$ (the choice of V_1, z ensures this in case $a \in L(u)$, the choice of V_0, z ensures it in case $a \notin L(u)$). In the sequel we find a large d-sparse reduction of $\mathcal{T}'$.

Let V_2 be the set of vertices in $\mathcal{T}'$ of degree at most two. The subgraph of T induced by $V_2 \setminus \{\mathrm{root}\}$ is a union of disjoint paths. We denote these paths $P_1, \ldots, P_m$ so that if $i < j$ then no vertex of P_i lies under a vertex of P_j.

We now consecutively construct vertex sets

$$\{\text{root}\} = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_n = W \subseteq \bigcup_{i=1}^{m} V(P_i) \cup \{\text{root}\},$$

where $V(P_i)$ is the vertex set of P_i, such that the W_i-reduction of $\mathcal{T}$ is d-sparse for each i and the final W-reduction will contain all but at most $O(n)$ vertices of $\mathcal{T}$. We set $W_0 := \{\text{root}\}$. For $i = 1 \ldots, n$, W_i is obtained from W_{i-1} by adding some of the vertices of P_i. We will make sure that all but at most $f_{au}(3d-4)+d$ vertices of P_i are put to W_i. If $|P_i| \le f_{au}(3d-4)+d$, then we set $W_i := W_{i-1}$. Otherwise, let $v_1, \ldots, v_{f_{au}(3d-4)+d}$ be the topmost $f_{au}(3d-4)+d$ vertices in P_i (listed from top to bottom), and let L_1 be the set of $d-1$ labels labeling the bottommost $d-1$ vertices of W_{i-1} on the $[\text{root}, v_1]$-path (if fewer than $d-1$ vertices of W_{i-1} lie on the $[\text{root}, v_1]$-path, then all their labels form L_1). Further, let $L_2 = \{l(v_j) : j = f_{au}(3d-4)+2, \ldots, f_{au}(3d-4)+d\}$. Since $\mathcal{T}'$ is au-free and $\mathcal{T}$ is d-sparse, the sequence $l(v_1), \ldots, l(v_{f_{au}(3d-4)+1})$ is au-free and d-sparse. Thus, it contains at least $3d-3$ distinct labels $l(v_{j_1}), \ldots, l(v_{j_{3d-3}})$. Among these labels, we find $d-1$ labels, $l(v_{k_1}), \ldots, l(v_{k_{d-1}})$, not lying in $L_1 \cup L_2$ (note that $|L_1| \le d-1$, $|L_2| = d-1$). We set

$$W_i := W_{i-1} \cup (V(P_i) \setminus \{v_1, \ldots, v_{f_{au}(3d-4)+1}\}) \cup \{v_{k_1}, \ldots, v_{k_{d-1}}\}.$$

It is easily checked that the W_i-reduction of $\mathcal{T}$ is d-sparse (assuming the W_{i-1}-reduction was d-sparse). Thus, the $(W = W_n)$-reduction $\mathcal{T}''$ of $\mathcal{T}'$ is $(d = ||au||)$-sparse and au-free. Hence, $|\mathcal{T}''| \le \sigma_{au}(n)$. On the other hand, $\mathcal{T}''$ is constructed so that $|\mathcal{T}''| \ge |V_2| - m(f_{au}(3d-4)+d)$. In the sequel we derive from this that $|\mathcal{T}''| \ge |\mathcal{T}| - O(n)$, and the lemma will easily follow.

First, we estimate $|V_0|$. Throughout the estimate we consider the V_0-reduction $\mathcal{T}_0$. Let V_0' be the set of leaves (in $\mathcal{T}_0$) different from the root. Vertices in V_0' are labeled by distinct labels. Therefore, $|V_0'| \le n$. Let V_0'' be the set of vertices lying in V_0' or at distance at most $d-1$ above a vertex of V_0'. Clearly, $|V_0''| \le d \cdot |V_0'| \le dn$.

For any vertex $x \in V_0 \setminus V_0''$. there is an inner vertex x' below x at distance $d-1$ from x. At least one of the d vertices on the $[x, x']$-path lies in V_1 — otherwise the $[\text{root}, x']$-path would be in contradiction with the definition of $\mathcal{T}_0$. Thus, there are at least $|V_0 \setminus V_0''|$ pairs (x, x'') of vertices such that $x \in V_0 \setminus V_0''$, $x'' \in V_1$, and that either $x'' = x$ or x'' lies below x at distance at most $d-1$. On the other hand, each vertex $x'' \in V_1$ can be in at most d such pairs. It follows that $|V_0 \setminus V_0''| \le d|V_1|$. We get $|V_0| \le |V_0''| + d|V_1| \le dn + dn = 2dn$.

It is not hard to verify that $m < |V_1| + 2n = 3n$. Consequently,

$$\begin{aligned}
\sigma_{au}(n) \geq |\mathcal{T}''| = |W| &\geq |V_2| - m(f_{au}(3d-4)+d) \\
&> (|\mathcal{T}'| - (n-1)) - 3n(f_{au}(3d-4)+d) \\
&\geq (|\mathcal{T}| - |V_0| - |V_1| - (n-1)) - 3n(f_{au}(3d-4)+d) \\
&> |\mathcal{T}| - (2d+2+3(f_{au}(3d-4)+d))n \\
&= \sigma_{a^2u}(n) - O(n).
\end{aligned}$$

On the other hand,

$$\sigma_{au}(n) \leq \sigma_{a^2u}(n).$$

This finishes the proof of the first part ($\sigma_{a^2u}(n) = \sigma_{au}(n) + O(n)$).

The second part, the estimate $\sigma_{ua^2}(n) = \sigma_{ua}(n) + O(n)$, is somewhat easier to prove. Let $\mathcal{T} = (T, l)$ be a (n, ua^2)-critical spider. We define V_1 be the set of vertices v in $\mathcal{T}$ with the property that no $l(v)$-vertex lies below v. (Thus, V_1 contains just the bottommost i-vertex of $\mathcal{T}$ for each $i \in L(\mathcal{T})$.) Then the $((V \setminus V_1) \cup \{\text{root}\})$-reduction $\mathcal{T}'$ is ua-free. Analogously as above, we can show that it has a $||ua||$-sparse reduction $\mathcal{T}''$ of size $|\mathcal{T}'| - O(n) = |\mathcal{T}| - O(n) = \sigma_{ua^2}(n) - O(n)$. This, combined with the trivial estimate $\sigma_{ua}(n) \leq \sigma_{ua^2}(n)$, gives the second part of the lemma. □

The following lemma is an easy consequence of the sparsity lemma:

Lemma 16 (monotone property) *If u_1 contains u_2, then there is a constant $\varepsilon > 0$ such that, for any n,*

$$\sigma_{u_1}(n) \geq \varepsilon \cdot \sigma_{u_2}(n).$$

Proof. Let $\mathcal{T}$ be a (n, u_2)-critical spider. By the sparsity lemma (Lemma 12), $\mathcal{T}$ has a $||u_1||$-sparse reduction of size at least $\varepsilon|\mathcal{T}|$ for some $\varepsilon = \varepsilon(u_2, ||u_1||) > 0$. Thus, $\sigma_{u_1}(n) \geq \varepsilon|\mathcal{T}| = \varepsilon \cdot \sigma_{u_2}(n)$. □

Lemma 17 (weak super-additivity) *If u is a sequence with $|u| > \min\{||u||, 2\}$, then there is a positive constant $\varepsilon = \varepsilon(u) > 0$ such that*

$$\sigma_u(\sum_{i=1}^{r} n_i) \geq \varepsilon \sum_{i=1}^{r} \sigma_u(n_i)$$

holds for any positive integers $r, n_1, n_2, \ldots, n_r$.

Proof. The star in Fig. 2 gives $\sigma_u(n) \geq n$ for all n and u with $|u| > \min\{||u||, 2\}$.

We choose a (n_i, u)-critical spider $\mathcal{T}_i$ for each $i = 1, 2, \ldots, r$, so that $L(\mathcal{T}_i) \neq L(\mathcal{T}_j)$ for all $i \neq j$. For each i, let z_i be the root of $\mathcal{T}_i$. We consider the spider $\mathcal{T}$ with the root in z_1 obtained from the disjoin union of the spiders $\mathcal{T}_i$ by adding the edges $z_1 z_i, i = 2, 3, \ldots, r$ (see Fig. 5). Clearly $\mathcal{T}$ has

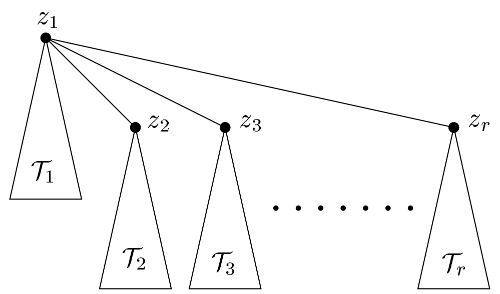


Figure 5: The spider $\mathcal{T}$ constructed in the proof of Lemma 17.

$\sum_{i=1}^{r} \sigma_u(n_i)$ vertices labeled by at most n labels and is au-free, where a is the first letter of u. Thus,

$$\sigma_{au}(n) \geq \sum_{i=1}^{r} \sigma_u(n_i),$$

where $n = n_1 + \ldots + n_r$. By Lemma 15, there is a $c > 0$ such that

$$\sigma_{au}(n) \leq \sigma_u(n) + c \cdot n \leq (c+1)\sigma_u(n).$$

It follows that

$$\sigma_u(n) \geq \frac{1}{c+1} \sum_{i=1}^{r} \sigma_u(n_i).$$

$\square$

Lemma 18 (additive property) *Let u_1, u_2 be two sequences with $L(u_1) \cap L(u_2) = \emptyset$. Then there are two positive constants $c_1, c_2 > 0$ such that, for any n,*

$$c_1(\sigma_{u_1}(n) + \sigma_{u_2}(n)) \leq \sigma_{u_1 u_2}(n) \leq \sigma_{u_1}(n) + c_2(\sigma_{u_2}(n) + n).$$

Proof. The first inequality follows from the monotone property (Lemma 16). For the second inequality, consider a (n, u_1u_2)-critical spider $\mathcal{T}$. Let $\mathcal{T}'$ be the maximum u_1-free induced reduction of $\mathcal{T}$.

The size of $\mathcal{T}'$ is at most $\sigma_{u_1}(n)$. If we remove vertices of $\mathcal{T}'$ from $\mathcal{T}$, we get a disjoint union of subspiders of $\mathcal{T}$. Let $\mathcal{T}_1, \ldots, \mathcal{T}_p$ be these subspiders. We have $L(\mathcal{T}_i) \cap L(\mathcal{T}_j) = \emptyset$ for $i \neq j$.

Let $i \in \{1, \ldots, p\}$. The sequence of labels along the $[\mathrm{root}(\mathcal{T}), \mathrm{root}(\mathcal{T}_i)]$-path contains a subsequence s_i isomorphic to u_1 (by the maximality of $\mathcal{T}'$). Let W_i be the set of vertices of $\mathcal{T}_i$ labeled by a label not lying in $L(s_i)$. Then the $(W_i \cup \{\mathrm{root}(\mathcal{T}_i)\})$-reduction of $\mathcal{T}_i$ is au_2-free, where a is the element appearing in u_2 on the first position (otherwise $\mathcal{T}$ would contain u_1u_2). Thus, $|W_i| \leq \sigma_{au_2}(||W_i||) = O(\sigma_{u_2}(||W_i||) + ||W_i||)$ (the second estimate follows from Lemma 15).

In two ways, we now count the number Z of pairs (w, x), where $w \in W_i$, $x \notin W_i$, and x lies in $\mathcal{T}_i$ below w at distance at most $||u_1u_2|| - 1$ from w. We get

$$||u_2||(|\mathcal{T}_i| - ||\mathcal{T}_i|| - |W_i|) \leq Z \leq ||u_1|||W_i|.$$

Consequently,

$$|\mathcal{T}_i| = O(|W_i| + ||\mathcal{T}_i||).$$

We get that

$$\begin{aligned}
\sigma_{u_1u_2}(n) = |\mathcal{T}| &\leq |\mathcal{T}'| + \sum_{i=1}^{p} |\mathcal{T}_i| \\
&\leq \sigma_{u_1}(n) + \sum_{i=1}^{p} O(|W_i| + ||\mathcal{T}_i||) \\
&\leq \sigma_{u_1}(n) + \sum_{i=1}^{p} O\left(\sigma_{u_2}(||W_i||) + ||W_i|| + ||\mathcal{T}_i||\right) \\
&\leq \sigma_{u_1}(n) + \sum_{i=1}^{p} O\left(\sigma_{u_2}(||\mathcal{T}_i||) + ||\mathcal{T}_i||\right) \\
&\leq \sigma_{u_1}(n) + O(\sigma_{u_2}(n) + n)
\end{aligned}$$

(the last inequality follows from Lemma 17). □

Lemma 19 (weak sub-additivity) *For any sequence u and for any integer $k \geq 1$, there is a $c = c(u, k) > 0$ such that $\sigma_u(kn) \leq c \cdot \sigma_u(n)$, for any integer n.*

Proof. Let $\mathcal{T}$ be a (kn, u)-critical spider. By the sparsity lemma (Lemma 12), $\mathcal{T}$ has a $k||u||$-sparse reduction $\mathcal{T}_0$ of size at least $\varepsilon|\mathcal{T}|$. Let $L(\mathcal{T}_0) = L_1 \cup \ldots \cup L_k$ be any partition of the label set $L(\mathcal{T}_0)$ into k sets of size at most n. By the pigeon-hole principle, labels of one of the sets L_i occur altogether at least $\frac{1}{k}|\mathcal{T}_0|$ times in $\mathcal{T}_0$. Let V_1 be the set of vertices labeled by these labels, and let $\mathcal{T}_1$ be the V_1-reduction of $\mathcal{T}_0$. Thus, $|\mathcal{T}_1| \geq \frac{1}{k}|\mathcal{T}_0| \geq \frac{\varepsilon}{k}|\mathcal{T}|$ and $||\mathcal{T}_1|| \leq n$. By the reduction lemma (Lemma 13, with $l = ||u||$ and $\varepsilon = \frac{1}{k}$), $\mathcal{T}_1$ has a $||u||$-sparse reduction of size at least $\varepsilon'|\mathcal{T}_1| \geq \frac{\varepsilon'\varepsilon}{k}|\mathcal{T}| = \frac{\varepsilon'\varepsilon}{k} \cdot \sigma_u(kn)$. The lemma with $c = \frac{k}{\varepsilon'\varepsilon}$ follows. $\square$

3 Letter Insertions

In this section we prove the following generalization of Theorem 1(v).

Theorem 20 (letter insertions) *Let $u = au_1a^2u_2$, $u' = bau_1abau_2$, where u_1, u_2 are two sequences (possibly empty), a, b are letters, and $b \notin L(u)$. Then $\sigma_{u'}(n) = O(\sigma_u(n))$.*

The following lemma is the key part of the proof of Theorem 20.

Lemma 21 (density lemma) *For every sequence u and for every $p \geq 1$, there exists a constant $\varepsilon = \varepsilon(u, p) > 0$ with the following property: Let $\mathcal{T}$ be a $||u||$-sparse u-free spider, and let $<$ be any linear order on $L(\mathcal{T})$. Then $\mathcal{T}$ has a $||u||$-sparse reduction $\mathcal{T}^{(p)}$ on at least $\varepsilon|\mathcal{T}|$ vertices such that, for any label i and for any two different i-vertices v_1, v_2 in $\mathcal{T}^{(p)}$, at least p labels j with $j < i$ occur on the path in $\mathcal{T}$ connecting v_1, v_2.*

Before proving Lemma 21, we derive Theorem 20 from it.

Proof of Theorem 20. Let $\mathcal{T}$ be a (n, u')-critical spider. Consider the partial order $\prec$ on $L(\mathcal{T})$ such that $i \prec j$ if and only if there is an i-vertex lying above all j-vertices. Let $<$ be any linear extension of $\prec$. We apply the density lemma (Lemma 21) with $p = ||u|| + 1$, obtaining a $||u||$-sparse spider $\mathcal{T}^{(p)}$.

We prove that $\mathcal{T}^{(p)}$ is u-free. Suppose to the contrary that $\mathcal{T}^{(p)}$ contains u. Thus, $\mathcal{T}^{(p)}$ has a path P from the root such that the sequence $l(P)$ of labels along P has a subsequence $\tilde{u} = \tilde{a}\tilde{u}_1\tilde{a}^2\tilde{u}_2$ isomorphic to u. By the choice of $\mathcal{T}^{(p)}$, $l(P)$ has a subsequence $\tilde{a}\tilde{u}_1\tilde{a}\tilde{b}\tilde{a}\tilde{u}_2$ for at least $||u||+1 = |L(\tilde{u})|+1$ labels $\tilde{b}, \tilde{b} < \tilde{a}$. Choose $\tilde{b}$ such that $\tilde{b} \notin L(\tilde{u})$. Since $\tilde{b} < \tilde{a}$, at least one occurrence of $\tilde{b}$ precedes any occurrence of $\tilde{a}$ in $l(P)$. Consequently, $l(P)$ has a subsequence $\tilde{b}\tilde{a}\tilde{u}_1\tilde{a}\tilde{b}\tilde{a}\tilde{u}_2$ isomorphic to u' – a contradiction. Thus, $\mathcal{T}^{(p)}$ is u-free.

From $|\mathcal{T}^{(p)}| \geq \varepsilon|\mathcal{T}|$ we get

$$\sigma_u(n) \geq |\mathcal{T}^{(p)}| \geq \varepsilon|\mathcal{T}| = \varepsilon \cdot \sigma_{u'}(n).$$

Consequently,

$$\sigma_{u'}(n) = O(\sigma_u(n)),$$

as required. □

It remains to prove the density lemma (Lemma 21). We first prove the following weaker statement:

Lemma 22 *For every sequence u and for every $p \geq 1$, there exists a constant $\varepsilon_0 = \varepsilon_0(u,p) > 0$ with the following property: Let $\mathcal{T}$ be a $||u||$-sparse u-free spider, and let $<$ be any linear order on $L(\mathcal{T})$. Then $\mathcal{T}$ has a $||u||$-sparse reduction $\mathcal{T}'$ on at least $\varepsilon_0|\mathcal{T}|$ vertices such that, for any label i and for any two different i-vertices v_1, v_2 in $\mathcal{T}'$, at least p vertices on the $[v_1,v_2]$-path in $\mathcal{T}$ are labeled by labels smaller than i (smaller with respect to $<$).*

Proof. We may suppose that $p \geq 2$ (the case $p = 2$ implies the case $p = 1$). Since the lemma is easy to prove for $||u|| = 1$ (with $\varepsilon_0 = \frac{1}{|u|-1}$), we may also assume that $||u|| > 1$.

Suppose that $L(\mathcal{T}) = \{1, 2, \ldots, n\}$ and that $<$ coincides with the usual order of integers. By the sparsity lemma (Lemma 12), there is a $84p||u||$-sparse reduction $\mathcal{T}'$ of $\mathcal{T}$ with $|\mathcal{T}_0| \geq \varepsilon|\mathcal{T}|$. From now on, we consider the spider $\mathcal{T}_0$.

If $|\mathcal{T}_0| < (210p-21)n$, then we may take $\mathcal{T}'$ as any reduction of $\mathcal{T}$ containing exactly one i-vertex for each $i \in L(\mathcal{T})$. Then the size of $\mathcal{T}'$ is $n > \frac{|\mathcal{T}_0|}{210p-21} \geq \frac{\varepsilon}{210p-21}|\mathcal{T}|$ and the lemma holds with any $\varepsilon_0 \leq \frac{\varepsilon}{210p-21}$. We further suppose that $|\mathcal{T}_0| \geq (210p-21)n$.

Let $i \in L(\mathcal{T}_0)$. All but at most $10p-1$ i-vertices in $\mathcal{T}_0$ can be covered by pairwise disjoint i-paths, each containing exactly $10p$ i-vertices. We fix such pairwise disjoint i-paths and call them *i-blocks*. We partition the $10p$ i-vertices in each i-block into $5p$ pairs of i-vertices determining pairwise disjoint

i-paths further called *i-arcs*. A path in $\mathcal{T}_0$ is called *block* or *arc*, if it is an i-block or an i-arc for some $i \in L(\mathcal{T}_0)$, respectively. Certainly, the spider $\mathcal{T}_0$ contains at least $\frac{|\mathcal{T}_0|-(10p-1)n}{10p} \geq \frac{|\mathcal{T}_0|-(10p-1)\cdot\frac{|\mathcal{T}_0|}{210p-21}}{10p} = \frac{2|\mathcal{T}_0|}{21p}$ (and at most $\frac{|\mathcal{T}_0|}{10p}$) blocks. The number of arcs is $5p$ times bigger.

Now, we construct $n+1$ auxiliary sets $\mathcal{U}^{(0)}, \mathcal{U}^{(1)}, \dots \mathcal{U}^{(n)}$ of pairwise disjoint arcs in $\mathcal{T}_0$. We start with $\mathcal{U}^{(0)} = \emptyset$. We now describe how to construct $\mathcal{U}^{(i)}$ from $\mathcal{U}^{(i-1)}$. We say that an i-block in $\mathcal{T}_0$ is *good* if it has non-empty intersection with at least $p+1$ arcs of $\mathcal{U}^{(i-1)}$, otherwise we say that it is *bad*. The set $\mathcal{U}^{(i)}$ is obtained from $\mathcal{U}^{(i-1)}$ in the following way: For each bad i-block B in $\mathcal{T}_0$, remove from $\mathcal{U}^{(i-1)}$ the (at most p) arcs intersecting B and add to $\mathcal{U}^{(i-1)}$ the $5p$ i-arcs contained in B. The obtained set of arcs is taken for $\mathcal{U}^{(i)}$.

The next step in the proof is to estimate the number of bad blocks from above and consequently the number of good blocks from below.

To get the first estimate, consider the set $\mathcal{U} = \mathcal{U}^{(n)}$. Certainly, $\mathcal{U}$ is a set of pairwise disjoint arcs in $\mathcal{T}_0$. Since $\mathcal{T}_0$ is 2-sparse, every arc contains at least 3 vertices. Thus, $|\mathcal{U}| \leq \frac{|\mathcal{T}_0|}{3}$. On the other hand, we can easily estimate the size of $\mathcal{U}$ from below by the number of bad blocks. If b is the number of bad blocks in $\mathcal{T}_0$, then $|\mathcal{U}| \geq (5p-p)b = 4pb$, since to every bad block we can assign an addition of (its) $5p$ arcs to $\mathcal{U}$ and a removal of at most p arcs from $\mathcal{U}$. Thus, $4pb \leq |\mathcal{U}| \leq \frac{|\mathcal{T}_0|}{3}$, which implies $b \leq \frac{|\mathcal{T}_0|}{12p}$.

Since $\mathcal{T}_0$ contains at least $\frac{2|\mathcal{T}_0|}{21p}$ blocks, there are at least $\frac{2|\mathcal{T}_0|}{21p} - \frac{|\mathcal{T}_0|}{12p} = \frac{|\mathcal{T}_0|}{84p}$ good blocks in $\mathcal{T}_0$.

Now, we show for any i that if an i-block is good then it contains at least p vertices labeled by labels smaller than i. Suppose B is a good i-block. Then it intersects at least $p+1$ arcs of $\mathcal{U}^{(i-1)}$. These arcs are pairwise disjoint. Thus, at least p of them do not contain the topmost vertex of B. The topmost vertex of each such arc lies in B. This gives us the desired set of p vertices in B labeled by labels smaller than i.

Let V_1 be the set of the topmost vertices of good blocks in $\mathcal{T}_0$. Thus, the size of V_1 equals the number of good blocks in $\mathcal{T}_0$, and is therefore at least $\frac{|\mathcal{T}_0|}{84p} \geq \frac{\varepsilon}{84p}|\mathcal{T}|$.

Suppose first that V_1 contains the root of $\mathcal{T}_0$, and define $\mathcal{T}_1$ as the V_1-reduction of $\mathcal{T}_0$. Any i-path in $\mathcal{T}_0$ connecting two i-vertices of V_1 contains the (good) i-block starting from its higher end-vertex. Therefore, it contains at least p vertices labeled by labels smaller than i.

If V_1 does not contain the root of $\mathcal{T}_0$, then we remove from V_1 the topmost λ-vertex lying in V_1 ($\lambda :=$ the label of the root), replace it in V_1 by the root of $\mathcal{T}_0$, and define $\mathcal{T}_1$ as the V_1-reduction of $\mathcal{T}_0$. Again, any i-path in $\mathcal{T}_0$ connecting two i-vertices of V_1 contains at least p vertices labeled by labels smaller than i.

In either case, the spider $\mathcal{T}_1$ partially satisfies Lemma 22. The only problem is that $\mathcal{T}_1$ does not have to be $||u||$-sparse. To achieve the $||u||$-sparsity, we apply the reduction lemma. By the reduction lemma (Lemma 13, with $l = ||u||$, $\varepsilon = \frac{1}{84p}$), $\mathcal{T}_1$ has a $||u||$-sparse reduction $\mathcal{T}'$ of size at least $\varepsilon'|\mathcal{T}_0| \geq \varepsilon'\varepsilon|\mathcal{T}|$. $\Box$

We are ready to prove the density lemma.

Proof of Lemma 21. Let $\mathcal{T}^{(p,j)}$ be the spider obtained from $\mathcal{T}$ by j applications of Lemma 22. We set $\mathcal{T}^{(p)} := \mathcal{T}^{(p,p)}$, i.e., $\mathcal{T}^{(p)}$ is the reduction of $\mathcal{T}$ obtained from $\mathcal{T}$ by p applications of Lemma 22. Clearly, $\mathcal{T}^{(p)}$ has the required size. It remains to verify the "density" property of paths in $\mathcal{T}^{(p)}$.

If two vertices u, v in $\mathcal{T}^{(p)} = \mathcal{T}^{(p,p)}$ are labeled by the same label, then the path in $\mathcal{T}^{(p,p-1)}$ between them contains p vertices labeled by smaller labels. If all p vertices are labeled by pairwise different labels then we are done. Otherwise, consider two of these vertices, u_1, v_1, labeled by the same label. Then the path in $\mathcal{T}^{(p,p-2)}$ between them contains p vertices labeled by smaller labels. Again, if all p vertices are labeled by pairwise different labels then we are done. Otherwise, consider two of these vertices, u_2, v_2, labeled by the same label, etc. Continuing this process, we either find p vertices labeled by pairwise different labels or we construct p vertices $u_1, u_2, \ldots, u_p$ with decreasing (thus, pairwise different) labels smaller than $l(u) = l(v)$. $\Box$

4 Sequence Insertions

In this section we prove the following result, which gives Theorem 1(vi).

Theorem 23 (sequence insertion) *Let $u = u_1a^2u_2$, w be two sequences, where u_1, u_2 are two sequences, a is a letter, and $L(u) \cap L(w) = \emptyset$. Then*

$$\sigma_{u_1awau_2}(n) = O(\sigma_w(\sigma_u(n)) + \sigma_u(n)).$$

In particular, if $\sigma_u(n), \sigma_w(n) = O(n)$, then also $\sigma_{u_1awau_2}(n) = O(n)$.

Proof. Consider the sequence $w' = w_1 w_2 \ldots w_{||u||+2}$, $L(w_i) \cap L(w_j) = \emptyset$ $(i \neq j)$, of length $(||u||+1)||w||+1$ consisting of $||u||+2$ intervals $w_1, \ldots, w_{||u||+2}$ over pairwise disjoint alphabets such that $w_1, \ldots, w_{||u||+1}$ are isomorphic to w and $w_{||u||+2}$ is a sequence of length 1. By Lemma 18 (additive property),

$$\sigma_{w'}(n) = O\left(\sum_{i=1}^{||u||+2} \sigma_{w_i}(n) + n\right) = O(\sigma_w(n) + n). \tag{1}$$

For a spider $\mathcal{S}$, we define $\mathcal{S}^*$ as the maximal induced w'-free reduction of $\mathcal{S}$. More generally, if $\mathcal{G}$ is a "multispider" $(S_i, i \in I)$ (i.e., a disjoin union of spiders $\mathcal{S}_i, i \in I$), then we define $\mathcal{G}^*$ as the multispider $(\mathcal{S}_i^*, i \in I)$.

Let $\mathcal{T}$ be a (n,t)-critical spider, where $t = u_1 awau_2$. We set $\mathcal{G}_1 = \mathcal{T}$, and inductively define multispiders $\mathcal{G}_2, \mathcal{G}_3, \ldots$ by taking $\mathcal{G}_{i+1}$ as the induced subgraph of $\mathcal{G}_i$ obtained from $\mathcal{G}_i$ by removing all vertices of $\mathcal{G}_i^*$. In each component of $\mathcal{G}_{i+1}$, the topmost vertex is taken for the root of the component. Clearly, each $\mathcal{G}_i$ (and also each $\mathcal{G}_i^*$) is a multispider such that no label is used on more than one spider belonging to $\mathcal{G}_i$ (to $\mathcal{G}_i^*$), and each vertex of $\mathcal{T}$ lies in exactly one of the graphs $\mathcal{G}_1^*, \mathcal{G}_2^*, \ldots$. Since each component in $\mathcal{G}_i^*$ is w'-free, the weak super-additivity of $\sigma_{w'}$ (Lemma 17) and (1) imply

$$|\mathcal{G}_i^*| = O(\sigma_{w'}(||\mathcal{G}_i^*||)) = O(\sigma_w(||\mathcal{G}_i^*||) + ||\mathcal{G}_i^*||). \tag{2}$$

For each $i > 2$, we choose arbitrarily a set V_i of $||\mathcal{G}_i^*||$ vertices in $\mathcal{G}_i^*$ such that no two vertices of V_i are labeled by the same label. For $i = 1, 2$, a set V_i is chosen arbitrarily so that it contains the root of $\mathcal{T}$ and other $||\mathcal{G}_i^*|| - 1$ vertices of $\mathcal{G}_i^*$ such that no two vertices in V_i are labeled by the same label.

Let $V_{\text{odd}} = V_1 \cup V_3 \cup V_5 \cup \ldots$ and $V_{\text{even}} = V_2 \cup V_4 \cup V_6 \cup \ldots$. Further, let $\mathcal{S}_{\text{odd}}$ and $\mathcal{S}_{\text{even}}$ be the V_{odd}- and V_{even}-reductions of $\mathcal{T}$, respectively. We now get estimates on the size of $\mathcal{S}_{\text{odd}}$ and $\mathcal{S}_{\text{even}}$ by showing that they are u-free and have relatively large $||u||$-sparse reductions.

First, we show that $\mathcal{S}_{\text{odd}}$ is u-free. Suppose to the contrary that there is a path P in $\mathcal{S}_{\text{odd}}$ from the root such that the sequence $l(P)$ of labels appearing along the path P has a subsequence $\tilde{u} = \tilde{u}_1 \tilde{a}^2 \tilde{u}_2$ isomorphic to u. By the construction of $\mathcal{S}_{\text{odd}}$, the two $\tilde{a}$'s in $\tilde{u}$ (between $\tilde{u}_1$ and $\tilde{u}_2$) correspond to two vertices x, y lying in different multispiders $\mathcal{G}_{2j+1}^*, \mathcal{G}_{2k+1}^*$, $j < k$. By the maximality of $\mathcal{G}_{2j+2}^*$, if P' denotes the part of P lying in $\mathcal{G}_{2j+2}^*$, then the sequence of labels along P' has a subsequence $\tilde{w}_1 \tilde{w}_2 \ldots \tilde{w}_{||u||+1}$ isomorphic to $w_1 \ldots w_{||u||+1}$ (otherwise $\mathcal{G}_{2j+2}^*$ could be enlarged by the topmost vertex of $P \cap \mathcal{G}_{2j+3}^*$ — a contradiction with the definition of $\mathcal{G}_{2j+2}^*$).

The label set $L(\tilde{u})$ has a non-empty intersection with at most $|L(\tilde{u})| = ||u||$ of the sets $L(\tilde{w}_i)$. Thus, there is an index $i \in \{1, 2, \ldots, ||u|| + 1\}$ such that $L(\tilde{u}) \cap L(\tilde{w}_i) = \emptyset$. Consequently, $\tilde{u}_1\tilde{a}\tilde{w}_i\tilde{a}\tilde{u}_2$ is a subsequence of $l(P)$ isomorphic to t — a contradiction. Thus, $\mathcal{S}_{\text{odd}}$ is u-free.

We now show that $\mathcal{S}_{\text{odd}}$ has a $||u||$-sparse reduction $\mathcal{S}'_{\text{odd}}$ of size $\Omega(|\mathcal{S}_{\text{odd}}|)$. In fact, $\mathcal{S}'_{\text{odd}}$ can be obtained from $\mathcal{S}_{\text{odd}}$ by running the algorithm $\mathcal{B}(||u||)$ (introduced below Lemma 12).

Let $\mathcal{R}$ be one of the spiders forming the multispider $\mathcal{G}^*_{2i+1}$ for some i. Then we say that $\mathcal{R}$ is *internal*, if at least one path in $\mathcal{T}$ from the root to $\mathcal{G}^*_{2i+2}$ goes through $\mathcal{R}$. Otherwise we say that $\mathcal{R}$ is *external*.

Obviously, each label appears in at most one external spider (in the whole $\mathcal{S}_{\text{odd}}$). Thus, all external spiders contain altogether at most $||\mathcal{S}_{\text{odd}}||$ vertices of $\mathcal{S}_{\text{odd}}$. Consequently, all internal spiders contain altogether at least $|\mathcal{S}_{\text{odd}}| - ||\mathcal{S}_{\text{odd}}||$ vertices of $\mathcal{S}_{\text{odd}}$.

If $\mathcal{R}$ is internal and P is any path from the root to $\mathcal{G}^*_{2i+2}$ intersecting $\mathcal{R}$, then the label set $L(P \cap \mathcal{R})$ of their intersection has size at least $||w'|| - 1 = ||w||(||u|| + 1) \geq ||u|| + 1$ (otherwise $\mathcal{G}^*_{2i+1}$ together with the topmost vertex of $P \cap \mathcal{G}^*_{2i+2}$ is w'-free — a contradiction with the maximality of $\mathcal{G}^*_{2i+1}$). Thus, if $\mathcal{R}$ is internal, then it contains $||\mathcal{R}|| \geq ||u|| + 1$ vertices of $\mathcal{S}_{\text{odd}}$ and, by the definition of the algorithm $\mathcal{B}(||u||)$, at least $||\mathcal{R}|| - (||u|| - 1) \geq ||\mathcal{R}|| - (||u|| - 1)\frac{||\mathcal{R}||}{||u||+1} = \frac{2}{||u||+1}||\mathcal{R}||$ vertices of $\mathcal{S}'_{\text{odd}}$.

Since internal spiders contain altogether at least $|\mathcal{S}_{\text{odd}}| - ||\mathcal{S}_{\text{odd}}||$ vertices of $\mathcal{S}_{\text{odd}}$, we get that $|\mathcal{S}'_{\text{odd}}| \geq \frac{2}{||u||+1}(|\mathcal{S}_{\text{odd}}| - ||\mathcal{S}_{\text{odd}}||)$. If we sum up this inequality with the inequality $|\mathcal{S}'_{\text{odd}}| \geq ||\mathcal{S}_{\text{odd}}|| \geq \frac{2}{||u||+1}||\mathcal{S}_{\text{odd}}||$ (following directly from the definition of $\mathcal{B}(||u||)$ and from $||u|| \geq 1$), we get

$$2|\mathcal{S}'_{\text{odd}}| \geq \frac{2}{||u|| + 1}|\mathcal{S}_{\text{odd}}|.$$

Thus,

$$|\mathcal{S}_{\text{odd}}| \leq (||u|| + 1)|\mathcal{S}'_{\text{odd}}| \leq (||u|| + 1)\sigma_u(n). \tag{3}$$

Similarly, we could prove that

$$|\mathcal{S}_{\text{even}}| \leq (||u|| + 1)\sigma_u(n). \tag{4}$$

Now we are ready to finish the proof. By (2), (3), (4), by the weak super-additivity of σ_w (Lemma 17), and by the weak sub-additivity of σ_w

(Lemma 19),

$$\begin{aligned}
\sigma_t(n) = |\mathcal{T}| &= \sum_{i=1}^{\infty} |\mathcal{G}_i^*| \\
&\overset{(2),\ \text{Lemma 17}}{=} O\left(\sigma_w(\sum_{i \text{ odd}} ||\mathcal{G}_i^*||) + \sum_{i \text{ odd}} ||\mathcal{G}_i^*|| + \sigma_w(\sum_{i \text{ even}} ||\mathcal{G}_i^*||) + \sum_{i \text{ even}} ||\mathcal{G}_i^*||\right) \\
&= O\left(\sigma_w(|\mathcal{S}_{\text{odd}}|) + |\mathcal{S}_{\text{odd}}| + \sigma_w(|\mathcal{S}_{\text{even}}|) + |\mathcal{S}_{\text{even}}|\right) \\
&\overset{(3),\ (4)}{=} 2 \cdot O(\sigma_w((||u|| + 1)\sigma_u(n)) + (||u|| + 1)\sigma_u(n)) \\
&\overset{\text{Lemma 19}}{=} O(\sigma_w(\sigma_u(n)) + \sigma_u(n)).
\end{aligned}$$

□

We now summarize the proof of Theorem 1 contained in Sections 2–4.

Proof of Theorem 1. Part (i) is trivial, part (ii) follows from Lemma 16. Part (iii) is a corollary of Lemma 14, part (iv) a corollary of Lemma 15. Part (v) follows from Theorem 20, and part (vi) from Theorem 23. □

5 A Lower Bound Construction

Here we describe a construction giving Theorem 2. For any two positive integers k and m, we construct a spider $\mathcal{T}(k, m)$ with the following five properties:

(P1) $\mathcal{T}(k, m)$ is *ababa*-, *aba*^{2}*b*-, *abcabc*-, and *abcacb*-free,

(P2) $c(k, m) := \frac{||\mathcal{T}(k,m)||}{m}$ is an integer,

(P3) for each label $i \in L(\mathcal{T}(k, m))$, there is a (unique) leaf $z(i)$ labeled by i; moreover, the leaves $z(i)$ form $c(k, m)$ pairwise disjoint fans, where *fan* is a set of m leaves having a common neighbor (father),

(P4) in each fan in $\mathcal{T}(k, m)$, there is a leaf labeled by the same label as its father; otherwise no two neighbors in $\mathcal{T}(k, m)$ are labeled by the same label,

(P5) $\frac{|\mathcal{T}(k,m)|}{||\mathcal{T}(k,m)||} > \frac{k}{2}$.

The spiders $\mathcal{T}(k,m)$ are constructed by double induction on k and m as follows.

For $k=1$ and any m, $\mathcal{T}(1,m)$ consists of the root and a single fan of m leaves connected to the root (see Fig. 6).

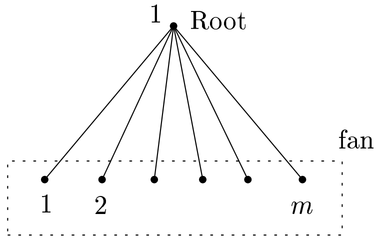


Figure 6: The spider $\mathcal{T}(1,m)$

For $k>1$ and $m=1$, we construct $\mathcal{T}(k,1)$ from $\mathcal{T}(k-1,2)$ by adding an edge and a vertex in each fan as shown in Fig. 7. We get two one-term fans in $\mathcal{T}(k,1)$ in place of each two-term fan in $\mathcal{T}(k-1,2)$. For $k>1$

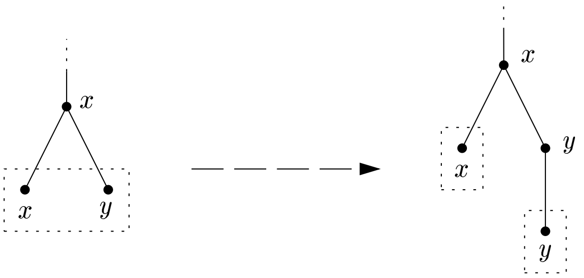


Figure 7: Obtaining $\mathcal{T}(k,1)$ from $\mathcal{T}(k-1,2)$ (dotted rectangles denote fans).

and $m>1$, we construct $\mathcal{T}(k,m)$ from the spiders $\mathcal{T}=\mathcal{T}(k,m-1)$ and $\mathcal{V}=\mathcal{T}(k-1,M)$, where $M=c(k,m-1)=\frac{||\mathcal{T}(k,m-1)||}{m-1}$ is the number of fans in $\mathcal{T}$, as follows (see Fig. 8). Let $F_1,F_2,\ldots,F_P$ be the fans in $\mathcal{V}$ (listed in an arbitrary order). For each $i=1,\ldots,P$, we take a spider $\mathcal{T}_i$ isomorphic to $\mathcal{T}=\mathcal{T}(k,m-1)$ such that the label sets $L(\mathcal{V}),L(\mathcal{T}_1),\ldots,L(\mathcal{T}_P)$ are pairwise disjoint. For $i=1,\ldots,P$, we denote the leaves in F_i by $f_{i1},f_{i2},\ldots,f_{iM}$ (in an arbitrary order) and their common father by f_i. Further, we denote the root of $\mathcal{T}_i$ by r_i and the fans in $\mathcal{T}_i$ by $G_{i1},G_{i2},\ldots,G_{iM}$. For $j=1,\ldots,M$, let g_{ij} be the common father of the leaves in G_{ij}.

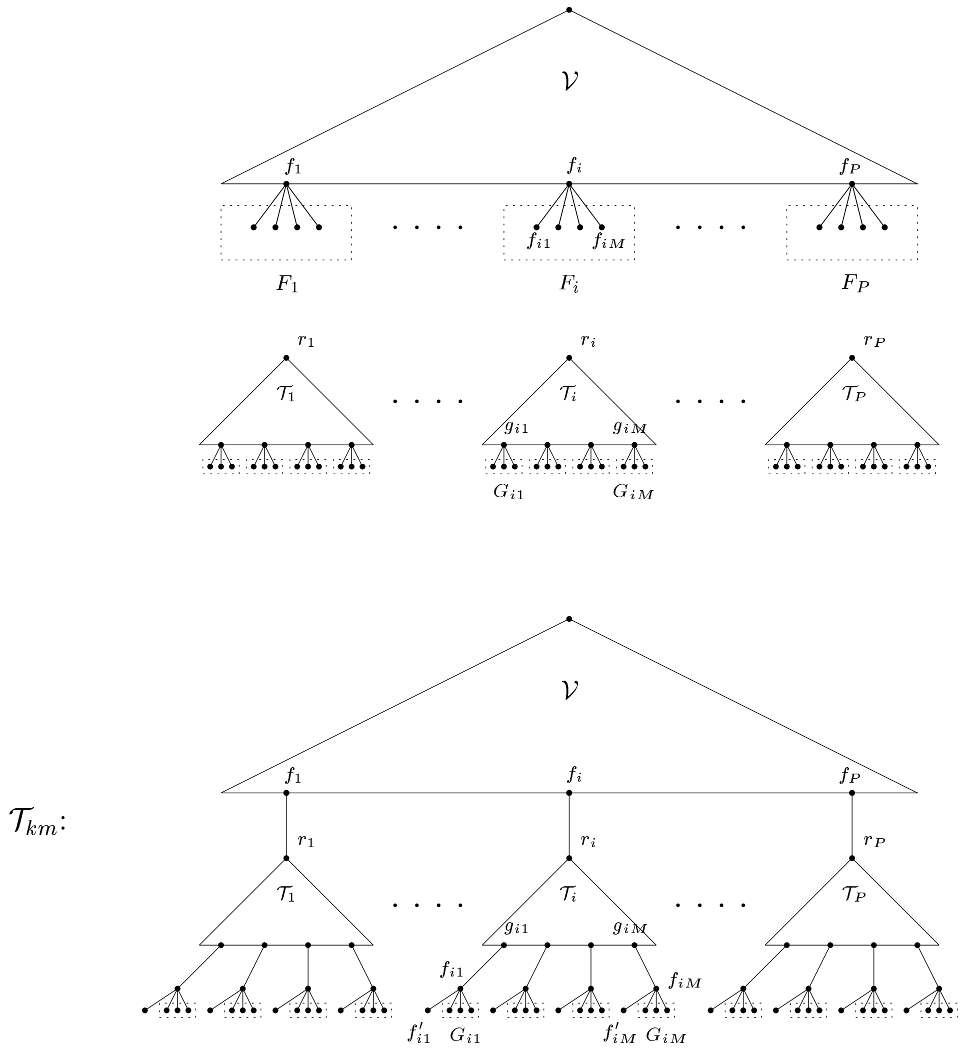


Figure 8: The inductive construction of $\mathcal{T}(k, m)$, $k > 1$, $m > 1$.

The spider $\mathcal{T}(k,m)$ is obtained from the disjoint union of the spiders $\mathcal{V}$, $\mathcal{T}_1\ldots,\mathcal{T}_P$ in the following four steps made for each $i=1,\ldots,P$ (see Fig. 8):

1. remove all edges connecting the vertices in the fans $F_i, G_{i1}, G_{i2},\ldots,G_{iM}$ with their fathers,

2. connect f_i with r_i by an edge,

3. for each $j=1,\ldots,M$, connect f_{ij} by an edge with the vertex g_{ij} and with all vertices of the fan G_{ij},

4. for each $j=1,\ldots,M$, add a new vertex (leaf) f'_{ij} labeled by $l(f_{ij})$ and connect it by an edge to f_{ij}.

The fans in $\mathcal{T}(k,m)$ are $G_{ij}\cup\{f'_{ij}\}$, $i=1,\ldots,P, j=1,\ldots,M$.

Properties (P2)–(P5) follow directly from the construction. We now verify (P1). It follows from the construction that if $\mathcal{T}(k,m), k,m>1$, contains some sequence u, then u is a subsequence of some sequence $u_1\alpha u_2\alpha\beta$ or $u_1\alpha u_2\alpha^2$, where $|\alpha|=|\beta|=1$, $L(u_1\alpha)\cap L(u_2\beta)=\emptyset$, $\mathcal{S}(k-1,c(k,m-1))=\mathcal{V}$ contains $u_1\alpha$, and $\mathcal{T}(k,m-1)=\mathcal{T}$ contains $u_2\beta$.

Consequently, it easily follows by induction that if any spider $\mathcal{T}(k,m)$ contains a sequence u with $||u||\le 2$, then u is isomorphic to some sequence $a^pb^qa^rb^s$, $p,q,s\ge 0$, $r\in\{0,1\}$. Thus, the spiders $\mathcal{T}(k,m)$ are *ababa*-free and aba^2b-free.

Consequently, again by induction, we get that if any spider $\mathcal{T}(k,m)$ contains a sequence with $||u||\le 3$, then u is isomorphic to some sequence $a^pb^qa^rb^sd^tb^xd^y$ or to some sequence $d^ta^pb^qa^rb^sd^xb^y$. It follows that the spiders $\mathcal{T}(k,m)$ are also *abcabc*-free and *abcacb*-free. This finishes the verification of property (P1).

We do not know if there is a *ababa*-, aba^2b-, *abcabc*-, and *abcacb*-free sequence u with $\sigma_u(n)\ne O(n)$ (the above construction does not help here — every *ababa*-, aba^2b-, *abcabc*-, and *abcacb*-free sequence u is contained in at least one spider $\mathcal{T}(k,m)$).

We are ready to prove Theorem 2.

Proof of Theorem 2. We prove only part (i), part (ii) then easily follows from the monotone property (Lemma 16).

It follows from properties (P1)–(P5) that $\frac{\sigma_{ababa}(n)}{n}$ and $\frac{\sigma_{aba^2b}(n)}{n}$ are unbounded, since the removal of all leaves in $\mathcal{T}(k,m)$ gives a $(||\mathcal{T}(k,m)||, ababa)$-spider (and a $(||\mathcal{T}(k,m)||, aba^2b)$-spider) of size at least $(\frac{k}{2}-1)||\mathcal{T}(k,m)||$.

Since $\mathcal{T}(k,m)$ is also *abcabc*-free and *abcacb*-free, Lemma 12 (sparsity lemma, with $u = ababa$, $l = 3$) implies that also $\frac{\sigma_{abcabc}(n)}{n}$ and $\frac{\sigma_{abcacb}(n)}{n}$ are unbounded. Now, let u be any sequence containing at least one of the sequences $ababa$, aba^2b, $abcabc$, and $abcacb$. Then from above and by Proposition 10, $\frac{\sigma_u(n)}{n} \to \infty$.

Actually, the numbers $c(k,m)$ grow analogously as the "Ackermann" numbers $A(k,m)$:

- $c(1,m) = 1$, for $m \geq 1$,
- $c(k,1) = 2 \cdot c(k-1,2)$, for $k > 1$,
- $c(k,m) = c(k,m-1) \cdot c(k-1, c(k,m-1))$, for $k > 1, m > 1$.

Therefore, the numbers $||\mathcal{T}(m,m)|| = m \cdot c(m,m)$, $m \geq 1$, grow asymptotically as the Ackermann function $A(m) = A(m,m)$. From this and from reasoning as above we get $\sigma_u(||\mathcal{T}(m,m)||) = \Omega(m \cdot ||\mathcal{T}(m,m)||)$ for the four sequences $u = ababa, aba^2b, abcabc, abcacb$. It follows that the spiders $\mathcal{T}(k,m)$ give (i) (and also (ii)) in Theorem 2. $\square$

Our construction also gives the following result:

Theorem 24 ([V99b]) *If u is a 2-sparse sequence with $|u| > 3||u||-2$, then*

$$\sigma_u(n) \geq \hat{\sigma}_u(n) = \Omega(n\alpha(n)).$$

Theorem 24 is obtained from the above construction in [V99b, Thm. 43, page 386]. The bound $|u| > 3||u|| - 2$ in it is best possible (see [V99b]).

6 Proofs of the Corollaries

In this section we derive Theorems 3–5 from Theorems 1 and 2.

Proof of Theorem 3. **(i)** Let u be a sequence with $||u|| \leq 2$. By Theorem 2, if u contains $ababa$ or aba^2b, then $\sigma_u(n) = \Omega(n\alpha(n))$. Suppose now that u is $ababa$-free and aba^2b-free. Then u is isomorphic to some sequence $a^ib^ja^kb^l$, $i,j,l \geq 0, k \in \{0,1\}$. By Theorem 1(ii), we may suppose that $j \geq 2, k = 1$. By Theorem 1(i), $\sigma_u(n) = O(n)$ holds for the sequence $u = b^{j+l}$. Consequently, by Theorem 1(v), $\sigma_u(n) = O(n)$ holds also for the sequence $u = ab^jab^l$. Part (i) now follows from Theorem 1(iv).

(ii) Let $||u|| \leq 2$. If u is isomorphic to some sequence $a^i b^j a^k$, $i, j, k \geq 0$, then $\hat{\sigma}_u(n) \leq \sigma_u(n) = O(n)$ by (i), and (ii) holds.

If u contains $ababa$ then $\hat{\sigma}_u(n) \geq \hat{\sigma}_{ababa}(n) = \sigma_{ababa}(n) = \Omega(n\alpha(n))$ according to Theorem 2, and (ii) holds again.

It remains to verify (ii) for every sequence $u = a^i b^j a^k b^l$, $i, j, k, l \geq 1$. If u contains ab^2a^2b (i.e., $j, k \geq 2$), then Theorem 2(ii) gives $\hat{\sigma}_u(n) = \Omega(n\alpha(n))$. If u is ab^2a^2b-free, then $j = 1$ or $k = 1$. If $k = 1$, then $\hat{\sigma}_u(n) \leq \sigma_u(n) = O(n)$ by (i). If $j = 1$, then $\hat{\sigma}_u(n) = \hat{\sigma}_{\overline{u}}(n) \leq \sigma_{\overline{u}}(n) = O(n)$ also by (i). □

Proof of Theorem 4. We proceed by induction on $|u|$. If $|u| \leq 1$, then the statement clearly holds. Suppose now that $|u| \geq 2$ and that $\hat{\sigma}_v(n) \leq \sigma_v(n) = O(n)$ holds for any $abab$-free sequence with $|v| < |u|$. Write $u = x^{j_1} u_1 x^{j_2} u_2 \ldots u_{r-1} x^{j_r} u_r$, where $x \in L(u)$, $x \notin L(u_i)$, $j_i > 0$ (for $i = 1 \ldots, r$).

Since u is $abab$-free, the alphabets $L(u_i)$ are pairwise disjoint. By the induction assumption, $\sigma_{u_i}(n) = O(n)$ for each i. By Theorem 1(i), we have $\sigma_t(n) = O(n)$ for the sequence $t = x^{j_1+\ldots+j_r+1}$. By repeated applications of Theorem 1(vi), we then get $\sigma_t(n) = O(n)$ consecutively for the sequences $t = x^{j_1+\ldots+j_r} u_r x$, $t = x^{j_1+\ldots+j_{r-1}} u_{r-1} x^{j_r} u_r x, \ldots$, $t = x^{j_1} u_1 \ldots u_{r-1} x^{j_r} u_r x = ux$. Hence, $\hat{\sigma}_u(n) \leq \sigma_u(n) \leq \sigma_{ux}(n) = O(n)$. □

Proof of Theorem 5. If u is $abab$-free, then Theorem 5 follows from Theorem 4.

If u contains $abab$, then it also contains ab^2a^2b and thus $\sigma_u(n) \geq \hat{\sigma}_u(n) = \Omega(n\alpha(n))$ by Theorem 2(ii) and by Theorem 1(ii). □

7 Relation between $\hat{\sigma}_u(n)$ and $\sigma_v(n)$

In this section we prove Theorem 6 (derivation rule) and derive Theorem 7 from it.

Proof of Theorem 6. Let u be a sequence with $|u| \geq ||u||$, and let $u = vw, L(v) \cap L(w) = \emptyset$. We have to show that $\hat{\sigma}_u(n) = O(\sigma_{\overline{v}}(n) + \sigma_w(n))$.

Let $\mathcal{T}$ ba a $(\hat{u}, n)$-critical spider (i.e., it is a $\hat{u}$-free $||u||$-sparse spider with $|\mathcal{T}| = \hat{\sigma}_u(n)$, $||\mathcal{T}|| \leq n$). For an induced V'-reduction $\mathcal{T}'$ of $\mathcal{T}$, consider the collection of subspiders $(\mathcal{S}_i, i \in I)$ of $\mathcal{T}$ obtained from $\mathcal{T}$ by removing vertices of $\mathcal{T}'$. From now on, let $\mathcal{T}'$ be minimal with the property that all spiders $\mathcal{S}_i, i \in I$, are t-free, where $t = \alpha_1 \ldots \alpha_{||u||+2}$ is a sequence obtained by concatenating $||u||+2$ sequences α_i, $i = 1, \dots, ||u||+2$, which are isomorphic to $\overline{v}$ each and have pairwise disjoint label sets $L(\alpha_i)$.

We now separately estimate $|\mathcal{T}'|$ and $|\mathcal{T}| - |\mathcal{T}'| = \sum_{i \in I} |\mathcal{S}_i|$. The spiders $\mathcal{S}_i$ have distinct label sets $L(\mathcal{S}_i)$. Thus, by the weak super-additivity (Lemma 17) and by the additive property (Lemma 18, repeatedly applied),

$$|\mathcal{T}| - |\mathcal{T}'| = \sum_{i \in I} |\mathcal{S}_i| \leq \sum_{i \in I} \sigma_t(||S_i||) = O(\sigma_t(n)) = O(\sigma_{\overline{v}}(n) + n).$$

To estimate $|\mathcal{T}'|$, we consider the structure of $\mathcal{T}'$. Let y be the topmost vertex in $\mathcal{T}'$ such that it has at least two neighbors lying under it (see Fig. 9). $\mathcal{T}'$ can be partitioned into the $[\text{root}, y]$-path P and a spider $\mathcal{V}$ rooted in y

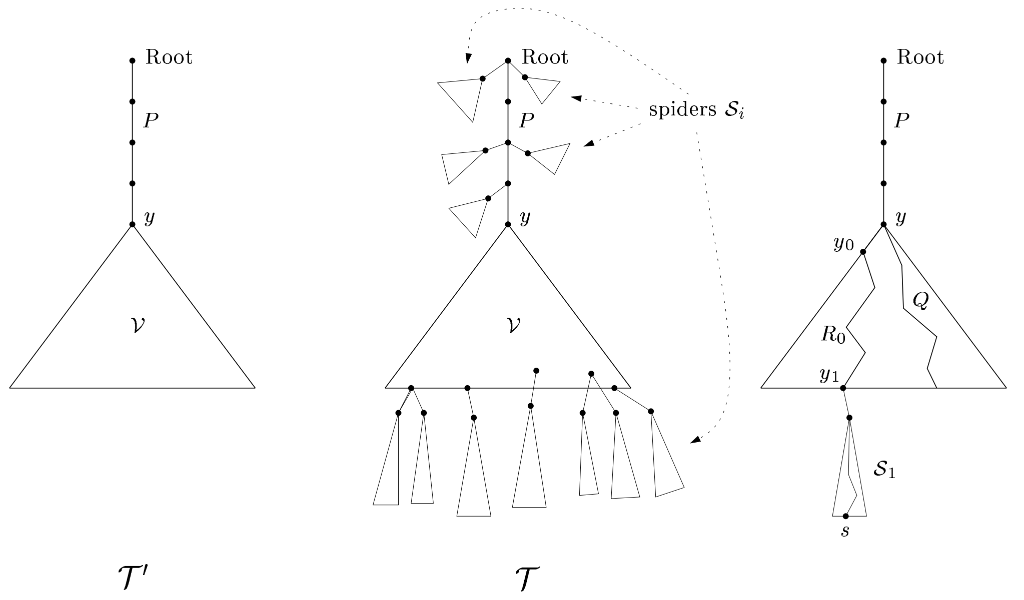


Figure 9: The trees $\mathcal{T}'$, $\mathcal{T}$, and the considered paths in $\mathcal{T}$.

and containing y and all vertices under y (P and $\mathcal{V}$ intersect in the unique vertex y).

We recall from [V99b, Lemma 15] that $f_u(n) = O(f_v(n) + f_w(n))$. Since the sequence of labels along P has to be u-free, we get

$$\begin{aligned} |P| &\leq f_u(n) \\ &= O(f_v(n) + f_w(n)) \end{aligned}$$

$$
\begin{aligned}
&= O(f_{\overline{v}}(n) + f_w(n)) \\
&= O(\sigma_{\overline{v}}(n) + \sigma_w(n)).
\end{aligned}
$$

We now show that $\mathcal{V}$ is w-free. Suppose to the contrary that there is a path Q from the root y such that the sequence of labels along Q has a subsequence $\tilde{w}$ isomorphic to w. Let y_0 be a neighbor of y lying under y and not lying on Q. Further, let R_0 be any maximal path in $\mathcal{V}$ from y with the first edge yy_0. Then the last vertex of R_0, denoted y_1, is adjacent to the root of some spider S_i (S_1, say) such that the sequence of labels along the path from y_1 to some vertex s in $\mathcal{S}_1$ has a subsequence $\tilde{t} = \tilde{\alpha}_1 \dots \tilde{\alpha}_{||w||+2}$ isomorphic to $t = \alpha_1 \dots \alpha_{||w||+2}$. At least one of the $||w|| + 1$ sequences $\tilde{\alpha}_2, \dots, \tilde{\alpha}_{||w||+2}$ contains no element of $L(\tilde{w})$ (since the sets $L(\alpha_i)$ are pairwise disjoint and $|L(\tilde{w})| = ||w||$). We denote this sequence by $\tilde{\alpha}_j$. Then the sequence $\overline{\tilde{\alpha}_j}\tilde{w}$ is isomorphic to u. Since the vertices corresponding to it lie on a path, $\mathcal{T}$ contains u — a contradiction.

Thus, $\mathcal{V}$ is w-free and $|\mathcal{V}| \le \sigma_w(n)$. It follows that

$$|\mathcal{T}'| = |P| + |\mathcal{V}| - 1 = O(\sigma_{\overline{v}}(n) + \sigma_w(n)).$$

Consequently,

$$\hat{\sigma}_u(n) = |\mathcal{T}| = |\mathcal{T}'| + (|\mathcal{T}| - |\mathcal{T}'|) = O(\sigma_{\overline{v}}(n) + \sigma_w(n) + n) = O(\sigma_{\overline{v}}(n) + \sigma_w(n)).$$

□

Proof of Theorem 7. Let u be isomorphic to $\overline{u}$, and let $u = u_0 u_1 \dots u_r$ be the partition of u into the maximum number of non-empty intervals such that their label sets $L(u_i)$ are pairwise disjoint. Then $\overline{u_0 u_1 \dots u_j} = \overline{u}_j \dots \overline{u}_1 \overline{u}_0$ is isomorphic to $u_{r-j} u_{r-j+1} \dots u_r$ for each $j = 0, \dots, r$. Hence, setting $q = \lceil \frac{r}{2} \rceil$, we get

$$\min\{\sigma_{\overline{v}}(n) + \sigma_w(n) : u = vw, L(v) \cap L(w) = \emptyset\} \quad = \quad \Theta(\sigma_{u_q u_{q+1} \dots u_r}(n))$$

from the monotone property (Lemma 16).

Since the sequences $\overline{u_0 u_1 \dots u_{\lfloor \frac{r}{2} \rfloor}}$, $u_q u_{q+1} \dots u_r$ are isomorphic, any $u_q u_{q+1} \dots u_r$-free spider is $\hat{u}$-free. We get

$$\hat{\sigma}_u(n) \ge \sigma_{u_q u_{q+1} \dots u_r}(n) = \Theta(\min\{\sigma_{\overline{v}}(n) + \sigma_w(n) : u = vw, L(v) \cap L(w) = \emptyset\}).$$

On the other hand,

$$\hat{\sigma}_u(n) = O(\min\{\sigma_{\overline{v}}(n) + \sigma_w(n) : u = vw, L(v) \cap L(w) = \emptyset\})$$

by Theorem 6. □

8 2-Sparse Sequences over 3 Letters

It can be verified from Theorem 1, Theorem 6, and from results on $f_u(n)$ given in [AKV92] (see also [V99b]) that all three function $f_u(n)$, $\sigma_u(n)$, $\hat{\sigma}_u(n)$ are of order $O(n)$ for any 2-sparse sequence u of length at most 5 not isomorphic with *ababa*. On the other hand, the functions $f_u(n)$, $\sigma_u(n)$, and $\hat{\sigma}_u(n)$ are of order $\Omega(n\alpha(n))$ for any sequence u containing *ababa* according to the monotone property of f_u (see [V99b]), $f_{ababa}(n) = \Theta(n\alpha(n))$ (proved in [HS86], see also [Ko88, WS88, SA95]), and obvious inequalities $\sigma_u(n) \geq f_u(n)$ and $\hat{\sigma}_u(n) \geq f_u(n)$.

Here is a table showing what we know about the validity of $f_u(n) = O(n)$, $\sigma_u(n) = O(n)$, and $\hat{\sigma}_u(n) = O(n)$ for the remaining sixteen 2-sparse sequences u with $||u|| = 3$:

u	$\sigma_u(n) = O(n)$?	$\hat{\sigma}_u(n) = O(n)$?	$f_u(n) = O(n)$?
ababcb	+	+	+
abacac	+	+	+
abacab	−	+	+
abcbab	+	+	+
abcacb	−	?	+
abcbac	?	?	+
abcabc	−	−	+
abacbc	+	+	+
abcbca	+	+	+
ababcbc	+	+	+
abcbabc	−	−	+
abacabc	−	−	?
abcacbc	−	−	?
abacacb	−	?	?
abcbcac	?	?	?
abacacbc	−	−	?

The +'s in the last column are explained in [V99b], they all follow from the result mentioned below in the first paragraph in Section 9. The other $\pm$'s follow from Theorems 1, 2, and 6.

The sequence $u = abcbab$ satisfies $\sigma_u(n) = O(n)$ and $\sigma_{\overline{u}}(n) \neq O(n)$ (by Theorem 3, this holds also for $u = ab^2ab$). This shows that $\sigma_u(n)$ and $\sigma_{\overline{u}}(n)$ may be significantly different. One may compare this with the equalities

$\hat{\sigma}_u(n) = \hat{\sigma}_{\overline{u}}(n)$ and $f_u(n) = f_{\overline{u}}(n)$ holding clearly for all u and n. By these equalities, in the above table the row for u may differ from the row for $\overline{u}$ only in the first column (pairs $u, \overline{u}$ of "mutually reversed" sequences lie in these rows: 1./2., 3./4., 5./6., 12./13., 14./15.).

9 Concluding Remarks

1. An analogue of Theorem 1 for generalized Davenport–Schinzel sequences. A result very similar to Theorem 1 holds for the set of sequences u with $f_u(n) = O(n)$. In fact, we can get the wording of a result of [KV94] from the wording of Theorem 1 by the following three changes: 1. change $\sigma_u(n) = O(n)$ to $f_u(n) = O(n)$ on the first line, 2. change $u' = bau_1abau_2$ to $u' = bau_1ab^2au_2$ in part (v), 3. add a new rule: (vii) (symmetry) *If u lies in $\mathcal{L}$ then $\overline{u}$ lies also in $\mathcal{L}$.*

It is interesting that the little changes allow to prove $f_u(n) = O(n)$ even for many sequences for which $\sigma_u(n) \neq O(n)$ is known. In this respect the two results are quite different. Of course, we were strongly motivated by [KV94]. In the proof of Theorem 1 we used some of the ideas of the proof [KV94] of the above mentioned analogue. However, on many places the proof of Theorem 1 is more difficult and we had to find new techniques to accomplish it.

2. Comparison of $f_u(n)$, $\sigma_u(n)$, and $\hat{\sigma}_u(n)$. The above mentioned analogue of Theorem 1 gives also an analogue of Theorem 3 for the function $f_u(n)$ (see also [V99b]):

Theorem 25 (two-letter theorem [AKV92]) *If u is a sequence with $||u|| \leq 2$ then $f_u(n) = O(n)$ if and only if u is ababa-free.*

Thus,

$$f_u(n) = O(n), \quad \hat{\sigma}_u(n) \neq O(n), \quad \sigma_u(n) \neq O(n)$$

holds for all sequences $u = a^i b^j a^k b^l, i, l \geq 1, j, k \geq 2$, and

$$f_u(n) = O(n), \quad \hat{\sigma}_u(n) = O(n), \quad \sigma_u(n) \neq O(n)$$

holds for all sequences $u = a^i b a^k b^l, i, l \geq 1, k \geq 2$.

This shows that the functions $f_u(n)$, $\sigma_u(n)$, $\hat{\sigma}_u(n)$ are significantly different already for sequences over two letters.

3. Other models of Davenport–Schinzel trees. Let $\mathcal{T}$ be a labeled tree. For a label i, let $\mathcal{T}_i$ be the smallest (labeled) subtree of $\mathcal{T}$ containing all vertices labeled by i. We say that $\mathcal{T}$ is

- *weakly spidery* if $\mathcal{T}_i$ is a path for every label i,
- *skeletal* if, for every label i, all vertices in $\mathcal{T}_i$ of degree more than 2 are labeled by i.

Clearly, every spider is weakly spidery and every weakly spidery labeled tree is skeletal. Define

$\sigma'_u(n) =$ the maximum size (i.e., number of vertices) of a weakly spidery (n, u)-tree,

$\tau_u(n) =$ the maximum size of a skeletal (n, u)-tree (unless $||u|| = 1$).

If $||u|| = 1$ then we set $\tau_u(n) := \sigma'_u(n)$ for technical reasons (the above definition would give $\tau_u(n) = +\infty$ for $|u| > 2, ||u|| = 1$). Here are "unrooted" analogues of the above functions:

$\hat{\sigma}'_u(n) =$ the maximum size of a weakly spidery $(n, \hat{u})$-tree,

$\hat{\tau}_u(n) =$ the maximum size of a skeletal $(n, \hat{u})$-tree (unless $||u|| = 1$).

If $||u|| = 1$ then we set $\hat{\tau}_u(n) := \hat{\sigma}'_u(n)$ for technical reasons (the above definition would give $\hat{\tau}_u(n) = +\infty$ for $|u| > 3, ||u|| = 1$).

The function $\hat{\tau}_u(n)$ (originally denoted by $Ex(u, n)^{\mathrm{T}}$) was introduced in [Kl95] and further investigated in [Kl97, Kl99, V99a], while the function $\hat{\sigma}_u(n)$ is considered in [V00, V99b]. The remaining four functions $\sigma_u(n)$, $\sigma'_u(n)$, $\tau_u(n)$, $\hat{\sigma}'_u(n)$ were introduced in [V00], where it was shown that the functions $\hat{\sigma}_u(n)$, $\hat{\sigma}'_u(n)$, $\hat{\tau}_u(n)$ (and also the functions $\sigma_u(n)$, $\sigma'_u(n)$, $\tau_u(n)$) have the same asymptotic behavior for each u:

Theorem 26 (Valtr [V00]) *For any sequence u,*

$$\Theta(\sigma_u(n)) = \Theta(\sigma'_u(n)) = \Theta(\tau_u(n)) \quad \text{and} \quad \Theta(\hat{\sigma}_u(n)) = \Theta(\hat{\sigma}'_u(n)) = \Theta(\hat{\tau}_u(n)).$$

Let us remark that the definitions give

$$f_u(n) \le \sigma_u(n) \le \sigma'_u(n) \le \tau_u(n), \qquad \hat{\sigma}_u(n) \le \hat{\sigma}'_u(n) \le \hat{\tau}_u(n),$$

$$\hat{\sigma}_u(n) \le \sigma_u(n), \quad \hat{\tau}_u(n) \le \tau_u(n), \quad \hat{\sigma}'_u(n) \le \sigma'_u(n),$$

for each u and n. If u is isomorphic to $\overline{u}$ (u is "symmetric"), then also $f_u(n) \le \hat{\sigma}_u(n)$.

Due to Theorem 26, asymptotic results on the function $\sigma_u(n)$ hold also for the functions $\sigma'_u(n)$ and $\tau_u(n)$. Similarly, asymptotic results on the function $\hat{\sigma}_u(n)$ hold also for the functions $\hat{\sigma}'_u(n)$ and $\hat{\tau}_u(n)$. In this paper we worked with the functions $\sigma_u(n)$ and $\hat{\sigma}_u(n)$, since the model of spiders is much easier to handle with than the models of skeletal and weakly spidery trees.

4. Applications of (generalized) Davenport–Schinzel sequences. Our main motivation for the study of Davenport–Schinzel trees is their close relation to (generalized) Davenport–Schinzel sequences, which have many applications. Classical Davenport–Schinzel sequences were applied to numerous problems in computational and combinatorial geometry (see [SA95]), their generalized version was applied to Turán–type questions for geometric graphs [V97, V98] (see also [V99b]) and to an enumeration problem of Stanley and Wilf concerning the number of permutations avoiding a fixed permutation-pattern [AF00, Kl00].

5. Davenport–Schinzel theory of matrices. Problems for matrices avoiding fixed "forbidden" submatrix-patterns were considered in several papers, e.g. see papers [AF00, FH92, FP94] containing also other references.

References

[AKV92] R. Adamec, M. Klazar and P. Valtr, Generalized Davenport–Schinzel sequences with linear upper bound, *Discrete Math.* **108** (1992), 219–229.

[AF00] N. Alon and E. Friedgut, On the number of permutations avoiding a given pattern. *J. Comb. Theory, Ser. A,* **89** (2000), 133–140.

[FH92] Z. Füredi and P. Hajnal, Davenport-Schinzel theory of matrices, *Discrete Mathematics* **103** (1992), 233–251.

[FP94] Z. Füredi and J. Pach, Traces of finite sets: extremal problems and geometric applications, in: *Extremal Problems for Finite Sets*, Proc. Colloq., Visegrád, Hungary, 1991. (P. Frankl et al., Eds.) *Bolyai Society Mathematical Studies* **3** (1994), 251–282.

[HS86] S. Hart and M. Sharir, Nonlinearity of Davenport–Schinzel sequences and of generalized path compression schemes, *Combinatorica* **6** (1986), 151–177.

[Kl92] M. Klazar, A general upper bound in extremal theory of sequences, *Comment. Math. Univ. Carolinae* **33** (1992), 737–747.

[Kl94] M. Klazar, A linear upper bound in the extremal theory of sequences, *J. Comb. Th., Ser. A*, **68** (1994), 454–464.

[Kl95] M. Klazar, *Combinatorial aspects of Davenport–Schinzel sequences*, doctoral thesis, Charles University, Prague, 1995.

[Kl97] M. Klazar, Combinatorial aspects of Davenport–Schinzel sequences, *Discrete Math.* **165/166** (1997), 431–445.

[Kl99] M. Klazar, Extremal problems for colored trees and Davenport–Schinzel sequences, *Discrete Math.* **197/198** (1999), 469–482.

[Kl00] M. Klazar, The Füredi–Hajnal conjecture implies the Stanley–Wilf conjecture, in: D. Krob, A. A. Mikhalev, and A. V. Mikhalev (Eds.), *Formal Power Series and Algebraic Combinatorics*, [Proceedings of the 12th FPSAC conference, Moscow, June 25-30, 2000,] Springer, Berlin, 2000, pp. 250-255.

[KV94] M. Klazar and P. Valtr, Generalized Davenport–Schinzel sequences, *Combinatorica* **14** (1994), 463–476.

[Ko88] P. Komjáth, A simplified construction of nonlinear Davenport–Schinzel sequences, *J. Comb. Th., Ser. A,* **49** (1988), 262–267.

[Ro67] H. Rogers, *Theory of recursive functions and effective computability*, McGraw-Hill, New York (1967).

[SA95] M. Sharir and P. K. Agarwal, *Davenport–Schinzel sequences and their geometric applications*, Cambridge University Press, Cambridge, 1995.

[V97] P. Valtr, Graph drawings with no k pairwise crossing edges, in: G. DiBattista (ed.), *Graph Drawing '97*, Rome, September 1997, pp. 205–218.

[V98] P. Valtr, On geometric graphs with no k pairwise parallel edges, *Discrete Comput. Geom.* **19** (1998), 461–469.

[V99a] P. Valtr, On an extremal problem for colored trees, *Europ. J. Comb.* **20** (1999), 115-121.

[V99b] P. Valtr, Generalizations of Davenport–Schinzel trees, in: *Contemporary Trends in Discrete Mathematics* (Edited by: J. Nešetřil, J. Kratochvíl, F.S. Roberts, R.L. Graham), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Volume 49 (1999), 349-389.

[V00] P. Valtr, Three asymptotically equivalent definitions of Davenport–Schinzel trees, manuscript.

[WS88] A. Wiernik and M. Sharir, Planar realization of nonlinear Davenport-Schinzel sequences by segments, *Discrete Comput. Geom.* **3** (1988), 15–47.